Cogent Arts & Humanities

ISSN: 2331-1983 (Online) Journal homepage: www.tandfonline.com/journals/oaah20

Indonesian audio description of paralinguistic contexts in abstract and metaphorical movie scenes

Raden Arief Nugroho, Anisa Larassati , Mangatur Rudolf Nababan & Stephani Diah Pamelasari

To cite this article: Raden Arief Nugroho, Anisa Larassati , Mangatur Rudolf Nababan & Stephani Diah Pamelasari (2026) Indonesian audio description of paralinguistic contexts in abstract and metaphorical movie scenes, Cogent Arts & Humanities, 13:1, 2629087, DOI: 10.1080/23311983.2026.2629087

To link to this article: https://doi.org/10.1080/23311983.2026.2629087

Published online: 23 Feb 2026.

COGENT ARTS & HUMANITIES
2026, VOL. 13, NO. 1, 2629087
https://doi.org/10.1080/23311983.2026.2629087

TRANSLATION STUDIES | RESEARCH ARTICLE

OPEN ACCESS Check for updates

Indonesian audio description of paralinguistic contexts in abstract and metaphorical movie scenes

Raden Arief Nugroho[a], Anisa Larassati[a], Mangatur Rudolf Nababan[b] and Stephani Diah Pamelasari[c]

[a]Faculty of Humanities, English Department, Universitas Dian Nuswantoro, Semarang, Indonesia; [b]Faculty of Cultural Science, English Department, Universitas Sebelas Maret, Surakarta, Indonesia; [c]School of Education, University of Szeged, Szeged, Hungary

ABSTRACT

Linguistic research on audio description (AD) in Indonesian remains very limited, especially on the relationship between AD and the representation of abstract and metaphorical paralinguistic elements in film. We aimed to identify which abstract and metaphorical paralinguistic elements are described by Indonesian AD and what paralinguistic descriptive strategies are used to convey abstract and metaphorical scenes. We obtained the data from four films with Indonesian subtitles. We collected all data using ELAN software, analyzed them using qualitative content analysis, then further examined them through domain, taxonomy, and componential analyses. The results suggest that Indonesian AD places heavy emphasis on visual paralinguistic elements. Contextual aspects such as the use of color and lighting (which signify a film's identity or production house) appear quite often but do not directly contribute to the plot or the main meaning of the story. Description and naming emerged as the most prominent strategies for referring to visual, vocal, and contextual features. We believe that using these descriptive strategies can facilitate comprehensive explanations of abstract and metaphorical scenes while adapting to the time constraints between them.

ARTICLE HISTORY
Received 30 October 2025
Revised 30 December 2025
Accepted 4 February 2026

KEYWORDS
Abstract scenes; audio description; Indonesian; metaphorical scenes; paralinguistic contexts

SUBJECTS
Translation; Interpreting; Language and Linguistics; Audio Description

Introduction

Audio description (AD) facilitates accessibility for people with visual impairments by converting visual information in films into spoken narratives. This allows visually impaired individuals to experience audio-visual elements fully and meaningfully. Thus, people with visual impairments can benefit from technology to understand the paralinguistic context of film (Karpiński, 2012; Marpaung & Gonzalez, 2021; Okorji et al., 2022), which involves nonverbal aspects of meaning such as voice intonation, facial expressions, and gestures (Masharipova, 2024). In movies, paralinguistic details build context through unspoken features (Temirova, 2020). The paralinguistic context cannot be ignored in film because it significantly supports the embedded meaning of a story. Various movies, from horror films (Jianchao et al., 2010) to cartoons (Aini et al., 2023), require paralinguistic context to convey a scene's message and atmosphere, which are not fully revealed through the characters' dialogue. For example, a horror film builds suspense through the actors' movements as well as the situation, setting, and facial expressions.

Figures 1 and 2 show paralinguistic elements in dramas that support the atmosphere, plot, and message of each respective movie. Park (2018) and Niazi et al. (2024) argued that the paralinguistic components present in dramas play a central role in constructing meaning and conveying emotions to the audience. Through aspects such as silhouette lighting, uplighting, chiaroscuro, and deliberate shadows, horror films create visual ambiguity that triggers the audience's imagination of hidden dangers that are not immediately visible. The use of settings such as abandoned houses, dark forests, and remote areas

CONTACT Raden Arief Nugroho arief.nugroho@dsn.dinus.ac.id Faculty of Humanities, English Department, Universitas Dian Nuswantoro, Semarang, Indonesia.

Figure 1. (Abigail et al., 2024).

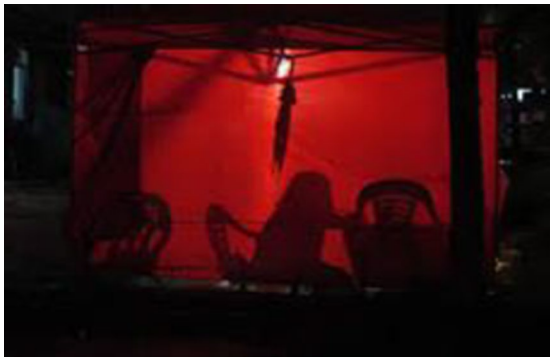

Figure 2. (Ammer, 2020).

further reinforces a sense of isolation and helplessness, giving the impression that no escape or help is available. Furthermore, the mysterious, eerie mood created by fog, dim lighting, and unearthly sounds (such as whispers or high-pitched, unstable music) evokes fear while reflecting the psychological state of the characters and adds depth to the narrative meaning. Thus, the paralinguistic context is crucial for creating an immersive cinematic experience while interpreting hidden thematic messages in horror films.

Movies from a wide range of genres (not just horror) make use of visual elements (such as lighting and setting) as paralinguistic cues; their meanings can also be made accessible to visually impaired audiences (VIA). AD translates visual features (including paralinguistic ones) such as facial expressions, gestures, and emotional ambience into a verbal narrative that can be heard. With the help of AD, VIA can follow the storyline, feel the tension, and understand the psychological state of the characters (Sim, 2017). AD not only improves accessibility but also enriches the user's cinematic experience by providing important information that is not conveyed through dialogue (Lopez et al., 2018). According to Minervino et al. (2018), the most significant challenge in AD is the ability to describe scenes containing abstract and metaphorical concepts. Conceptual representations such as a 'bitter smile' or 'blank stare' are not always fully accepted or understood by VIA (especially those who are congenitally visionless) because the meaning of such expressions is highly dependent on their previous experiences or background information (i.e. semantic references) about abstract and metaphorical matters (Jelec, 2014).

The effectiveness of AD faces several challenges, particularly in terms of integration and narrative quality. Manual AD production is time-consuming and expensive and often does not involve the director or creative team, resulting in descriptions that tend to be generic and less reflective of the film's artistic vision (Wang et al., 2021). This poses a particular challenge in horror, fantasy, and similar genres that rely heavily on the atmosphere, visual intent, and subtle psychological effects. Furthermore, technical limitations (such as the overlap between AD narration and the film's music or sound effects) can diminish the clarity of information, as demonstrated by Lopez and Pauletto (2009) and Park et al. (2025). Hence, future development of AD should consider a more adaptive, aesthetic approach so that an immersive cinematic experience can be enjoyed by all members of the audience, including those with visual impairments.

In Indonesia, improvements in accessibility for the VIA have been increasing (Fauzi & Wicaksana, 2024; Muhammadiah & Selao, 2025; Nurhayati & Kardi, 2024), including the use of AD in a documentary film titled *Sejauh Mata Memandang* (*As Far As The Eye Can See*) (Madani et al., 2023). However, the use of AD in Indonesia remains sporadic and has not yet become an established practice in the film industry. Research on AD—particularly its use in films for the visually impaired—is very limited and has not developed significantly. In many cases, the delivery of visual information for the visually impaired still relies on human prompters to explain the paralinguistic context of Indonesian films (Karolina et al., 2020, 2021; Mahameru et al., 2023).

Global technological advances are beginning to present new opportunities, such as the presence of Indonesian AD on Netflix, which allows VIA to enjoy film content more independently (Sitanggang, 2022). Unfortunately, to date, no research has examined how AD characterizes abstract and metaphorical paralinguistic elements in Netflix films aimed at VIA in Indonesia. Furthermore, we were unable to find research on descriptive strategies used to explain paralinguistic elements. To the best of our knowledge, research on AD centers on the responses of VIA to AD services in their respective countries (Darwish et al., 2022; Hättich & Schweizer, 2020; Tor-Carroggio, 2021; Yang et al., 2025); qualitative analysis of the depictions of scenes and actors, particularly scene changes (Salehi & Razavi, 2023); the portrayal of the body and its physical form (Haider et al., 2025); AD creation, both automatically (Wang et al., 2021) and by professional AD makers and students (Matamala, 2018); AD translation analysis (Bardini, 2020); and linguistic analysis, especially for coherence and cohesion in AD (Giovanni, 2014). Thus, we developed the following research questions:

a. What abstract and metaphorical paralinguistic elements are described by Indonesian AD?
b. How are paralinguistic descriptive strategies applied to Indonesian AD in abstract and metaphorical film scenes?

Literature review

Audio description

As explained in the introduction, AD represents visual elements in spoken language, making them accessible to people with visual impairments. According to the Royal National Institute for Blind People (RNIB), AD functions as a narrator telling a story, providing additional commentary that explains body language, facial expressions, and movement in a film or performance. Through AD, people with visual impairments can understand the storyline because visual information is conveyed through verbal descriptions (Fryer, 2016).

From a historical angle, Fryer (2016) explained that AD was used in Spain in the 1940s and even earlier in England in 1917, when Eleanor, the wife of the famous British painter Sir Ernest Waterlow, described images of visually impaired soldiers. Similar traditions emerged in Japan and other countries, helping audiences understand silent films. AD then expanded to television, with its first broadcast in Japan in 1983. In practice, AD can be broadcast in an open (audible to everyone) or closed (only for certain users via a headset) manner; each way has its pros and cons, particularly regarding the comfort of hearing the sounds of a performance and interacting with other audience members.

Several important aspects must be considered when drafting a plan of visual information to ensure optimal visual access for the visually impaired. First, the types of information that must be described include the appearance, action, position, written text, sound markers or sound sources (indexicals), points of view, and conditions (Piety, 2004). Marzà Ibañez (2010) added that AD should include the setting, time, actions, characters' identity and expression, clothing, professions, lighting, decoration, sound (e.g. difficulty in recognizing voices, song lyrics, foreign languages), and onscreen text such as logos, lyrics, and credits. Second, the presentation mode of AD can be divided into narration and description (Dávila-Montes & Orero, 2014). Narration serves to introduce basic elements such as who, what, when, and where, whereas description provides visual details of how an event occurred. Careful linguistic choices should also be made; excessive use of adverbs, long strings of adjectives, and overly technical or unfamiliar terms should be avoided. Instead, verbs that are concrete, clear, and contextually appropriate are preferred (Salway, 2007).

Third, cultural and linguistic factors play a role in the effectiveness of AD. Strategies that can be used for unfamiliar cultural references include description and naming, explication, and generalization (Maszerowska & Mangiron, 2014; Pedersen, 2005). Additionally, because AD is intended to be an oral medium, descriptors must be wary of potential ambiguities such as homonyms, double *entendres*, or juxtapositions that could create ambiguity. Descriptions should be synchronized with the rhythm of the soundtrack; even sounds, alliteration, and sound symbolism can be used to enhance immersion (Fryer & Freeman, 2014). Finally, the description of color remains important because, in addition to its visual qualities, color has cultural and emotional values that are recognized by many visually impaired people (Piety, 2004).

AD must adapt to the characteristics of film genres because each type of film places different demands on the form and style of delivery. In action films—which are characterized by brief shots, a fast pace, and a predominance of explosions and fights—descriptions must be concise and fit between the intensities of loud sounds (Rasheed & Shah, 2002). In contrast, romance and drama films, which tend to have longer shots and slower pacing, allow AD to be delivered in longer, more poetic sentences, including descriptions of the character's physical attractiveness, although this has the potential to stir up debate (Matamala & Remael, 2015). In comedies, AD must prioritize timing, facial expressions, and gestures that support humor without ruining the punchline. In horror films, the focus is on creating a tense atmosphere through the use of limited space, visual symbols, and ambiguity that is deliberately maintained so that the surprise effect is preserved (Michalewicz, 2015).

Other genres demand distinctive descriptive strategies. For example, large-scale musicals or spectacles require AD that stresses prosody, synchronization with music, and the use of metaphors to recreate stunning visual effects (Matamala & Remael, 2015). Thrillers and documentaries about animals and nature emphasize spatial aspects (such as the distance between characters and threats or predators), although in certain cases, such as in action films like *Mission Impossible* or *Die Hard*, spatial accuracy can be compromised for the sake of narrative flow (Matamala & Remael, 2015). Meanwhile, films with complex genres, such as horror-comedy (e.g. *Shaun of the Dead* or *Happy Death Day*), require descriptive flexibility to switch between suspenseful and humorous nuances (Michalewicz, 2015). Thus, awareness of differences between genres is important in AD to support an optimal viewing experience for the visually impaired.

Paralinguistic context

To provide a complete film-watching experience for the visually impaired, AD provides contextual paralinguistic portrayals of information in scenes. Mukhtarovna (2020) asserted that the paralinguistic context in films facilitates nonverbal communication (such as intonation, gestures, facial expressions, and body language). The paralinguistic context complements spoken conversation and conveys intention, emotions, and information. Paralinguistics functions as a nonverbal channel of meaning that enriches verbal communication through the speakers' emotional nuances, attitudes, and intentions. In a cross-cultural context, Abduazizova (2022) emphasized that intonation, pauses, and vocal stress have significant stylistic, communicative-pragmatic, cognitive, and cultural functions. This means that, even though the words spoken may be the same, the way they are delivered (such as by pausing before answering or increasing vocal volume) can transmit different emotional messages. Audiovisual media (especially AD) play a crucial role in conveying the atmosphere and attitudes of characters that are not captured visually, such as hesitation seen through pauses or tension readable through intonation.

Paralinguistic forms have strong cultural dimensions. Irgin (2017) examined how the use of kinesics and proxemics in cross-cultural communication revealed striking differences between groups based on gender and cultural background, so that gestures familiar in one culture can have different meanings in another. Hence, in practicing AD, it is not sufficient to simply mention the presence of a movement or expression; descriptions should ideally mediate the cultural meanings inherent in a given movement so that VIA can gain an accurate understanding. For example, mentioning a 'shy smile' might be enough for a local audience, but in the global context, the description needs to consider whether the smile indicates embarrassment, politeness, or cultural respect. Thus, paralinguistics are not the only important emotional and interactional markers. Each gesture, facial expression, and vocal intonation does not stand alone but

is always connected to the sociocultural context in which it occurs. In practicing AD, this means that descriptions must go beyond simply mentioning the physical form of the paralinguistic expression; they must also consider the nuances of its meaning.

Paralinguistics includes a variety of nonverbal elements that accompany speech (including vocal qualities such as pitch, volume, and rhythm), gestures, and facial expressions. According to Guyer et al. (2021), vocal features (such as laughter or emotional intonation) significantly improve the understanding of emotional intensity compared to relying solely on words. In the context of AD, this is vital because VIA require vocal cues to capture subtle nuances, such as the trepidation of a sad voice or the hidden emphasis of anger in dialogue. In addition to vocals, visual and contextual paralinguistic elements significantly contribute to comprehension of a scene. Sumekto et al. (2021) used a multimodal analysis tool (EUDIOC Linguistic Annotator (ELAN)) to show that gestures, facial expressions, and vocal loudness significantly enriched students' narrative interactions in storytelling. Thus, in AD practice, descriptions are not sufficient to simply mention objects or actions; they must include visual and vocal dimensions framed by the context of the scene so that the entire audiovisual experience can be fully accessed.

To facilitate the mapping of these elements, McGonigle (2013) categorized the types of paralinguistics that appear in films and how they can be analyzed and translated into AD. In McGonigle's dissertation (2013), he categorized the paralinguistic context into several main forms. Paralinguistic cues include body movements, posture, facial expressions, gaze, movement, and vocal components (such as tone, tempo, volume, and silence) that accompany speech. From a visual perspective, paralinguistics include facial expressions, gestures, eye contact, and the use of space (proxemics), indicating social relations between characters. Paralinguistic vocal elements refer to intonation, stress, volume, laughter, crying, sighs, and other emotional sounds that accompany dialogue. Finally, there are contextual paralinguistic features (namely, factors related to the setting, distance between characters, and temporal aspects such as pauses or instances of silence) that convey emotional or dramatic meaning. These four dimensions complement each other in creating nuances of communication in films and are important to consider in practicing AD so that an emotionally rich cinematic experience remains accessible to viewers with visual impairments.

In many modern films, descriptions relate not only to concrete gestures or expressions but also to abstract and metaphorical scenes, which are often employed to express meanings that are not merely literal but also emotional and conceptual. Classic examples include Christopher Nolan's 2010 film *Inception,* in which the city folds into the sky as a metaphor for the flexibility of dreams, and Ang Lee's 2012 film *Life of Pi*, in which the luminous ocean becomes a symbol of spiritual transcendence. Similar contexts have been found in Indonesian films. For instance, Joko Anwar's 2019 film *Perempuan Tanah Jahanam* (*Impetigore*) uses old homes and dim lighting as metaphors for inherited collective trauma and fear, while Garin Nugroho's 2018 film *Kucumbu Tubuh Indahku* (*I Kissed My Beautiful Body*) presents the dancing body as a symbol of a fluid, vulnerable identity. In line with the notion of audio-vision proposed by Chion (1994), such meanings emerge not only from dialogue but also from the combination of visuals, sound, and rhythm, which builds the symbolic atmosphere. The challenge in audiovisuals is to convey metaphorical nuance without losing the openness of interpretation so that VIA can still grasp the film's aesthetic depth.

Methods

To systematically conduct this research, we developed a methodology that encompasses a qualitative approach, data collection strategies, and analytical procedures tailored to the characteristics of the research object and goals. The following subchapters outline the approach used in the study, the sources and types of data analyzed, and the collection and analysis techniques applied to interpret the findings.

Approach

We relied on a qualitative paradigm because the data are in the form of verbal language—especially AD in voice narratives—and involve humans as the main instruments in the process of interpreting and validating the data (Neuman, 2007). In addition, since we focused on representation in audio descriptions as an experience of meaning, this approach can be classified as phenomenological research (Maykut

& Morehouse, 1997). In several other studies, this approach is called qualitative content analysis (QCA) (Romero-Muñoz, 2023, 2025). Hedenus (2016) referred to this type of research as qualitative audiovisual analysis. QCA is appropriate for research on multimodal data or data in different modes or forms (such as visual, audio, and written data). QCA0020research is used to understand the meaning of qualitative data by grouping pieces of data into specific categories, which are labeled using particular codes based on systematic guidelines. These codes are called codebooks (Schreier, 2012), emphasizing that QCA is a method for understanding the meaning of qualitative data by classifying pieces of data into specific categories labeled with codes that consist of words or phrases, which represent the attributes of the phenomenon being examined. This is also known as open coding (Neuman, 2007).

Sources of the data

Our data sources were Indonesian films on Netflix that offer AD support. Several over-the-top (OTT) streaming services are available in Indonesia, including Netflix, McDonald Hotstar+, Viu, WeTV, Iflix, Vidio, Catchplay, Amazon Prime Video, HBO Go, KlikFilm, Apple TV, Mola TV, and GoPlay (Ramadhan et al., 2023). However, of these platforms, only Netflix has AD capabilities, including Indonesian AD (Agirre-Miguelez et al., 2023). We selected three films—*Tersanjung (Flattered)*, *Kabut Berduri (Borderless Fog)*, and *A Normal Woman*—along with a television series, *Nightmare and Daydreams*, which comprises eight episodes. We chose these samples because, in addition to having AD features, they represent four different genres. *Tersanjung (Flattered)* is a drama, *A Normal Woman* is a psychological drama, *Kabut Berduri (Borderless Fog)* is crime fiction, and *Nightmare and Daydreams* is a supernatural science fiction series, which is rare in Indonesia. The total data used were ten videos, namely three films and seven episodes of the abovementioned series.

Our selection of films was also based on the limited availability of Indonesian films equipped with AD, even though Indonesian is included in the top 20 languages supported by AD, in 17th place (Agirre-Miguelez et al., 2023). The data obtained from these films are multimodal. The term 'multimodal' refers to the way films transmit meaning by combining diverse modes of communication (such as visual, audio, verbal, and paralinguistic aspects) (Holsanova, 2020; Matamala, 2019). Thus, our data include linguistic data in the form of speech in Indonesian AD as well as visual data from movie scenes that form the basis for the emergence of the descriptions, especially those containing multimodal elements such as facial expressions, gestures, and the atmosphere of a given scene.

Data collection and analysis

We used ELAN to gather data. We chose ELAN as the primary tool because it can create complex annotations directly linked to videos, store data in XML format, and provide powerful analysis features for multimodal corpora (Matamala & Villegas, 2016). Similar to Bonsignori and Cappelli (2020) and Yoon (2025), after we collected the films, we first used ELAN to load the videos, create annotation types and tiers according to the research focus, manually insert annotations in each tier, synchronize them with the time of their appearance in the video, and export the data for further analysis. Figure 3 below provides an illustration of how ELAN was employed in the multimodal data collection process.

After gathering the data, we analyzed them using the stages of domain, taxonomy, and componential analyses (Spradley, 1980). In the domain stage, we grouped the AD description strategies and paralinguistic elements into categories based on their similarities in meaning and occurrence in the data. In the taxonomic analysis, we arranged these categories hierarchically to observe the patterns of relationships among description strategies and the types of paralinguistic elements they represented. Finally, in the phase of componential analysis, we examined the differences in characteristics between categories to understand how certain strategies are used to describe the types of elements unique to AD.

Findings

Based on data collection and analysis of two Indonesian Netflix films equipped with AD, we identified several paralinguistic scenes containing abstract and metaphorical representations, along with their descriptive strategies. The details of these scenarios are listed in the following table. For illustration, the numbers indicate the frequency of occurrence; for example, 1 refers to one scene, 2 signifies two scenes, and so on.

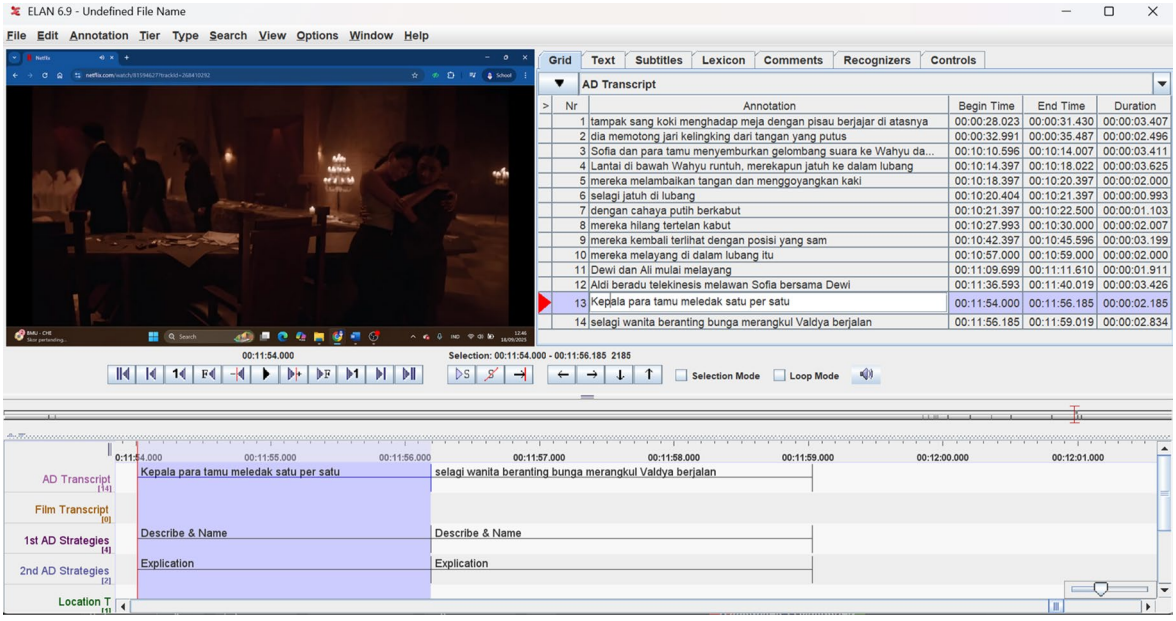

Figure 3. Data codification process with ELAN (Researchers' Own Illustration).

Based on the data in Table 1, there were 703 types of AD in seven scene locations, with four kinds of description strategies. The largest number of types of Indonesian AD was in visual paralinguistic elements, with 320 data points (45,51%), followed by vocal paralinguistic elements with 210 data points (29,87%), and contextual paralinguistic elements with 173 data points (24,60%). In describing abstract and metaphorical scenes, clarification was the only strategy not employed. In contrast, description and naming was the most frequent strategy, with 583 instances (82,9%). Other descriptive strategies included simplification and condensation with 70 findings (10%), generalization with 30 findings (4,3%), and explanation with 20 findings (2.8%).

The following are some examples of the codification and classification results using ELAN, along with their analysis. One of the analyzed scenes from the film *Kabut Berduri (Borderless Fog)* depicts a woman's decapitated head with a cell phone in her mouth (timestamp: 38:12 – 38:40). There is no dialogue between the characters in the scene—only the notification sound of an incoming message on the cell phone and an AD explaining the scene. The first AD reads, '*Ada pesan masuk di ponsel wanita itu* (There's an incoming message on the woman's cell phone.)' This AD uses the description and naming strategy because it clearly refers to the incoming message along with the notification sound on the cell phone, as shown on the screen. In addition, the use of 'that woman' to indicate the head of the corpse is included in the naming strategy. The location or focus of this AD mentions the current state or situation in the scene. Meanwhile, the paralinguistic cues in the AD are contextual elements within the category of scene-setting and objects because they describe the setting of the situation and objects in the film, namely, the head of the woman's corpse and the cell phone in her mouth.

Just as before, the second AD that appears in the scene uses the description and naming strategy; in particular (timestamp: 38:23 – 38:46), '*Muncul tulisan korban nomor 4, Umi* (The message shows that this is Umi, victim number 4)'. The location or focus of this AD is on reading because he reads the writing that appears in the scene. Paralinguistic cues of AD are contextual elements in the form of scene-setting and objects.

The following example from the data was taken from the first episode of the *Nightmare and Daydreams* series, '*Old House*'. Several scenes in this episode depict two alien creatures. There is no official description of the creatures' appearance, but they are portrayed as monsters with baby heads, sharp teeth, and sharp fingers and toes. At first glance, their bodies resemble vaginas. These creatures may symbolize rebirth, as they can transfer energy from a child to an aging mother; the mother is rejuvenated when the child dies. This is certainly difficult to describe in detail because the production team must adjust the film's length and the length of each scene. Hence, one strategy involves generalization, simplification, and condensation.

Table 1. Types, position, and description strategies of Indonesian audio description.

			Audio description strategies										
			Simplification & condensation		Description & naming		Explication		Generalization		Clarification		Total
Paralinguistic cues		Location	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	freq. (%)
Visual Paralinguistic Elements	Gesture	Action	30	4,3%	70	10,0%	10	1,4%	10	1,4%			320 (45,51%)
		Position	20	2,8%									
		State			10	1,4%							
	Posture	Appearance			90	12,8%			10	1,4%			
	Facial Expressions	Appearance			20	2,8%							
	Gaze	Appearance			10	1,4%							
	Movements	Action			40	5,7%							
Vocal Paralinguistic Elements	Intonation	Position	10	1,4%									210 (29,87%)
		Indexical			10	1,4%							
	Stress	Viewpoint			10	1,4%							
	Volume	Appearance			10	1,4%							
	Laughter	Action			10	1,4%							
	Crying	Viewpoint			10	1,4%							
		Action			20	2,8%							
		Viewpoint			10	1,4%							
	Sighs	Action			10	1,4%							
		Viewpoint			10	1,4%							
	Loudness	Action			20	2,8%							
		State			20	2,8%							
	Pitch	State			20	2,8%							
	Silence	Viewpoint			30	4,3%							
		State							10	1,4%			
Contextual Paralinguistic Elements	Colors and Lighting	Appearance			26	3,7%							173 (24,60%)
		Reading			30	4,3%							
		State			20	2,8%							
	Proxemics	Action	10	1,4%									
	Pauses	Appearance					10	1,4%					
	Scene Setting and Objects	Action			10	1,4%							
		Position			20	2,8%							
		Reading			27	3,8%							
		State			10	1,4%							
	Symbolic Environment	Viewpoint			10	1,4%							
Total Occurrences of Each Description Strategies			70	10,0%	583	82,9%	20	2,8%	30	4,3%	0	0,0%	
Overall Audio Description Strategies								703					

One scene depicted in Figure shows the process of the two creatures absorbing life from a child and transferring it to the mother (timestamp: 38:06 – 38:10), which is described as '*Jari makhluk di belakang ibu Bambang mencengkeram kepalanya* (The creature's fingers behind Bambang's mother gripped her head)'. In the film, the long fingers not only grip but also pierce the heads of the two people. In this case, the translator uses generalizations for visual paralinguistic cues in the form of gestures (grabbing and stabbing movements simplified to simple gripping) that demonstrate actions (the actions of the characters in the film, namely, two sharp-handed monsters).

The film also explains that the life-energy transfer process can fail if one of them is unwilling. If this happens, both mother and child will be transformed into exiled beings. Exiled beings are depicted as physically deformed humans with painful facial expressions who walk in a bridge position. When the main character encounters this creature (timestamp: 48:10 – 48:18), AD describes it as '*Makhluk buangan berjalan kayang, mundur ke belakang* (The exiled being walks in a bridge position, backwards)'. However, the translator used simplification and condensation strategies without explaining the physical appearance of the exiled person.

Overall, the paralinguistic elements found in the dramas analyzed in this study are fairly clear in conveying narrative content and supporting VIA's enjoyment of the films. However, in supernatural science fiction genres, such as the *Nightmare and Daydreams* series, many contextual scenes are not described through AD, making it difficult to follow the storyline in the series. One example is in the episode '*The Other Side*', which tells the story of Bandi and his family. Bandi experiences time manipulation when he meets the creatures of Agartha. He has disappeared for several years but feels like he has left his family

for only a few hours. There are several clues from the scene in which Bandi returns to his family: When the camera focuses on Bandi from the front, he still seems young and clean, as when he leaves his family, but when the camera focuses on Bandi from behind, he appears disheveled and dressed in shabby clothes with long hair (timestamp: 28:38 – 30:50). The difference between reality and the timeline experienced by Bandi is not explicitly spelt out, making it difficult for VIA to follow the storyline because they do not obtain contextual clues from the series.

Discussion

Indonesian AD shows a dominant tendency toward describing visual paralinguistic elements, particularly in abstract and metaphorical scenes, whereas vocal and contextual elements appear to comprise a more limited portion. Visual elements are crucial because, according to the International Telecommunications Union (ITU), AD should focus on conveying visual details so that viewers with visual impairments can still understand the context, storyline, and characters (ITU, 2022). According to Vercauteren (2012), visual communication is more implicit than verbal language; thus, the transmitted meaning is often not immediately identifiable. In other words, vocal aspects are generally already present directly in the film's audio and therefore do not always require additional description. Furthermore, in the case of audiovisual aids, visual information can be conveyed immediately after a scene occurs, whereas verbal information can only be communicated sequentially, thus requiring more time to describe it (Remael & Vercauteren, 2010; Vercauteren, 2012; Yan & Luo, 2024). We found that vocal aspects such as intonation, stress, volume, laughter, crying, sighs, loudness, pitch, and silence were still included, especially when the camera was not on the speaker and when these aspects were intended to convey emotions in the film. However, not all vocal elements need to be mentioned, as AD descriptors emphasize only the important features. For example, long silences must be stated so that VIA will not think the film has ended (Sueroj, 2023). In this paragraph, we do not highlight many contextual paralinguistic elements, especially the aspects of color, lighting, scene, and object settings, even though these elements have important roles. As presented in Table 1, the most dominant findings are in the position of colors and lighting in the appearance and reading sections, as well as in the position of scene-setting and objects in the position and reading sections. These findings indicate the existence of a descriptive reading of written texts such as titles, opening credits, closing credits, cast lists, diegetic text, and logos or identities in the film. According to the Audio Description Coalition (ADC, 2009), these texts must still be described according to the provisions, even though they do not make a direct narrative contribution to the plot or overall meaning of the film and generally only appear concisely at the beginning of the film, when the relevant production house is mentioned (Matamala & Orero, 2011; Vercauteren & Remael, 2007).

Indonesian AD predominantly use description and naming, far exceeding those found in other descriptive strategies. As a basic principle of advertising, things that need to be described include sounds with unclear sources but which are relevant to the storyline, logos, titles, and credits; on-screen text such as signs or lyrics; and colors that have visual, emotional, or cultural meanings for visually impaired viewers (ITU, 2022). According to Dávila-Montes and Orero (2014), naming serves to clarify an object on three levels simultaneously: narrative, symbolic, and ideological. For example, in the 2008 film *RocknRolla*, the mention of a car brand such as Bentley in the AD serves to emphasize the image of luxury and glamour attached to the car's character while also highlighting the symbolic value of high social status and an exclusive lifestyle (Dávila-Montes & Orero, 2014). In the 2013 film *Ida*, the English AD adds a touch of glamour to the actress by including the word *classy* in 'she wears a classy coat with a fur collar'; this does not actually appear in the Polish AD (Jankowska et al., 2017). The goal was to help VIA grasp the symbolic meaning of the film. In connection with the limited number of descriptive strategies found beyond description and naming, we believe this condition is closely related to the limited time available to deliver descriptions, especially when representing something abstract and metaphorical. One of the fundamental principles of AD is timing, namely, ensuring that descriptions do not overlap with dialogue (ADC, 2009; ITU, 2022), although Vera (2006) asserted that overlap is tolerated if it does not affect the film's plot. In this regard, we believe that the exploitation strategy has the potential to increase the duration of a description because not all details can be mentioned without wasting the time available (Vercauteren, 2016). Conversely, the use of simplification and condensation strategies, as well as

generalization, reduces the completeness of the meaning conveyed in the film. This view aligns with that of Rai et al. (2010), who emphasized that AD should 'move to details to enhance understanding and appreciation'.

As a form of intersemiotic translation, the descriptive strategies listed in Table 1 can have a dual impact. Some strategies have the potential to improve the quality of AD, whereas others can degrade it. The basic principle of translation highlights that the message in the source language must be transferred to the target language as accurately as possible, without reducing or adding meaning (Nugroho et al., 2021, 2022). Similar to the AD, all visible scenes must be completely and accurately portrayed (ADC 2009; ITU 2022). Considering the timing of scenes, not all scenes can be described completely; thus, according to Peaty et al. (2025), machine describers are not recommended, and human descriptors should be prioritized as they have 'warmth and personality'. However, human descriptors also have weaknesses, one of which is their objectivity (Schaeffer-Lacroix et al., 2023). Rai et al. (2010) revealed that one of the most complex issues in AD research is answering 'What should be described in AD?' Each country or AD association has different policies to address this issue (ITU, 2022). According to Bittner (2010), the ideal elements to mention in AD include shape, movement, color, sound, camera perspective, and supporting information. However, in practice, not all these aspects can be fully communicated owing to time and spatial constraints; hence, the descriptor must prioritize the scene. To select the scene, we suggest that the descriptor refer to the perimeter of translation quality, especially the quality of translation accuracy (Nababan et al., 2012). By definition, an accurate translation is characterized by the precise transfer of meaning from the source language to the target language without causing the slightest distortion of meaning (Nababan, 2008). Referring to the AD data we collected, we obtained an abstract and metaphorical description of a scene from *Nightmare and Daydreams* that states 'his hand clutches his head', whereas the description reads, 'his fingernails press firmly against his head'. This indicates inaccuracy in the description produced by the AD. In this study, we stress that the descriptor has the freedom to determine the priority of the scene being described; however, accuracy must be maintained to avoid multimodal discrepancies between the visual scene and its verbal description.

Conclusion

We successfully formulated a componential pattern, namely the interrelationships between elements in Indonesian AD. The interconnected AD elements included the type of AD, location of the scene, and the descriptive strategy used. The relationships among these three aspects have not yet been reported on in previous studies. Among the three types of AD, visual paralinguistic elements were the most dominant in the Indonesian descriptions. This indicates that gestures, body shape, facial expressions, eye contact, and body movements often determine meaning in storylines that prioritize abstract and metaphorical scenes. Although found in considerable numbers, contextual paralinguistic elements—which manifest in the use of color and lighting to mark the identity of a film or production house—do not directly contribute to the storyline or the film's main meaning. Description and naming strategies are most frequently used to describe visual, vocal, and contextual aspects. This strategy is considered the most appropriate for comprehensively explaining abstract and metaphorical scenes while still adapting to the limited timing between scenes.

This study has some limitations. One is the lack of involvement of people with visual impairment as respondents; therefore, the actual quality of Indonesian advertisements cannot be directly measured. As such, further research should involve more groups of people with visual impairment to evaluate the level of satisfaction with Indonesian advertisements, either through questionnaires or interviews. Furthermore, the findings are expected to provide an important basis for technological development, particularly in software or artificial intelligence (AI)-based systems capable of producing Indonesian advertisements automatically or semi-automatically. Such innovations are expected to accelerate advertising production and improve quality. Finally, our results are expected to encourage stakeholders (including the film industry and streaming service providers) to increase the availability of films with Indonesian advertisements so that the rights of Indonesian people with visual impairment to access inclusive entertainment media can be more equitably fulfilled.

Acknowledgments

We express our sincere gratitude to the Ministry of Higher Education, Science and Technology of Indonesia for funding this research under the *Hibah Penelitian Fundamental Regular* Research Scheme 2025 (Grant Number 127/C3/DT.05.00/PL/2025, 028/LL6/PL/AL.04/2025, 118/F.9-05/UDN-09/2025). We would also like to thank Universitas Dian Nuswantoro and Universitas Sebelas Maret for their continuous institutional support in helping us complete this study.

Author contributions

CRediT: **Raden Arief Nugroho**: Conceptualization, Formal analysis, Funding acquisition, Methodology, Supervision, Writing – original draft; **Anisa Larassati**: Data curation, Project administration, Software, Visualization; **Mangatur Rudolf Nababan**: Investigation, Resources, Validation, Writing – review & editing; **Stephani Diah Pamelasari**: Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was supported by Kementerian Riset, Teknologi dan Pendidikan Tinggi, 127/C3/DT.05.00/PL/2025.

About the authors

Raden Arief Nugroho earned his doctorate from Universitas Sebelas Maret in 2017 and is a lecturer in the Department of English Language Studies, Faculty of Law and Global Studies, Universitas Dian Nuswantoro, Indonesia. He has published and reviewed articles in Scopus-indexed journals and has received multiple competitive research grants. His research focuses on translation studies, English for specific purposes, computational linguistics, and disability studies.

Anisa Larassati is currently pursuing a doctoral degree in linguistics at Universitas Gadjah Mada. She earned her master's degree in linguistics from Universiti Malaya. Her academic interests include discourse analysis, phonetics and phonology, sociolinguistics, translation studies, and forensic linguistics, with a particular focus on language use in legal and investigative contexts.

Prof. Mangatur Nababan is a full-time professor who has been with Sebelas Maret University since 1992. He specializes in Translation Theory, Translation Research Methods, and Critical Reading in Translation. He is actively engaged in government-funded research and vocational training for translators at the Ministry of Cabinet Secretary. His research outputs are directly integrated into his teaching practices. As an academic expert, he frequently serves as a keynote speaker, journal reviewer for Scopus-indexed journals, and thesis supervisor for students at all levels.

Stephani Diah Pamelasari is currently pursuing a doctoral degree at the University of Szeged, Hungary. Her research interests include ESP in higher education and teaching for global competence. She can be contacted at stephani.diah.pamelasari@edu.u-szeged.hu

ORCID

Raden Arief Nugroho http://orcid.org/0000-0002-2495-033X
Mangatur Rudolf Nababan http://orcid.org/0000-0001-9848-4696

Data availability statement

The data supporting the findings of this study are available upon reasonable request from the corresponding author, Dr. Raden Arief Nugroho. Requests for data access should be directed to arief.nugroho@dsn.dinus.ac.id. It is important to note that the data are provided for research purposes only and may be subject to certain restrictions based on privacy, ethical, or other considerations.

References

Abduazizova, D. A. (2022). Paralinguistic Means in Intercultural Communication Communications. *European International Journal of Multidisciplinary Research and Management Studies*, *2*(10), 289–292. https://doi.org/10.55640/eijmrms-02-10-54

Abigail, M. D., Intyaswati, D., & Dharma, N. S. (2024). The Shadow of Horror: A Semiotic Analysis of Cinematic Transitions in Ivanna (2022). *Komunikator*, *16*(2), 208–219. https://doi.org/10.18196/jkm.24223

ADC. (2009). *Standards for audio description and code of professional conduct for describers*. https://share.google/W1XvOuObwzP8rGxxW

Agirre-Miguelez, K., MALópez-Gómez, A., Larrakoetxea-Salgado, N., Arias-Badia, B., Deogracias-Horrillo, M., & Amezaga-Albizu, J. (2023). *Monitoring language diversity and accessibility of streaming platforms in the EU* (Issue December). http://hdl.handle.net/10810/66294

Aini, H., Wahab, A. A., Kiromi, I. H., & Aini, H. (2023). Analysis of nonverbal communication in the cartoon series Shaun the sheep episode we wish a Merry Christmas. *Cendikia: Media Jurnal Ilmiah Pendidikan*, *13*(6), 1059–1069. https://doi.org/10.35335/cendikia.v13i6.3945

Ammer, S. M. E. I. (2020). Content analysis of lighting and color in the embodiment of fear concept in horror movies: A semiotic approach. *Information Sciences Letters*, *9*(2), 135–142. https://doi.org/10.18576/isl/090210

Bardini, F. (2020). Audio description and the translation of film language into words. *Ilha Do Desterro A Journal of English Language, Literatures in English and Cultural Studies*, *73*(1), 273–296. https://doi.org/10.5007/2175-8026.2020v73n1p273

Bittner, H. (2010). *Audio description guidelines-A comparison*. University of Hildesheim. https://www.uni-hildesheim.de/media/_migrated/content_uploads/AD_Guidelines_Comparison_-_Read.pdf

Bonsignori, V., & Cappelli, G. (2020). Specialized and culture-bound knowledge dissemination through spoken tourism discourse: Multimodal strategies in guided tours and documentaries. *Lingue e Linguaggi*, *40*, 213–239. https://doi.org/10.1285/i22390359v40p213

Chion, M. (1994). *Audio-vision: Sound on screen* (C. Gorbman, Ed., 2nd ed.). Columbia University Press. https://doi.org/10.1007/978-1-60327-219-3

Darwish, N. T., Haider, A. S., & Saed, H. (2022). Views of visually-impaired Jordanians on the Netflix Arabic audio description service. *Journal of Southwest Jiaotong University*, *57*(6), 816–828. https://doi.org/10.35741/issn.0258-2724.57.6.71

Dávila-Montes, J., & Orero, P. (2014). Strategies for the audio description of brand names. *CULTUS the Journal of Intercultural Mediation and Communication*, *7*(January), 96–108.

Fauzi, M., & Wicaksana, H. H. (2024). Realizing an inclusive policy for persons with disabilities in indonesia through the art of expressionism gesture photography. *Jurnal Manajemen Pelayanan Publik*, *8*(3), 801–824. https://doi.org/10.24198/jmpp.v8i3.54253

Fryer, L. (2016). An Introduction to audio description: A practical guide. In *An introduction to audio description: A practical guide* (1st ed.). Routledge. https://doi.org/10.4324/9781315707228

Fryer, L., & Freeman, J. (2014). Can you feel what i'm saying? The impact of verbal information on emotion elicitation and presence in people with a visual impairment. *Proceedings of the International Society for Presence Research, April* (pp. 99–107). https://ispr.info/presence-conferences/previous-conferences/ispr-2014/

Giovanni, E. D. (2014). Audio description and textuality. *Parallèles*, *26*(December), 69–83.

Guyer, J. J., Briñol, P., Vaughan-Johnston, T. I., Fabrigar, L. R., Moreno, L., & Petty, R. E. (2021). Paralinguistic features communicated through voice can affect appraisals of confidence and evaluative judgments. *Journal of Nonverbal Behavior*, *45*(4), 479–504. https://doi.org/10.1007/s10919-021-00374-2

Haider, A. S., Tair, S. A., & Dagamseh, M. (2025). A qualitative analysis of body representation in Netflix audio descriptions. *Training, Language and Culture*, *9*(2), 71–83. https://doi.org/10.22363/2521-442X-2025-9-2-71-83

Hättich, A., & Schweizer, M. (2020). I hear what you see: Effects of audio description used in a cinema on immersion and enjoyment in blind and visually impaired people. *British Journal of Visual Impairment*, *38*(3), 284–298. https://doi.org/10.1177/0264619620911429

Hedenus, A. (2016). Embodiment and materialization in "Neutral" Materials: Using Audio-Visual Analysis To Discern Social Representations. *Forum Qualitative Sozialforschung*, *17*(1), 25. https://doi.org/10.17169/fqs-17.1.2404

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

Irgin, P. (2017). Paralinguistics in spoken English: Investigating the use of proxemics and kinesics in an EFL context. *International Journal of Linguistics*, *9*(3), 64. https://doi.org/10.5296/ijl.v9i3.11178

ITU. (2022). *Guidance on Audio Description* (T.701.21). https://www.itu.int/rec/T-REC-T.701.21-202203-I

Jankowska, A., Milc, M., & Fryer, L. (2017). Translating audio description scripts … into English. *SKASE Journal of Translation and Interpretation*, *10*(2), 2–16.

Jelec, A. (2014). *Are abstract concepts like dinosaur feathers? Conceptual metaphor theory and conceptualisation strategies in gesture of blind and visually impaired children* (A. Strugielska, Ed.; 2014th ed.). Uniwersytet im. Adama Mickiewicza w Poznaniu.

Jianchao, W., Bing, L., Hu, W., & Ou, W. (2010). Horror movie scene recognition based on emotional perception. *Proceedings of 2010 IEEE 17th International Conference on Image Processing* (pp. 1489–1492). https://ieeexplore.ieee.org/document/5651515/

Karolina, C. M., Maryani, E., & Sjchro, D. W. (2020). Implikasi Genre Film dan Pemahaman Penonton Film Tuna Netra di "Bioskop Harewos. *ProTVF*, *4*(1), 123. https://doi.org/10.24198/ptvf.v4i1.25035

Karolina, C. M., Maryani, E., & Sjuchro, D. W. (2021). Developing an alternative media for visually impaired audiences: 'Bioskop Harewos' bandung. *Jurnal Studi Komunikasi (Indonesian Journal of Communications Studies)*, *5*(1), 134. https://doi.org/10.25139/jsk.v5i1.2451

Karpiński, M. (2012). The boundaries of language: Dealing with paralinguistic features. *Lingua Posnaniensis*, *54*(2), 37–54. https://doi.org/10.2478/v10122-012-0013-1

Lopez, M. J., & Pauletto, S. (2009). The design of an audio film for the visually impaired. In K. K. Jensen (Ed.), *Proceedings of the 15th International Conference on Auditory Display. May 18–22* (pp. 1–7). Aalborg University. https://www.york.ac.uk/media/tfti/documents/npsC576.pdf

Lopez, M. J., Kearney, G., & Hofstädter, K. (2018). Audio description in the UK: What works, what doesn't, and understanding the need for personalising access. *British Journal of Visual Impairment*, *36*(3), 274–291. https://doi.org/10.1177/0264619618794750

Madani, S., Prasetyo, D., Nyoman Payuyasa, I., & Suardina, N. (2023). Penerapan Audio Description Pada Karya Dokumenter Pendek Sejauh Mata Memandang. *Calaccitra: Jurnal Film Dan Televisi*, *3*(2), 118. https://doi.org/10.59997/cc.v3i2.2934

Mahameru, A. U., Langgeng, P. A., Aminuddin, A. T., Afrida., & W., Andriawan. (2023). Pembisik Virtual sebagai Pemenuhan Kebutuhan Rekreasi Kelompok Difabel Netra dalam Menonton Film. *Jurnal ABDIMAS FISIB*, *1*(2), 9–16.

Marpaung, A. H., & Gonzalez, A. J. (2021). Role of contextual knowledge in paralinguistic affect recognition by humans. *Modélisation et Utilisation du Contexte*, *4*(Special Issue), 1–9. https://doi.org/10.21494/ISTE.OP.2021.0691

Marzà Ibañez, A. (2010). Evaluation criteria and film narrative. A frame to teaching relevance in audio description. *Perspectives*, *18*(3), 143–153. https://doi.org/10.1080/0907676X.2010.485682

Masharipova, V. (2024). Difficulties classifying paralinguistic means in a literary text. *Ижтимоий-гуманитар фанларнинг долзарб муаммолари / Актуальные проблемы социально-гуманитарных наук / Actual Problems of Humanities and Social Sciences*, *4*(10), 175–182. https://doi.org/10.47390/SPR1342V4I10Y2024N28

Maszerowska, A., & Mangiron, C. (2014). Strategies for dealing with cultural references in audio description. In A. Maszerowska, A. Matamala, & P. Orero (Eds.), *Audio description. New perspectives illustrated* (Vol. 112, pp. 159–177). John Benjamins Publishing Company. https://doi.org/10.1075/btl.112.10mas

Matamala, A, Universitat Autònoma de Barcelona. (2018). One short film, different audio descriptions. Analysing the language of audio descriptions created by students and professionals. *Onomázein Revista de Lingüística Filología y Traducción*, *3*(41), 185–207. https://doi.org/10.7764/onomazein.41.04

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

Matamala, A., & Orero, P. (2011). Opening credit sequences: Audio describing films within films. *International Journal of Translation*, *23*(2), 35–58.

Matamala, A., & Remael, A. (2015). Audio-description reloaded: An analysis of visual scenes in 2012 and hero. *Translation Studies*, *8*(1), 63–81. https://doi.org/10.1080/14781700.2014.943678

Matamala, A., & Villegas, M. (2016). Building an audio description multilingual multimodal corpus: The VIW project. In J. Edlund, D. Heylen, & P. Paggio (Eds.), *Multimodal corpora: Computer vision and language processing* (1st ed., pp. 29–32), Department of Translation and Interpreting and East Asia Studies, Transmedia Catalonia.

Maykut, P., & Morehouse, R. (1997). *Beginning qualitative research: A philosophic and practical guide* (P. Maykut & R. Morehouse, Eds.; Vol. 28, Issue 2). The Falmer Press.

McGonigle, F. (2013). *Audio description and semiotics: The translation of films for visually-impaired audiences*. University of Surrey. https://queens.ezp1.qub.ac.uk/login?url=https://search.proquest.com/docview/1651908695?accountid=13374%0Ahttp://resolver.ebscohost.com/openurl?ctx_ver=Z39.88-2004&ctx_enc=info:ofi/enc:UTF-8&rfr_id=info:sid/ProQuest+Dissertations+%26+Theses%3A+UK+%26+Irela

Michalewicz, I. (2015). Is it a monster? Audio describing horror film. *5th Advanced Research Seminar on Audio Description (ARSAD) 2015.* https://grupsderecerca.uab.cat/arsad/sites/grupsderecerca.uab.cat.arsad/files/ARSAD2015_programme_final_0.pdf

Minervino, R. A., Martín, A., Micaela Tavernini, L., & Trench, M. (2018). The understanding of visual metaphors by the congenitally blind. *Frontiers in Psychology*, *9*(JUL), 1242. https://doi.org/10.3389/fpsyg.2018.01242

Muhammadiah, M. J., & Selao, A. (2025). Improving road and sidewalk accessibility for persons with disabilities: Infrastructure challenges and legal compliance in Indonesia. *Advance Sustainable Science Engineering and Technology*, *7*(1),02501022. https://doi.org/10.26877/asset.v7i1.1465

Mukhtarovna, M. D. (2020). The role of paralinguistic units in speech (on the materials of English and Uzbek Languages). *JournalNX - a Multidisciplinary Peer Reviewed Journal*, *6*(11), 165–167.

Nababan, M. R. (2008). Kompetensi Penerjemahan dan Dampaknya pada Kualitas Terjemahan. In *Pidato Pengukuhan Guru Besar Penerjemahan* (Issue April). Universitas Sebelas Maret. https://digilib.uns.ac.id/dokumen/download/22561/NDc2NDE=/Kompetensi-Penerjemahan-Dan-Dampaknya-Pada-Kualitas-Terjemahan-abstrak.pdf

Nababan, M. R., Nuraeni., & A., Sumardiono. (2012). Pengembangan Model Penilaian Kualitas Terjemahan. *Kajian Linguistik Dan Sastra*, *24*(1), 39–57.

Neuman, W. L. (2007). *Basics of social research. Qualitative and quantitative approaches* (2nd ed.). Pearson Education, Inc.

Niazi, M., Shoukat, A., & Saif, A. S. (2024). Illuminating the cinematic palette: A comparative analysis of lightning techniques and their emotional impact in horror films. *Human Nature Journal of Social Sciences*, *5*(4), 143–156. https://doi.org/10.71016/hnjss/1td3wd35

Nugroho, R. A., Muljono, M., & Nababan, M. R. (2021). Visually impaired novice translators in using translation techniques. *Script Journal: Journal of Linguistic and English Teaching*, *6*(5), 103. https://doi.org/10.24903/sj.v6i2.782

Nugroho, R. A., Prananta, A. Y., Septemuryantoro, S. A., & Basari, A. (2022). Translation accuracy in the Indonesian translation of the COVID-19 guidebook: Understanding the relation between medical translation, reception, and risk. In F. M. Federici (Ed.), *Language as a social determinant of health* (pp. 121–145). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-87817-7_5

Nurhayati, S., & Kardi, K. (2024). Digital divide for persons with disabilities related to Indonesian information technology accessibility in law perspective. *ADLIYA: Jurnal Hukum Dan Kemanusiaan*, *18*(2), 105–126. https://doi.org/10.15575/adliya.v18i2.28728

Okorji, R. I., Eze, A. N., & Nwike, C. C. (2022). Paralinguistic communication in Igbo Nollywood subtitles: A case study of two Igbo films. *Arabic Language, Literature, & Culture*, *7*(1), 1–5. https://doi.org/10.11648/j.allc.20220701.11

Park, J., Ye, J., Lee, S., Ka, H. W., & Han, D. (2025 *NarrAD: Automatic Generation of Audio Descriptions for Movies with Rich Narrative Context* [Paper presentation]. Proceedings - 2025 IEEE Winter Conference on Applications of Computer Vision, WACV 2025 (pp. 409–419). https://doi.org/10.1109/WACV61041.2025.00050

Park, M. (2018). *The aesthetics and psychology behind horror films*. Long Island University. https://digitalcommons.liu.edu/post_honors_theses/31/

Peaty, Gwyneth, Locke, Kathryn, Kao, Kai-Ti, Ellis, Katie, Hersinta, (2025). A series of lively impressions: Quality narration and the rise of audio description. *Media International Australia*, *194*(1), 102–114. https://doi.org/10.1177/1329878X231183286

Pedersen, J. (2005 *How is Culture Rendered in Subtitles?* [Paper presentation]. In H. Gerzymisch-Arbogast & S. Nauert (Eds.), *EU-high-level scientific conference series MuTra 2005-challenges of multidimensional translation: Conference proceedings* (pp. 1–18). Marie Curie Euroconferences MuTra. https://www.euroconferences.info/proceedings/2005_Proceedings/2005_Pedersen_Jan.pdf

Piety, P. J. (2004). The language system of audio description: An investigation as a discursive process. *Journal of Visual Impairment & Blindness*, *98*(8), 453–469. https://doi.org/10.1177/0145482X0409800802

Rai, S., Greening, J., & Petré, L. (2010). *A comparative study of audio description guidelines prevalent in different countries*. http://audiodescription.co.uk/uploads/general/RNIB._AD_standards1.pdf

Ramadhan, H. A., Purwaamijaya, B. M., & Guntara, R. G. (2023). Pengaruh user experience terhadap Customer Satisfaction pada Aplikasi Seluler Streaming Vidio. *JTIM: Jurnal Teknologi Informasi Dan Multimedia*, *5*(2), 122–133. https://doi.org/10.35746/jtim.v5i2.367

Rasheed, Z., & Shah, M. (2002). Movie genre classification by exploiting audio-visual features of previews. In R. Kasturi, D. Laurendeau, & C. Suen (Eds.), *Proceedings - International Conference on Pattern Recognition* (Vol. 16, Issue 2, pp. 1086–1089). IEEE. https://doi.org/10.1109/icpr.2002.1048494

Remael, A., & Vercauteren, G. (2010). The translation of recorded audio description from English into Dutch. *Perspectives*, *18*(3), 155–171. https://doi.org/10.1080/0907676X.2010.485684

Romero-Muñoz, A. (2023). Multimodal analysis as a way to operationalise objectivity in audio description: A corpus-based study of Spanish Series on Netflix. *Journal of Audiovisual Translation*, *6*(2), 8–32. https://doi.org/10.47476/jat.v6i2.2023.251

Romero-Muñoz, A. (2025). The audio description script through the lens of multimodality. A qualitative and quantitative analysis of the meaning codes in elite. *Parallèles*, *37*(February), 1–15. https://doi.org/10.17462/para.2025.02.03

Salehi, S., & Razavi, M. S. M. (2023). Indicators of scene changes in Persian and English audio descriptions: A comparative study. *International Journal of Linguistics and Translation Studies*, *4*(4), 128–143. https://doi.org/10.36892/ijlts.v4i4.379

Salway, A. (2007). A corpus-based analysis of the language of audio description. In J. Díaz-Cintas, P. Orero, & A. Remael (Eds.), *Media for all: Subtitling for the deaf, audio description and sign language* (1st ed., pp. 151–174). Rodopi.

Schaeffer-Lacroix, E., Reviers, N., & Di Giovanni, E. (2023). Beyond objectivity in audio description: New practices and perspectives. *Journal of Audiovisual Translation*, *6*(2), 1–7. https://doi.org/10.47476/jat.v6i2.2023.309

Schreier, M. (2012). Qualitative content analysis in practice. In *Qualitative content analysis in practice* (1st ed.). SAGE. https://doi.org/10.4135/9781529682571

Sim, A. (2017). *Recommendations for Chinese audio description guidelines through case-study analysis*. Nanyang Technological University. https://dr.ntu.edu.sg/handle/10356/72390

Sitanggang, P. S. (2022). Strategi Pemasaran Global terhadap Netflix. *ULIL ALBAB: Jurnal Ilmiah Multidisiplin*, *1*(9), 3026–3035.

Spradley, J. P. (1980). *Participant Observation*. Holt, Rinehart, and Winston.

Sueroj, K. (2023). *Strategies for Audio Describing Gestures and Facial Expressions for Visually Impaired Thai Television Audiences* (Issue August) [University of York]. https://etheses.whiterose.ac.uk/id/eprint/33874/1/Sueroj_205047576_CorrectedThesisClean.pdf

Sumekto, D. R., Taufiqulloh, T., Setyawati, H., Hikmah, S., & Ghozali, I. (2021). Understanding the paralinguistic features disclosure depicted in the lecturer's visual modes of writing class instruction. *3L The Southeast Asian Journal of English Language Studies*, *27*(4), 178–197. https://doi.org/10.17576/3L-2021-2704-13

Temirova, F. T. (2020). Importance of paralinguistic features in communication and their theoretical study. *The American Journal of Social Science and Education Innovations*, *2*(09), 111–116. https://doi.org/10.37547/tajssei/Volume02Issue09-15

Tor-Carroggio, I. (2021). The customer is always right: Study on Chinese Persons with sight loss' opinion on audio description. *Disability & Society*, *36*(2), 306–325. https://doi.org/10.1080/09687599.2020.1713727

Vera, J. F. L. (2006 *Translating Audio description Scripts: The Way Forward? - Tentative First Stage Project Results* [Paper presentation]. In M. Caroll, H. Gerzymisch-Arbogast, & S. Nauert (Eds.), MuTra 2006 - Audiovisual Translation Scenarios (pp. 1–10). EU-High-Level Scientific Conference Series. http://www.euroconferences.info/proceedings/2006_Proceedings/2006_Lopez_Vera_Juan_Francisco.pdf

Vercauteren, G. (2012). A narratological approach to content selection in audio description: Towards a strategy for the description of narratological time. *MonTI. Monografías de Traducción e Interpretación*, *4*(4), 207–231. https://doi.org/10.6035/MonTI.2012.4.9

Vercauteren, G. (2016). *A narratological approach to content selection in audio description*. Universiteit Antwerpen. https://repository.uantwerpen.be/docman/irua/4a8d3c/11347.pdf

Vercauteren, G., & Remael, A. (2007). Audio describing the exposition phase of films: Teaching students what to choose. *Trans. Revista de Traductología*, *11*, 76–93. https://doi.org/10.24310/TRANS.2007.v0i11.3099

Wang, Y., Liang, W., Huang, H., Zhang, Y., Li, D., & Yu, L.-F. (2021). *Toward Automatic Audio Description Generation for Accessible Videos* [Paper presentation]. In Y. Kitamura (Ed.), Conference on Human Factors in Computing Systems - Proceedings (Vol. 843, pp. 1–12). Association for Computing Machinery New York, United States. https://doi.org/10.1145/3411764.3445347

Yan, J. X., & Luo, K. (2024). Assessment criteria in audio describer training: An investigation of learner perceptions in a University Interpreting Programme in Hong Kong of China. *Interpreting and Society*, *4*(1), 26–43. https://doi.org/10.1177/27523810231193356

Yang, H., Xie, D., & Li, Y. (2025). The impact of film audio description style on presence in Chinese visually impaired audiences. *Frontiers in Psychology*, *16*(May), 1541539. https://doi.org/10.3389/fpsyg.2025.1541539

Yoon, M. (2025). A cinematographic analysis of English-dubbed Korean films. *Humanities and Social Sciences Communications*, *12*(1), 1–12. https://doi.org/10.1057/s41599-025-04425-7