SKASE Journal of Translation and Interpretation, 2025; 18(2): 93–116
doi: 10.33542/JTI2025-S-6

Translating the visual into extended verbal description

Olga Davis*, Muhammad Ahmed Saeed*,
Dimitris Asimakoulas*, Sabine Braun*

*University of Surrey

Abstract
The article highlights a research gap in the extended (paused) format of audio description (AD) against the backdrop of prolific scholarly output focusing on the mainstream, standard AD format, including numerous user reception studies. Extended AD exists as a peripheral practice and has a largely untapped potential as an alternative method of describing audiovisual content. The project presented in this article aims at offering insights into extended AD from different perspectives in three empirical studies employing a distinct combination of research methods, which include natural language processing, qualitative research (focus groups, interviews) and empirical tools (think-aloud protocol). As such, the study proposes systematic extendibility and practical use of so-called paused commentaries. This, in turn, has the potential to offer a more informative viewing experience to conventional user groups and an opportunity to repurpose AD as assistive commentaries for a wider range of audiences.

Keywords: *accessibility; intersemiotic translation; extended/paused audio description; empirical translation studies; multi-method research design; reception studies.*

1. Introduction

Audio description (AD) is an established professional accessibility practice of making visual information accessible in a verbal format. It enables blind and partially sighted (BPS) audiences to access the visual content of audiovisual media and thus have an inclusive experience. AD is recognised as a modality of audiovisual translation (Fryer 2016). Braun (2008) labels it as a form of intersemiotic translation, following Jakobson's (1959) (modified) typology describing the translation between verbal and non-verbal sign systems.

In standard screen AD, descriptions are inserted into natural pauses in the soundtrack, thus preserving the original length and semiotic make-up of the audiovisual media. However, this approach imposes time boundaries on the description of salient visual elements. Extended AD is an alternative method, that enables describers to pause the source video as necessary, in order to deliver longer descriptions. Furthermore, standard AD is created for a homogeneous audience, whereas audiovisual product users have diverse needs, such as those due to different degrees of visual impairment (Chmiel & Mazur 2022). Extended AD can be tailored to meet such differing needs.

The benefits of standard AD have been established through user-oriented research (e.g. Pettitt et al. 1996; Schmeidler & Kirchner 2001; Fryer & Freeman 2014). Moreover, empirical research with BPS users has demonstrated a positive reception for alternative versions of AD that emphasise aspects such as cinematic language (Fryer & Freeman 2013; Bardini 2020), emotive language (Caro 2016), and the use of specific stylistic devices on syntactic and lexical levels, including different parts of speech, similes and metaphors (Chmiel & Mazur 2022).

However, alternative versions such as the above were created within the same time constraints as standard AD. This leads to a presumption that there might be some information loss to the end user in each of the alternative versions, as some information normally included as per standard AD specification would need to give way to the aspects emphasised in alternative versions.

Prior work in creating authoring tools for producing audio descriptions developed options for adding extended descriptions (Gagnon et al. 2010; Branje & Fels 2012; 3PlayMedia 2019; Pavel et al. 2020; The Smith-Kettlewell Eye Research Institute n.d.). Over the past decade, a substantial body of audiovisual content with extended AD has been created. The YouDescribe platform repository (Pitcher-Cooper 2023) is a prime example. Although extended AD has not yet been subject to systematic evaluation, it has been the subject of critical commentary. For example, when reviewing AD authoring software, Minutella (2022: 337) has contended that "this method is seldom used in real, professionally-produced audio descriptions, since it would disrupt the flow of the audiovisual product", and others have suggested that extended AD is used by non-professional describers (Branje & Fels 2012; Pavel et al. 2020).

The present project is a systematic study of extended AD that actively involves the participation of providers and current users of AD. It investigates whether user-relevant extended narratives can provide a more informative viewing experience for the conventional user group and explores whether it has the potential to enhance the experience of a wider range of audiences.

The specific focus of investigation is on the types of information present in standard and extended AD and the way this information is organised. The project aims at laying the foundations for improving the way information is delivered through extended AD and seeks to answer the following research questions (RQs):

- RQ1: What types of information are present in the standard and extended AD and how is the information organised in the example of a described film trailer?
- RQ2: What types of information would be most user-relevant, and therefore potential candidates to systematically include in an extended version?
- RQ3: How do BPS audiences experience the information delivered by standard and extended AD and what types of additional descriptive elements could enhance this experience?

2. Research design

The research design is a multi-method study enabling triangulation of key stakeholder perspectives. The project is based on the fundamental principles of Greco's (2019: 25) Poietic Design method including "Epistemic Inclusivity where the design draws on the knowledge […], active participation and inclusion of stakeholders". The knowledge that can contribute to our understanding and conceptualisation of extended AD is distributed between three stakeholder groups – advocates of existing practices of extended AD and extended AD products, describers as creators, and BPS people as users of AD. This project integrated consultation with users from its earliest stages.

Phase I of the research is reported in this paper. It comprises three studies: (i) textual analysis of existing extended descriptions, (ii) expert opinion collected through focus groups

with audio describers, and (iii) user perceptions collected through consultations with BPS AD users.

(i) Textual analysis examined the applied practice of extended AD to answer RQ1: A sample transcript of existing standard and extended AD was analysed to compare the organisation of information in each format. Then, the additional information present in extended AD was subject to qualitative textual analysis.

(ii) Expert opinion was sought among practitioners to address RQ2: Views on current AD practice and the possibilities for increasing the use of extended AD were collected through two online focus groups, with four professional audio describers in each group.

(iii) User perceptions were elicited for an initial indication of user perspectives in respect of RQ2 & RQ3: This was collected through combined interview and think-aloud protocol sessions with BPS users who had experience of giving feedback to audio describers or being engaged as accessibility consultants.

Phase II, which will be reported later, builds on the knowledge acquired in Phase I to design and evaluate the usability of samples with extended AD in a user reception study to more fully address RQ3.

3. Dataset of standard and extended audio description

Existing practical applications of extended AD were sought to analyse their underlying narrative scripting techniques and linguistic composition, and to compare them with standard AD. Our selection criteria for the visual material for the qualitative studies were:

(a) a complete work rather than an excerpt from a longer video material to ensure entirety of the narrative;

(b) the existence of parallel texts, standard and extended AD, for each video;

(c) relatively short duration (to be used as a stimulus in our qualitative studies);

(d) moderately fast-paced video material with natural pauses to insert standard explanatory commentary (as opposed to very fast-paced with almost no time for standard AD).

Film trailers offer a concise cinematic experience. They are "a long-standing popular form of promotional narrative (which both sells and tells a reconfigured version of a film narrative)" (Kernan 2004: 2). Trailers present a coherent, post-production narrative that is deemed to be complete, albeit heavily edited by selecting and combining images from the feature with on-screen text and a voice-over narrator. In storytelling terms, trailers guide spectators' attention from an initiating event at the beginning to a suspended outcome at the end. They offer short self-contained video material that could be used for different qualitative studies.

The works by Social Audio Description Collective (SADC)[1] were selected for more in-depth exploration. SADC position themselves as offering an inclusive approach to AD that reflects the diversity of the content and the audience. The audiovisual material available on

[1] About Social Audio Descriptive Collective: https://socialaudiodescription.com/

SADC's YouTube channel[2] are mostly audio described official film trailers/teasers. These were used as the dataset for the study, consisting of forty described videos.

3.1. Description of the dataset

A distinct advantage of SADC's film trailer videos is that they feature standard AD and extended AD of the same audiovisual material recorded back-to-back in a single video file. Both AD versions are written by the same creative team and narrated by the same voice artist for each trailer. Thus, we were able to compare the standard and extended formats of AD for the same trailer, where the decisions about what to include/omit in each AD format were made by the same authoring team. This also excluded the uncontrollable extraneous variable of a difference in the describers' creative intent and confounding differences in the voice delivery of AD.

The forty videos in our dataset were between 3:00 and 14:23 minutes long. All videos followed the same sequence: audio introduction, film trailer with standard AD, introduction of the extended AD, film trailer with extended AD, outro, credits. The intro, outro, and credits durations were disregarded, and the time codes (TCs) of the time-in and time-out for standard and extended descriptions were used to calculate their durations. The durations were converted into seconds to perform calculations summarised in Table 1. The outliers in terms of duration were four teaser videos, which are shorter than the trailers, with video duration of around 3–4 minutes, and some Adventure/Si-Fi trailers with detailed extended descriptions of fictional characters such as the Jurassic World's dinosaurs, with video duration of around 14 minutes. To account for these extremes, the inter-centile range (the range excluding the top and bottom 10% of the distribution) is also given in Table 1.

Table 1: Summary of the observed video durations for the whole dataset

Statistical value	**Standard AD sec**	**Extended AD sec**
Minimum duration of video	36	110
Maximum duration of video	186	672
Range of video duration	150	562
Inter-centile range of video duration	109	305
Mean duration	131	386
Standard deviation	37	129

The mean duration increase in extended AD compared with standard AD was 255 sec (SD = 107 sec), which represents a 202% increase in length (SD 86%). The mean length of an individual extended description was 6 sec (SD = 2 sec), and the mean interval between individual extended descriptions was 3 sec (SD =1 sec).

[2] SADC's YouTube channel https://www.youtube.com/@KittenEarbreak/videos

3.2. Quantitative analysis of the dataset

The Python toolkit pytube was used for downloading the dataset videos from YouTube, providing the ability to interact with them programmatically (Crummy 2021). Pytube allows users to download videos to a specified resolution and format. After downloading, the videos were passed through OpenAI's (2022) Whisper library to transcribe the audio.

The Whisper library was used specifically in this instance because it offers robust models, which are easy to use and have relatively fast processing times. After transcription, Python's Natural Language Toolkit (NLTK) was employed for analysing the text transcripts (Bird et al. 2009). NLTK is a widely used toolkit among researchers, analysts, and data engineers for text analysis. The toolkit is freely available, easy to use, and contains an extensive catalogue of training data gathered from books, articles, and other publications. The text analysis process involves tokenizing the text into individual words and uses NLTK's catalogue to tag each word with its Part of Speech (POS). Each POS tag is then counted to provide insights into the structure and content of the transcripts.

Naturally, based on the automated transcripts across all genres and trailers in the dataset, the word count in the extended AD transcripts is substantially higher than in the standard AD, because in the additional time the extended AD provides more detailed descriptions that add context and clarity. The differences in the word count suggest that there is higher information load in the extended AD, given that the differences in word count occur in relevant lexical categories (nouns, adjectives, etc.). Action film trailers, compared to other genres in the dataset, exhibit the highest counts across these lexical categories in the extended AD transcripts. The contributing factors, in addition to the difference in video duration, are likely to be the fast-paced and visually complex nature of action films, which require more detailed descriptions to effectively capture the intensity and dynamics of the scenes.

For quantitative analysis, it was decided to post-edit the automated transcripts to eliminate occasions of automated speech recognition errors (e.g. *follow was* instead of *followers*, *higher cliffs* in place of *hieroglyphs*). This ensured high accuracy and coherence of the scripts for analysis of POS present in standard and extended AD. Due to the project time restrictions, it was not possible to post-edit all forty transcripts. Therefore, a subset of five videos was selected, with one video representing each of the top genres in the dataset: action, adventure, horror, mystery, and comedy/drama. The genre comparison was drawn between the trailers listed in Table 2.

Table 2: Described film trailers in the analysed subset

#	**Film title**	**Year**	**Genre**
1	Black Panther: Wakanda Forever	2022	Action
2	Jurassic World Dominion	2022	Adventure
3	Scream	2022	Horror
4	Mr Harrigan's Phone	2022	Mystery
5	Cruella	2021	Comedy/Drama

The POS breakdown for these five trailers shows a considerable increase in the number of nouns, adjectives, verbs, and adverbs per minute of the original trailer in the extended compared to the standard AD (see Figure 1). This could be interpreted as evidence of more characters, objects, and scenery being described, and more action not only being mentioned but also described how it happens. It will be explored later through qualitative textual analysis.

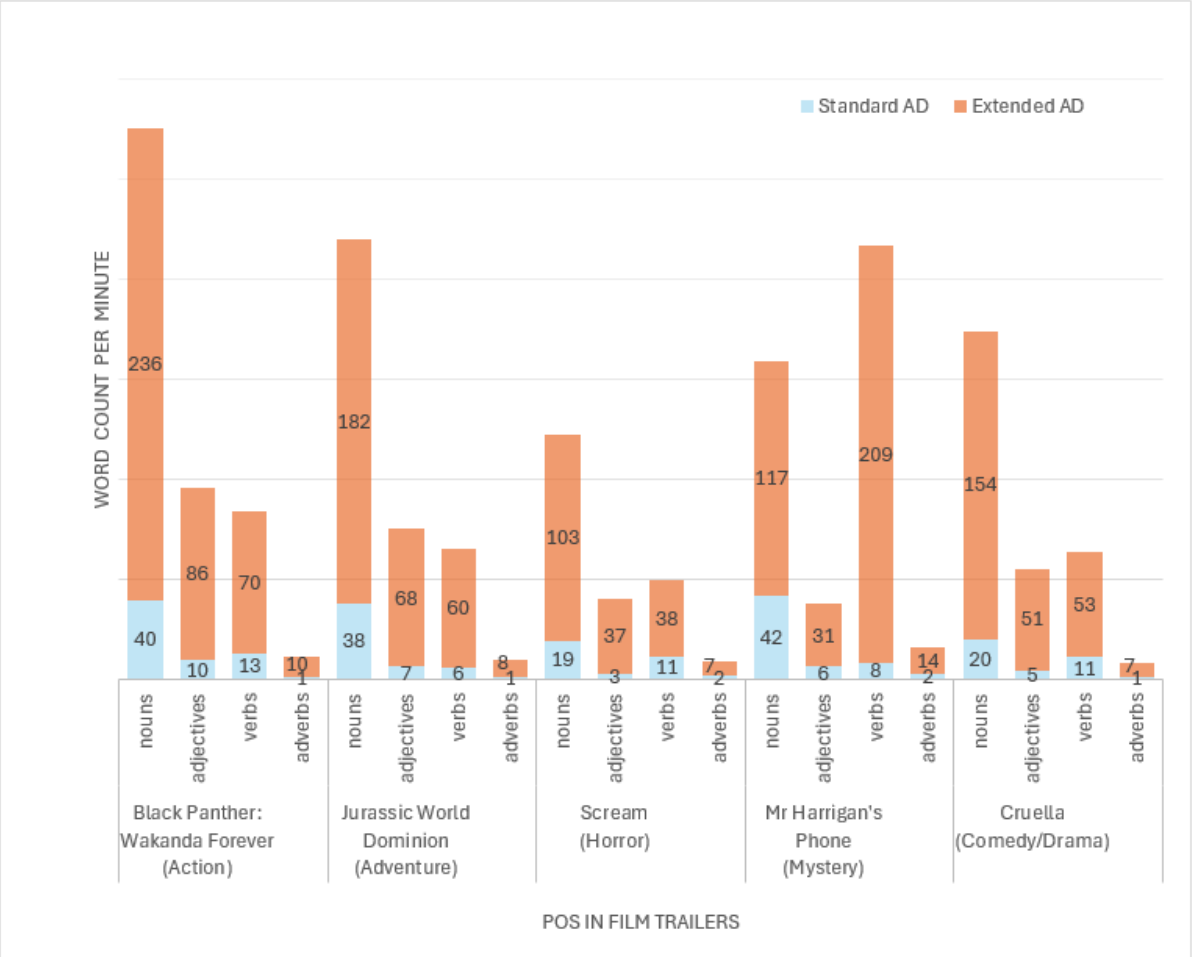

Figure 1: POS count comparison in standard and extended AD for trailers of different genres

4. Linguistic analysis of extended AD

The first step in our linguistic analysis was to observe patterns of information present in the extended AD script beyond those already delivered by standard AD and assign category types. The trailer selected for the linguistic analysis was for *Black Panther: Wakanda Forever* (2022). It was chosen because of its suitability for several studies in this project: Dynamically, it presented a transition from a moderate to fast pace and it satisfied several points we wanted participants of our studies to comment on such as diversity of characters, richness of costumes, variety of settings, presence of fictional objects, and special effects. It was longer than the average videos in our dataset so that we could elicit views on using extended AD in longer audio-visual media. The total runtime of the video is 12:25 min. The trailer runtime with standard AD is 2:10 min and with extended AD is 9:23 min; the remainder being the audio intro/outro and AD producer end credits.

The automated transcript was prepared in MS Word in a table format, allowing for parallel text comparison between the two AD versions, and post-editing of the transcription was performed to correct errors and/or reconstruct missing dialogue and AD utterances. Sound effects present in the original film track were not reconstructed in the transcript because the emphasis of the analysis was on the verbal representation of the visual cues. The automated

timecodes from the video soundtrack were retained in the transcription of the dialogues and the standard AD. Transcribed instances of extended AD were then added in an adjacent column, aligned with the corresponding text of the standard AD and the original dialogues. The transcript was used to perform two types of analysis:

(1) A comparative textual analysis of standard AD and extended AD to identify information present in both formats and how it is organised.

(2) Qualitative textual analysis of extended AD, identifying and classifying the information additional to that present in standard AD. MAXQDA 2022 (VERBI Software 2021) was used to code the different types of information identified.

4.1. Comparative textual analysis

The *Black Panther: Wakanda Forever* (2022) trailer has 27 utterances (as defined in Piety 2004: 458) in the standard AD and 70 in the extended AD. The extended AD includes 7 un-paused (i.e. descriptions inserted in natural pauses, as in standard AD) and 63 paused (extended AD approach) utterances. The comparative textual analysis shows that extended AD generally includes all the information delivered in standard AD (as highlighted in bold in Table 3). Identical wording between the two formats was observed in 4 instances (both, un-paused as in Table 3, TC 0:30, and as part of a paused commentary). While the core content remains consistent (e.g. TC 0:33 and 0:35), the extended AD often splits and interweaves the information into sentences differently (e.g. TC 0:24, 0:33, and 0:35), occasionally altering the wording. Sometimes standard AD information and extra information is intertwined (e.g. Table 3, TC 0:35). On other occasions information is concatenated, where the information from standard AD remains grouped together, often re-phrased, with extra details added as a subordinate part of a sentence (Table 3, TC 0:33) or a separate sentence (Table 3, TC 0:30).

There are parts of extended AD where the author uses the same available time intervals in the original soundtrack to deliver an un-paused AD commentary before pausing to give extra information. They are marked as "not paused" in the transcript (e.g. TC 0:24, 0:30). Interestingly, the author manages to include more detail in some of the un-paused extended AD lines in the same place compared to the corresponding standard AD (e.g. TC 0:33). This is most likely to be due to video editing in the extended AD part. The use of un-paused commentaries in the extended AD did not reveal any obvious patterns, with 5 instances observed throughout the trailer. This suggests that the insertion, layout (sentence structure), and delivery (paused or un-paused) of additional information is not always approached consistently and systematically in the analysed extended AD. Further analysis of the dataset is required for more substantial evidence to draw conclusions.

Whilst the delivery rate of standard and extended AD was not formally measured in this project, it is worth noting that there was no audible difference between the pace in the two AD formats. Besides, there does not seem to be any reason for a faster delivery pace of the extended commentary because extended AD is not time constrained. This, in fact, attracted a comment from an audio describer in the focus group study (see Section 5) that the narrator could have slowed the delivery of the extended AD.

Table 3: Fragment from the *Black Panther: Wakanda Forever* (2022) transcript showing a comparison of how the same information is presented in standard and extended AD

TC	ORIGINAL DIALOGUE	STANDARD AD	EXTENDED AD
0:21		Marvel Studios	Marvel Studios (not paused)
0:24		**Shuri and queen Ramonda lead a Wakandan funeral procession. Shui carries the Black Panther mask.**	**Queen Ramonda holds Shuri's wrist as they lead a funeral procession** with the Twin Towers of the Wakanda palace behind them (not paused) The black Wakandans are all dressed in white loose fabric in traditional African styles and many have white markings on their face while the crowd and followers dance joyously the queen looks around sombrely and Shuri's face is shrouded in a veil as she **holds the Black Panther's mask.** Queen Ramonda is a beautiful older black woman with shiny skin dark eyeshadow white face paint between her brow and on her chin and a white funnel-shaped hat with intricate carvings the Black Panther mask is a sleek matte black full head covering with pointed ears at the top on either side silver eyes and silver accents trim above and below the eyes.
0:30		**In a watery cavern Namor paints an Aztec mural.**	**In a structure inside a dark cavern Namor paints an orange mural (not paused)** It appears to be of a brown skin Artisan working on an ornate suit of armour the mural is surrounded by Aztec hieroglyphs square line drawings of symbols including human and animal heads in the distance are stalactites and pools of water.
0:33		**Wakandans perform a traditional dance dressed in white.**	In front of a stone temple overgrown with trees and vines **Wakandans perform a traditional dance dressed in white** (not paused)
0:35		**In gold jewellery, Namor is a muscular brown man with pointy ears.**	**Namor** continues to work on the mural he dips the brush and paint from a half-sliced conch shell **he is a muscular, brown-skinned man with pointy ears** short black hair beard and moustache he wears ornate **gold and jade gauntlets** and bicep bands a gold septum ring and a rust-coloured cape over his shoulders.
0:41	**Only the most broken people…**		

4.2. Qualitative textual analysis

Our qualitative analysis of extended AD in the same trailer aimed to assess and develop a typology of categories for information additional to that delivered by standard AD. The starting point for developing the coding scheme of information categories were the five categories of information used by Fresno (2016) in her study of criteria to prioritise information in the AD of film characters – age, height/weight, facial features, hair, clothes/other items. These five categories were already tested in her user reception study in terms of recall, comprehension, and the effect of segregated delivery of information on recall. Fresno (2016: 148) demonstrated that "some physical traits of the characters are recalled and recognized better than others". It is instrumental in assessing the potential information value of those categories to the AD users.

The results of the qualitative textual analysis show that the extended AD in the trailer goes beyond the information of standard AD to include, inter alia, description of character traits, setting, and action. All five character traits outlined by Fresno (2016), and adapted for our study, were found in the extended AD of the trailer, which indicates that they could be good candidates for further investigation and developing information typology. For the purposes of our coding scheme, the code 'height/weight' was included in a more general code 'physical appearance' to accommodate descriptions such as *athletic* or *muscular*. The category 'clothes/other items' used by Fresno was divided into two separate codes: 'clothing' and 'objects/items'. From the total of 121 classifiable character-related commentaries, 39 commentaries about the five character traits were present in the extended AD transcript, as shown in Table 4.

Table 4: Coded segments for the five categories of character traits adapted from Fresno (2016) in the extended AD the *Black Panther: Wakanda Forever* trailer transcript

Categories	**No. of coded segments = mentions (% of comments about characters)**
Age	2 (1.6 %)
Height/weight (Physical appearance)	3 (2.5 %)
Facial features	5 (4 %)
Hair	5 (4 %)
Clothing	24 (19.8 %)

Additional codes were also considered, such as ethnicity, following Ofcom's (2021: 7) recommendation that "when describing characters, aspects such as dress, physical characteristics, facial expression, body language, ethnicity and age may be significant". 'Body language' was included under the umbrella code 'action', which was added as a separate category based on the user feedback from the Layered AD study (Schneider et al. 2022). A 'camerawork' category was then added to this list based on the scholarly interest in this aspect of filmic description reported in studies such as cinematic AD (Bardini 2017) and auteur AD (Szarkowska 2013). Bardini (2020) investigated the cinematic style of AD which includes descriptions of camerawork in a user reception study, but this aspect has not been previously explored with audio describers.

Analysis of all additional information included in extended AD (compared to standard AD) indicated that 'action' and 'setting' make up almost half of the additional information: 'action' 37%, 'setting' 12% of all comments. Combined information about character traits (ethnicity, clothing, facial features, hair, age, physical appearance, etc.) made up 31% of all

comments. The remaining 20% of coded information included objects, special effects, references to colour, etc. The number of coded segments (i.e. instances of information present in the extended AD) for 'action' exceeds any other coded information in extended AD. This is consistent with our observations of the POS (see Figure 1) and suggests that the increased use of verbs and adverbs in extended AD is primarily due to the more detailed descriptions of action (e.g. "Queen looks around sombrely"; "warriors expertly climb the exterior"). The higher number of nouns and adjectives stems from more detailed descriptions of characters (e.g. "M'baku is a thick muscular dark-skinned black man with short natural hair on top and a close trim beard"), followed by setting (e.g. "in the cavernous underwater cathedral"). The additional description of action is also consistent with the findings from a user reception study of the Layered AD that provides low, medium, and high level of details (Schneider et al. 2022). The study reported that the participants considered low and medium levels insufficient for the tested scenario of an action scene where they wanted high level of detail described. Camerawork is not mentioned in the described *Black Panther: Wakanda Forever* (2022) trailer, but special effects are present and described in both AD formats. For instance, the standard version mentions "energy surging up her legs", and the corresponding utterance in the extended version reads, "the impact sending a surge of blue energy up her legs". Special effects were categorised with camerawork.

Qualitative textual analysis of the coding categories derived from existing literature helped to refine a shortlist of nine information categories on the basis of the types of information present in extended AD:

1. Physical appearance
2. Age reference
3. Hair
4. Clothing
5. Facial features
6. Ethnicity
7. Cultural references
8. Setting (incl. symbolic aspects of architecture)
9. Camerawork

This list was subsequently used in the consultation studies with audio describers and AD users with the aim of eliciting priority rankings, refining the categories, and adding further categories potentially recommended by the participants (as reported in Sections 5 and 6). 'Action' was not included in this list of categories because the idea was to observe the participants' unprompted comments about an action trailer as an audiovisual stimulus.

5. Expert consultation

This study aimed to obtain the views of AD creation experts on the practice of extended AD. The objectives were to gain a better understanding of the practicalities of AD script writing with regard to selecting relevant information and prioritising visual cues that need to be described under the time-constrained conditions as well as in the extended format of AD. Audio describers brainstormed ideas for the potential content of the extended descriptions. They contributed their professional knowledge and opinions to answering three interrelated questions:

(1) What information is typically selected or, conversely, left out due to time constraints?

(2) Is it possible to categorise what is included or omitted (time-dependent selections) and assign priority ranking?

(3) Is it possible to identify any generalisable patterns and categories to recommend for extended descriptions?

The method for this part of the research was to conduct online focus groups with audio describers. Focus groups have commonly been used in the initial exploratory phase of research design (Wilkinson 1998: 184) as well as in "phenomenological" research when rather little is known about the phenomenon of interest (Stewart et al. 2007: 9). Conducting focus groups online is an established qualitative technique, which is particularly useful when access to busy professionals is required (Brüggen & Willems 2009). Last but not least, it is a method that replicates the conditions under which describers work: individual access to viewing of filmic material on a computer screen (or client platform). The study was approved by the Ethics Committee of the University of Surrey (reference number FASS 22-23 074 EGA, July 2023).

Recruitment of professional audio describers was conducted through charitable and commercial organisations that provide accessibility services to BPS people in the UK. Participant recruitment through both kinds of organisations was designed to attract experts who may represent diverse working practices resulting from internal AD creation guidelines. Gatekeepers of the organisations were sent the researcher's recruitment materials for circulation to audio describers in their existing networks. There was no incentive to participate, and audio describers self-selected.

Two online focus group sessions were conducted, involving a total of eight professional audio describers, only two of whom had experience of creating extended AD. There were four describers in each focus group session to ensure sufficient discussion time in a 1.5–2-hour session. No specialist focus group software was used. The *Black Panther: Wakanda Forever* (2022) film trailer (see Section 4.1.) was used as a visual stimulus in both groups. The trailer was shown via MS Teams (screen share) during the focus group sessions. Respondents watched the trailer at the same time and then provided their comments – prompted by the researcher as moderator. Both sessions followed the same protocol, and the same question guide was used. The uniformity of tasks, runtime, and the number of participants in the two focus groups allowed all participants to have equal opportunities to contribute their professional opinions on the same subject matter, therefore the transcripts from the sessions could be combined and analysed as a single dataset. Data from both focus groups was collected using the MS Teams video recording option with automated transcription. The transcripts were anonymised, and minimal editing was done as was required to correct automated speech recognition errors, spelling, and removing double words to ease readability (Stewart et al. 2007). The content analysis of the transcripts was performed using MAXQDA.

The audio describers were first encouraged to reflect on what in their previous experience they would like to have included into standard AD but could not because of time constraints (Task 1 of the Focus Group). The task produced 30 classifiable responses (an average of four per participant) across a wide range of items, foremost being 'diversity' (23%), 'speaker (identifying who is speaking)' in 17% of all comments, followed by 'physical characteristics' (10%), and 'setting' (10%). The describers felt that the time constraints impinge on their ability to achieve fair representation of the diversity of characters and cultures, as shown in this quote:

> I realise how important it is to describe diversity where it occurs and to have representation for all kinds of cultures and subcultures, but it always feels unfair to me to be describing some characters and not others when there's just not time to devote the same kind of description to everybody. I still think we need to do it. I don't think that means we don't describe anyone's physical characteristics because there's no time to describe everyone's. But in an ideal world, there would be that time. (Participant DF2-3)

Next, the participants were guided to think about the hypothetical scripting of extended AD. The describers were asked whether they saw the omitted content discussed in Task 1 as adding value and whether they would have included it if there was a way to do so (Task 2). The task was designed to elicit spontaneous attitudes towards extended AD. The entire focus group yielded a very mixed response of 80 classifiable comments:

- Negative: 41% (e.g. disapproval and sceptical views);
- Positive: 16% (e.g. approval and views that extended AD works);
- Uncertain: 43% (participants suggested that further research is needed to ask the opinions of BPS users and to investigate how/where extended AD can provide additional benefits beyond standard AD).

As a next step, Task 3, the describers watched the *Black Panther: Wakanda Forever* (2022) film trailer with standard and then extended AD. The describers were asked not to comment on the quality of the standard AD, but to speak about the positive and negative aspects of the extended description compared to the standard AD in this trailer. The subsequent exchange of views generated 33 classifiable comments:

- Negative: 40% (e.g. extended AD gives "too much information" and "but actually, that's not what blind people want.")
- Positive: 24% (e.g. "what did occur to me was that this is a different form of audio description, and I think it would be interesting to embrace")
- Uncertain: 36% (e.g. comments about voice and pace illustrated by the quote below).

> Personally, I would have appreciated Thomas [the AD narrator] doing exactly his normal audio description and then having some awareness that he was in the extra explanatory phase, he could have taken it a bit slower, he could have almost put it into brackets in his head so that you think, oh yes, now we're in this bit and now we're going back to the action I would have found that easier to take. And it's also a way of exploiting the voice which is often underused. (Participant DF2-2)

Then, the discussion revolved around where to insert extended AD, what to include, and what to leave out (Task 4). It uncovered a wide range of opinions from the describers, illustrated in Table 5.

Table 5: Summary of coded comments on where to use extended AD, and what information to include/leave out

Code system	No. of coded comments
Where to use:	
Depends on film	2
Film credits	2
Introduction	2
Sections of the story	5
Title card	1
What to leave out:	
Some objects	3
Some action scenes	3
What to include:	
Action	1
Appearance	4
Camerawork	1
Colour	1
Context for characters	3
Movements	3
Onscreen text/logo	3
Setting	6
Visuals contradicting the audio	1
Vital story changes	1

Describers' overall cautious, sceptical frame of mind about extended AD, evident from the Tasks 3 and 4, can also be traced in their thinking about where to use it. They commented on specific places where to use the extended commentary, for instance, as an introduction, which is what a traditional audio introduction does (Fryer 2016), or over the film credits, rather than thinking of it as being inserted throughout a film.

A broad range of points was brought up in the discussion of what to include in extended AD. The top answers were 'setting' (6 comments) followed by 'appearance' (4 comments). In contrast, describing 'action' was only mentioned once. This is consistent with their answers in Task 1, where they mention 'action' twice when commenting on what they wanted to include but could not.

In the concluding part of the discussion, each describer was asked to name the most salient point in the group discussion on extended AD. Information saturation limits, i.e., that the extended AD contained too much detail that was not possible to remember, was a recurring theme. Another frequently brought up point was the need to ask the views of the BPS users on this subject rather than describers.

At the end of the focus group sessions, the participants were presented with the list of nine information categories identified in the textual analysis (Section 4.2.) and were asked to rank the categories in the order of perceived importance for BPS users. The rankings were collected through a post-session MS Forms questionnaire. It was an individual task for each describer to complete independently, without being influenced by opinions of others. The describers were not exposed to those categories during the focus group sessions so that their

reactions were unprompted by prior labelling. The rankings indicated that the most important category of supplementary information to add to standard AD was physical appearance, followed in the order of frequency by ethnicity, clothing, symbolic aspects of architecture/setting, and cultural references, etc. (see Table 6).

Table 6: Information categories ranking by audio describers from highest to lowest importance

Ranking	Audio describers' ranking of information that is perceived as important for AD users and exceeding standard AD.
1.	Physical appearance
2.	Ethnicity
3.	Clothing
4.	Setting (symbolic aspects of architecture)
5.	Cultural references
6.	Age reference
7.	Hair
8.	Camerawork
9.	Facial features

The describers were also given an opportunity to add categories not listed. A number of categories were mentioned, including 'action and movement', 'identity' and 'diversity', and naming of speakers. They also indicated that unnecessary details and plot spoilers should not be included.

6. BPS user consultation

The next step was to elicit information from BPS users in a consultation study. The comments made by the audio describers (see Section 5) also endorse this approach. The study involved BPS participants who have had experience of giving feedback about AD or acting as accessibility consultants. The objective of this study was to obtain more information about the following: the extent of user involvement in the AD script writing process and typical amendments the users have been able to advise during the AD creation process; user expectations, preferences, and wishes for improving AD; and gaining more information from AD through the use of extended AD. This study gave the participants an opportunity to reflect and give feedback on their experiences of current provision of AD created to standard AD guidelines. It gathered participants' views and helped with identifying the AD users' information needs and preferences. The main question of the study was:

> What types of information might enhance user experience if they were captured in extended AD?

Given the under-explored nature and complexity of the studied phenomenon of extended AD, we employed an exploratory sequential research design using different methods for data collection (Sun 2011). Think-aloud protocol (TAP), a method that allows participants to express their reactions as freely as possible, was used in combination with more specific semi-structured interview questions.

TAP methodology has been used in product usability testing as well as in the humanities and social sciences disciplines, notably in translation process research (e.g. Fraser 1996; Künzli 2009). TAP experiments entail participants engaging with a task or product and verbalising their thoughts in real time while they perform the set task or are exposed to a product. Their verbalisations are recorded, transcribed and analysed from a variety of viewpoints. In Translation Studies, TAPs have traditionally been used when participants translate a text. Jankowska (2021) used TAP as one of the data gathering methods to study AD creation process, with audio describers as participants. Our approach has been to adapt TAP to study the process of AD product reception by BPS users, following successful application of TAPs in this area by Carloni (forthcoming). In addition to increasing our potential for observation of the cognitive processes of AD reception, particularly the points where lack of comprehension occurs, these analyses have a purpose directly related to the researched phenomenon. The verbalisations of what causes problems in comprehension at certain points throughout the viewing process of a described video could help with identification of the types of additional information required at these points to aid user comprehension.

The current study explored user reactions to the two existing formats of AD in the free-thinking part of the session. Prior to the sessions, all participants received information about the project and gave their consent to take part in the study. The study was approved by the Ethics Committee of the University of Surrey (reference number FASS 23-24 036 EGA, May 2024).

Through gatekeeper organisations that provide accessibility services to BPS people in the UK, seven participants were recruited. Participants self-selected. There was no incentive offered in the recruitment materials. However, each participant received a gift voucher after their session to thank them for their time, which they were told about in the concluding part of their session. The individual sessions took place remotely using the videoconferencing platform preferred by the participants (Google Meets for one participant, Zoom for the others), and lasted approximately 1.5 hours each. The audio-only version of the described *Black Panther: Wakanda Forever* (2022) film trailer was used as a stimulus in this study. This was a suitable choice for the level of sight reported by all participants and saved bandwidth by not including the video, resulting in improved internet connection and a better quality of the automated recording/transcription of the sessions.

The sessions with BPS users revealed that none of our participants has been involved in a consultation capacity during the AD script writing process nor as a reviewer of the final script, which would have been an ideal scenario. In practice, the end user involvement was to give retrospective feedback for future reference on something that had already been broadcast. (En passant, only two audio describers reported occasionally involving a BPS user in their AD creation process).

The TAP component of the study was a combination of tasks with retrospective and concurrent verbalisations required from participants. They were first asked to listen to the trailer recorded with standard AD and to give their impressions after listening to the whole trailer. Then, the participants were asked to listen to the same trailer with standard AD again and this time to interrupt at any point where they felt there were issues with AD impacting their user experience such as either insufficient, or excessive, or unnecessary commentary. The researcher paused the trailer at the point of interruption to give the participant time to verbalise what the issue was and what could have helped to improve the description at that point, then resumed the playback until the next interruption.

The mean number of interruptions per session (i.e. per BPS user) was 4.5, with a range of 2–7. Often more than one issue was verbalised during an interruption. The mean number of issues raised per session was 7, with a range of 3–11.

The participants raised a variety of issues (see Table 7). As was expected, there were variations in their needs, which was reflected in them focusing on different issues. On the whole, the participants wanted more descriptions of characters, scenes and actions. The issue of identifying who is speaking was repeatedly brought up by one person (6 times).

Some participants found that the provided standard description of special effects was not detailed enough to interpret what was meant, e.g., mentioning of the energy blast or energy surging up the Black Panther's legs. Similarly, the AD users thought that more detailed descriptions of unfamiliar (genre-specific, fantasy) objects, such as the Black Panther's mask and characters' headdress, could have enhanced their appreciation of the trailer.

In addition, there were comments from two participants about the non-plot-related descriptions: a logo and on-screen text. The standard description reads "Marvel Studios", which was reported as not being clear in terms of whether the actual logo appeared on the screen at that point, or whether it was an on-screen text. The overlayed text advertising the date of the film release during the trailer (not the one announcing the date at the end of the trailer) caused some confusion and interrupted the flow of the dialogue by giving the information that was not related to the scene. The same participant also thought that it was not necessary to read the name of the production studio during the trailer.

Table 7: Issues raised by TAP participants sorted in the descending order of frequency

Issue	**No. of times came up**
Requests for more description of:	
Characters (incl. ethnicity and identity)	9
Scene	8
Speaker identity	8
Action	5
Special effects	4
Objects	3
Clothing	2
Context	1
Technical concerns:	
Warning needed for rapid scene changes	3
Onscreen text/logo unclear or confusing	2
Volume of AD was not enough to hear	2
Acronyms need to be spelled out	1
Subjective Choice of language	1
Voice similarity (between AD and characters)	1

After the TAP tasks with retrospective and concurrent verbalisation about the standard AD, the participants listened to the trailer with extended AD from start to finish and then gave feedback about their experience. This feedback was compared to the feedback they gave after watching the trailer with standard AD first time all the way through. Five participants (except P2 & P4) found the trailer with standard AD too fast, a bit confusing, and hard to follow. By contrast, all but one (P2) reported that the extended AD version was more comprehensive, gave them a better idea of what was going on, and resolved many questions they had whilst listening to the standard version. Excerpts from the participants' comments on each of the AD formats are presented in Table 8.

Table 8: Comparison table of participants' comments on standard and extended AD

Participant	**What did you think of that?**	
	Standard AD format	**Extended AD format**
P1	A lot going on there. I kind of got all kinds of different impressions from that.	That explains it an awful lot more. I found I got a lot more from it and I enjoyed being more connected with the trailer.
P2	It was very fast. I got most of it. To some extent I knew I wasn't gonna get all of it as soon as I worked out was a trailer clip.	Very distracting because it kept cutting out into what the speech was. There was too much stuff. I didn't need to know all that.
P3	So, in places I felt a bit confused. It's made me want to go and watch the film. Which is kind of the idea of trailer.	I really enjoyed that because I got answers to a lot of the questions that I was asking before.
P4	I thought that was fantastic. It was really, well done and he had a lot of material to cover in a really short time.	I think that was fantastic. The extended version actually resolved a lot of the comments that I had about the shorter version.
P5	I couldn't follow very well what was happening.	This was a lot more detailed. I would well prefer this kind of description.
P6	Quite hard for myself to get a real handle of what it's about.	Totally amazing. And it's how you would like it. My only thing that it's. It sounds sort of crazy, to say almost. Overly described. It's so much to take in your minds.
P7	Not a lot, actually. I could have badly done with some sort of context.	My immediate thought is that's really good. Very comprehensive. We've got a better idea of what it is about. No, I like that format very much.

The structured part of the session focused on eliciting AD users' views on how important certain categories of information delivered by extended AD were for them. The typology of information categories developed during the textual analysis and consolidated into nine types (clothing, physical appearance, ethnicity, cultural references, symbolic aspects of architecture/setting, age reference, hair, facial features, and camerawork) was read out to the participants. Although the BPS participants expressed their personal preferences for each category, precise rankings for the whole group proved difficult due to individual differences.

The responses varied more by participant than by category, suggesting that the participant's personal preferences and information needs played an important role. For example, P1 predominantly responded "not too concerned", P2 – "really important", P4 – "pretty important", P5 – "definitely important", and P6 – "quite important". For P7, 'physical appearance' was "important" and 'cultural references' were "quite important", with other categories being of lesser importance, to be mentioned only when pertinent to the storyline. The closest answers to ranking were from P3, who regarded 'physical appearance' as being of "top importance", with 'ethnicity' sitting "alongside" it, while considering other character traits "useful", followed by the remaining categories. After adding that an 'age reference' would be useful, P3 specified that it would be better to have a description of what makes a sighted person think of a character as being of a certain age. When commenting on the 'symbolic aspects of architecture/setting', P3 remarked that they were not important because symbolism explains the meaning and is "getting away from what a description is".

The overall finding was that the perceived importance of information categories varied among users, often influenced by whether they had previously reflected on the categories of information generally present in AD or simply accepted AD as presented. The category of 'physical appearance' emerged as the most universally valued, with participants noting that other character traits from the list were also "helpful" as they contributed to a fuller understanding of a person's appearance. The 'symbolic aspects of architecture/setting' were deemed important when relevant to the plot, especially in the case of iconic landmarks such as the Houses of Parliament in London, which were considered essential to include. The 'camerawork' category received a more mixed response of "yes", "no", and "maybe". The participants felt that the importance of describing camerawork depended on the film's storyline and should be used judiciously.

Similarly to the audio describers, the AD users were asked an open question about what additional information they would have liked to receive through the extended AD. Their responses included background contextual information, more detailed descriptions of scenery and locations, more details about characters, clothing, and objects. Contrary to the expectations in response to an action genre trailer, none of the participants explicitly mentioned 'action'. They were not prompted to do so because the idea was to obtain spontaneous reactions. However, one participant expressed a desire for "visual descriptions of what people were doing". For instance, in the scene described by the AD line, "Wakandans perform a traditional dance", this participant wanted a detailed account of the movements that constituted the dance.

Thinking of the audiovisual media genres that could benefit from extended AD, the participants named a wide range, including science fiction and fantasy "where you're dealing with things that are not regular occurrences in people's normal lives", period dramas where "costumes change all the time", documentaries, horror, and directors' cuts. The application of AD to an action film trailer, provoked differing views from participants: One thought that extended AD worked well for trailers and promotional materials, as it "gives a really good flavour for the film" whilst another participant commented that extended AD might not be suitable "for something that's like a trailer or something that's small and snappy". Some felt it could allow for more detailed descriptions of action and changes of scenery' while others noted that such films tend to have dense dialogue that helps infer what is happening, potentially reducing the need for extended AD.

For all participants, this session was the first time they had experienced extended AD of a video material. One participant said they had previously experienced extended AD once in an animated book reading. When an illustrator was reading the book, animated pictures were

coming up as he was speaking, and the extended commentary was explaining what was happening with the dog and the cat shown in the pictures.

Although the participants expressed some reservations about the extended AD format when asked about its usefulness, the majority of respondents (6 out of 7) thought it was a positive experience, with five participants saying it was "definitely useful". P2 said that extended AD could be useful "on certain occasions". Their reservations were about the extended AD potentially being overwhelming and overemphasising information. The practicalities of using the extended AD in cinema, theatre, or on TV were also a concern. One participant commented that when using DVDs, sighted assistance would normally be needed to turn on the audio description or to select additional commentaries. In other words, any platform used to deliver extended AD needs to be accessible in the first place. Another participant was under the impression that users might need to watch content multiple times or make a judgment call about the level of detail they wanted from the description.

Acknowledging that everyone's preferences may be different, the participants were supportive of extended AD as an optional feature, including when watching content together with sighted people. The participants' views on the genres where the extended AD could work covered a wide range of audiovisual content, which suggests that audiences across all genres could potentially benefit from this format and highlights the need for further research into the development and applicability of extended AD.

7. Discussion

The practices of using extended AD predominantly exist within video-on-demand environments, often without consistent guidelines or standards. While some collections of extended AD examples exist, they have been under-researched. One such collection featuring parallel texts of traditional and extended AD enabled a comparative analysis across all the studies presented in this paper. The common denominator of these studies was to gain a better understanding of the structure of the AD scripts, the types of information delivered to the end user in the extended format in addition to what is delivered by the standard format, and whether it would be possible to assign any user-relevant priority ranking to the information typology.

The textual analyses of the AD scripts revealed certain tendencies about how the information is organised in the extended AD. Firstly, the analysed example of extended AD consistently includes all information present in the corresponding standard AD. As a result, there is no information loss for the user. Secondly, we observed different creative approaches to extending AD (Section 4.1.), such as a) splitting existing standard AD utterances and enriching them with additional details, b) paraphrasing existing elements and adding details, or c) keeping existing elements in a block and supplementing them with additional information. These irregular patterns may be indicative of an ad-hoc application of creative solutions. However, it is also possible that the describers rendered the extended version differently from the standard AD to avoid repetition, because the two versions were designed for sequential watching in a single video playback. For our qualitative studies, we chose to focus on one out of forty trailer transcripts. We plan to extend the linguistic and textual analyses to the whole dataset which will include examples of other genres and may uncover differing compositions of POS and approaches to extending AD.

Among the explored categories of information, 'physical appearance' and other character traits formed over a third of the information delivered by the analysed extended AD script. Furthermore, the 'physical appearance' category was recognised as the top priority for inclusion in AD by the audio describers and was at the top of preferences reported by the AD users. 'Action' also formed more than a third of extended AD descriptions but was not one of the categories put before the audio describers nor AD users, whose unprompted reactions to extended AD were sought in the Phase I studies. Description of action needs further investigation with a larger sample of AD scripts which includes a variety of genres of audiovisual material, and user testing with a larger pool of respondents. Description of movement was not on our list of information categories but was suggested by the audio describers and one of the BPS users as being of potential value to AD users. This suggests that 'movement' could be explored separately or as a sub-category of the 'action' category.

In conducting the focus group study with AD creation experts, group dynamics were identified as a challenge because of the potential emergence of more experienced leaders whose opinions might influence the contributions from less experienced participants. Audio description used to be a niche practice, but considerably more describers have been trained in the past five to ten years following a surge in the demand for AD. Both focus groups included describers who worked since the dawn of AD, as well members of the next generation, who were trained by them. Group dynamics where group leaders emerge is a known issue which was considered when designing the study with BPS participants, initially also planned as a focus group. On reflection, it was decided to conduct individual sessions with the BPS participants and include the TAP component as part of the BPS user study.

The BPS user study had some methodological limitations. Namely, potential reactivity of TAP method, i.e. the impact the interruptions may have on the cognitive process. The studied phenomenon – extended AD – is, by its nature, an experience of interrupted film narrative. The concern here was whether "studying interruptions by means of interruptions" would be the best way to approach data collection. The design of the study attempted to offset this factor by letting the participants experience the full length of the standard AD format uninterrupted first, with a retrospective collection of data before exposing the participants to the TAP assessment during the second watch-through of the standard AD. This was followed by the participants experiencing the extended AD and retrospective collection of feedback.

The stimulus was a film trailer, which already had extra audio elements, such as the voice-over of the trailer's narrator and the industry requirement for the audio describer to read on-screen text. This caused additional confusion for some of the participants in working out who was speaking and whether it was one of the characters in the trailer or a voice over. User viewing habits were not the focus of this study, and familiarity with film trailers as a genre was not among the selection criteria for participants. During the study, the participants were asked about their general experience of using TV and video streaming platforms, but not about prior exposure to film trailers. Asking the participants about their familiarity with film trailers might have offered an explanation for P2's lack of clarity about who was speaking. Regardless, the tendency observed in the trailer was that clarifying verbal cues to identify the speaker were rarely used.

Although TAP as an assessment method was sufficiently accessible to use, going through the list of information categories with AD users by reading the categories out loud to obtain their rankings and subsequent analysis of the findings proved more challenging than expected. A challenging factor here could be whether the verbal delivery of the categories gives BPS participants sufficient time for a prioritisation task, or the prioritisation between the listed

categories requires a more delicate balance resulting in the repetitive answers we observed. In future studies of this kind, it would be advisable to adjust the design of the study to include a more accessible ranking mechanism or tool which may need to be developed in the absence of a ready-made solution. That said, the issue could be more about whether it is appropriate to try to rank a full league table of categories, or to look for groups of users who share common preferences or needs – whose requirements could be met by tailored extended AD.

8. Concluding remarks

This project examined the standard and extended AD formats through three consecutive empirical studies from the perspectives of current practice, practitioners, and users. The studies with audio description practitioners and users elicited participants' views on extended AD. The results revealed mixed views from expert describers, while highlighting their strong interest in being informed and guided by the opinions of BPS users. The response to extended AD from the majority of seven BPS participants was positive, indicating that additional information is welcome by AD users. This highlights the importance of including the voices and perspectives of the BPS users, which may, in the future, empower users to make informed choices about how they consume AD.

While different approaches to extending AD were observed, including enriching existing elements with additional details, paraphrasing, and supplementing information in blocks, the patterns appeared to lack consistency and systematicity in the analysed trailer. Further analysis within a larger data sample is needed to develop a better understanding of common patterns of AD extensions.

The three studies led to an initial typology of information categories that are conveyed in the additional descriptive elements of extended AD, beyond what is delivered by standard AD. This typology may need to be developed further. Triangulation of the results from all three studies will inform the design of a larger user reception study in Phase II, where more systematically designed extended AD samples will be tested.

Notes

1 The studies described in this article were part of a PhD project supported by the Arts and Humanities Research Council UK as part of the Techne Doctoral Training Partnership [Training Grant reference number AH/ R01275X/1]

2 The researcher is grateful to SADC team for responding to the request for the original AD script for the *Black Panther: Wakanda Forever* (2022) trailer. The script was obtained for the sole purpose of verifying our transcription, without the permission of sharing it.

Abbreviations

AD – audio description
BPS – blind and partially sighted
NLTK – Natural Language Toolkit
POS – parts of speech
SADC – Social Audio Description Collective

SD – Standard deviation
TAP – think-aloud protocol
TC – time code

References

3PlayMedia. 2019. How to Edit Audio Description. (*Https://Www.3playmedia.Com/Learn/How-to-Guides/Audio-Description-How-to-Guides/How-to-Edit-Audio-Description/*.) (Accessed 2024-11-28.)

Bardini, Floriane. 2017. Audio description style and the film experience of blind spectators. Design of a reception study. *International Journal of Translation* (19): 49–73. doi: 10.13137/2421-6763/17351.

Bardini, Floriane. 2020. Film language, film emotions and the experience of blind and partially sighted viewers. A reception study. *The Journal of Specialised Translation* (33): 259–80.

Bird, Steven & Klein, Ewan & Loper, Edward. 2009. Natural language processing with Python. Sebastopol, CA: O'Reilly Media.

Branje, Carmen J. & Fels, Deborah I. 2012. LiveDescribe. Can amateur describers create high-quality audio description? *Journal of Visual Impairment & Blindness* 106(3): 154–65. doi: 10.1177/0145482X1210600304.

Braun, Sabine. 2008. Audiodescription research. State of the art and beyond. *Translation Studies in the New Millennium* 6: 14–30.

Brüggen, Elisabeth C. & Willems, Pieter. 2009. A critical comparison of offline focus groups, online focus groups and e-Delphi. *International Journal of Market Research* 51(3): 363–81. doi: 10.2501/S1470785309200608.

Carloni, Arianna. Forthcoming. Watching and listening to dance. Investigating audio description for dance from a textual and audience reception perspective. Guildford: University of Surrey. (Doctoral dissertation.)

Caro, Marina Ramos. 2016. Testing audio narration. The emotional impact of language in audio description. *Perspectives: Studies in Translatology* 24(4): 606–34. doi: 10.1080/0907676X.2015.1120760.

Chmiel, Agnieszka, and Iwona Mazur. 2022. A homogenous or heterogeneous audience? Audio description preferences of persons with congenital blindness, non-congenital blindness and low vision. *Perspectives: Studies in Translation Theory and Practice* 30(3): 552–67. doi: 10.1080/0907676X.2021.1913198.

Crummy, Leonard R. 2021. Beautiful soup documentation. (https://www.crummy.com/software/BeautifulSoup/bs4/doc) (Accessed 2024-11-02.)

Fraser, Janet. 1996. The translator investigated. Learning from translation process analysis. *The Translator* 2(1): 65–79. doi: 10.1080/13556509.1996.10798964.

Fresno, Nazaret. 2016. Carving characters in the mind. A theoretical approach to the reception of characters in audio described films. *Hermeneus* 18: 59–92.

Fryer, Louise. 2016. *An Introduction to Audio Description* (1st edition). Abingdon: Routledge.

Fryer, Louise & Freeman, Jonathan. 2013. Cinematic language and the description of film. Keeping AD users in the frame. *Perspectives: Studies in Translatology* 21(3): 412–26. doi: 10.1080/0907676X.2012.693108.

Fryer, Louise & Freeman, Jonathan. 2014. Can you feel what I'm saying? The impact of verbal information on emotion elicitation and presence in people with a visual impairment. In *Proceedings of the International Society for Presence Research (2014)*. 99–107.

Gagnon, Langis & Chapdelaine, Claude & Byrns, David & Foucher, Samuel & Heritier, Maguelonne & Gupta, V. 2010. A computer-vision-assisted system for videodescription scripting. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition – Workshops*. San Francisco: IEEE. 41–48. doi: 10.1109/CVPRW.2010.5543575.

Greco, Gian Maria. 2019. Accessibility studies. Abuses, misuses and the method of poietic design. In Stephanidis, Constantine (ed.), *HCI International 2019 – Late Breaking Papers. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (vol. 11786). Cham: Springer. 15–27.

Jakobson, Roman. 1959. On linguistic aspects of translation. In Venuti, Lawrence (ed.), *The Translation Studies Reader*. London & New York: Routledge. 126–31.

Jankowska, Anna. 2021. Audio describing films. A first look into the description process. *The Journal of Specialised Translation Issue* 36a: 26–52.

Kernan, Lisa. 2004. *Coming attractions reading american movie trailers* (1st edition). Austin: University of Texas Press.

Künzli, Alexander. 2009. Think-aloud protocols – a useful tool for investigating the linguistic aspect of Translation. *Meta* 54(2): 326–41. doi: 10.7202/037684ar.

Minutella, Vincenza. 2022. Audio description software tools. In Taylor, Christopher & Perego, Elisa (eds.), *The Routledge Handbook of Audio Description*. London: Routledge. 331–52.

Ofcom. 2021. Ofcom's guidelines on the provision of television access services. (https://www.ofcom.org.uk/__data/assets/pdf_file/0025/212776/provision-of-tv-access-services-guidelines.pdf) (Accessed 2023-03-30.)

Open AI 2022. Whisper. A general-purpose speech recognition model. (https://openai.com/blog/whisper/) (Accessed 2024-10-11.)

Pavel, Amy & Reyes, Gabriel & Bigham, Jeffrey P. 2020. Rescribe. Authoring and automatically editing audio descriptions. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. New York, NY: ACM. 747–59.

Pettitt, Bridget & Sharpe, Katharine & Cooper, Steven. 1996. AUDETEL. Enhancing television for visually impaired people. *British Journal of Visual Impairment* 14(2): 48–52. doi: 10.1177/026461969601400202.

Piety, Philip J. 2004. The language system of audio description. An investigation as a discursive process. *Journal of Visual Impairment & Blindness* 98(8): 453–69. doi: 10.1177/0145482X0409800802.

Pitcher-Cooper, Charity. 2023. YouDescribe is a free, web-based platform for adding audio description to YouTube content. (https://www.ski.org/projects/youdescribe) (Accessed 2023-11-07.)

Schmeidler, Emilie & Kirchner, Corinne. 2001. Adding audio description. Does it make a difference? *Journal of Visual Impairment & Blindness* 95(4): 197–212. doi: 10.1177/0145482X0109500402.

Schneider, Remo & Ableitner, Tobias & Zimmermann, Gottfried. 2022. Layered audio descriptions for videos. In Miesenberger, Klaus & Kouroupetroglou, Georgios & Mavrou, Katerina & Manduchi,

Roberto & Covarrubias Rodriguez, Mario & Penáz, Petr (eds.), *Computers Helping People with Special Needs.* Cham: Springer. 51–63. doi: 10.1007/978-3-031-08645-8_7.

Stewart, David & Shamdasani, Prem & Rook, Dennis. 2007. *Focus Groups*. Thousand Oaks: SAGE Publications, Ltd.

Sun, Sanjun. 2011. Think-aloud-based translation process research. Some methodological considerations. *Meta* 56(4): 928–51. doi: 10.7202/1011261ar.

Szarkowska, Agnieszka. 2013. Auteur description. From the director's creative vision to audio description. *Journal of Visual Impairment & Blindness* 107(5): 383–86.

The Smith-Kettlewell Eye Research Institute. n.d. About YouDescribe. (https://Youdescribe.Org/Support/About.) (Accessed 2025-09-10.)

VERBI Software. 2021. MAXQDA 2022.

Wilkinson, Sue. 1998. Focus group methodology. A review. *International Journal of Social Research Methodology* 1(3): 181–203. doi: 10.1080/13645579.1998.10846874.

Olga Davis
Centre for Translation Studies
University of Surrey
Guildford, UK
e-mail: o.davis@surrey.ac.uk

Muhammad Ahmed Saeed
Centre for Translation Studies
University of Surrey
Guildford, UK
e-mail: m.a.saeed@surrey.ac.uk

Dimitris Asimakoulas
Centre for Translation Studies
University of Surrey
Guildford, UK
e-mail: d.asimakoulas@surrey.ac.uk

Sabine Braun
Centre for Translation Studies
University of Surrey
Guildford, UK
e-mail: s.braun@surrey.ac.uk