

## Audio Describing Humour in the Chinese context: A Creative and Collaborative Approach

Zi Ye\*, University of Bristol

Xiaochun Zhang\*\*, University College London

Carol O'Sullivan\*\*\*, University of Bristol

Cristina Marinetti\*\*\*\*, Cardiff University

### ABSTRACT

Audio description (AD) has seen rapid development internationally, including in China. AD principles have long focused on objectivity, and thus have yet to fully address the difficulties of audio describing humorous content, as the objective principle may conflict with comedy's main goal of generating affective responses. This article reflects on an exploratory study practising a creative, collaborative AD approach to prioritise the comedic value of the material. Following the action research approach, the first author, as the audio describer, worked with three external comedy creators and two persons with visual impairments to co-produce AD for a comedy fragment from *Never Say Die* (dir. Yang Song & Chiyu Zhang, 2017). Some creative approaches with which we experimented have considerable potential to give the AD its own humorous verbal charge, but in other cases, proposed creative solutions were rejected or replaced by a more objective approach. This suggests that, rather than moving away from an objective approach, comedy-first AD may expect to balance creativity and objectivity. Our small co-creation project underlines the importance of engaging with creatives as well as users with different sight conditions to create AD for comedy, and highlights the issue of power relations within multidisciplinary AD teams.

### KEYWORDS

Humour, creative audio description, comic effect, co-creation, action research, Chinese context, accessibility, humour translation.

## 1. Introduction

As an accessibility service, audio description (AD) refers to the “description of visual information delivered via an audio channel” (Salway, 2007, p. 151). It has been commonly accepted as a type of translation, semiotically speaking (e.g., Díaz Cintas, 2005, p. 4). It can help people with visual impairments to enjoy audiovisual media. With a large community of 16.91 million individuals with visual impairments, as estimated by the Accessibility Research Association (n.d.), China<sup>1</sup> has witnessed rapid developments in AD, enabling target users to better enjoy audiovisual media (Ye & Jie, 2023). Comedy is an important genre in China, as elsewhere. Of the ten highest-grossing films in the Chinese film market in 2024 (Liu, 2025, p. 16), the top five were all comedies. A survey investigating Chinese users' experiences with AD also shows that comedy is the genre the respondents feel most interested in (Tor-Carroggio, 2020, p. 81).

While comedy is a capstone genre in media creation and consumption, humour presents significant challenges for audio describing due to its clash with an established rule in AD production, namely objectivity. As Soler Gallego and Luque Colmenero (2023, p. 59) put it, objectivity in the AD field refers to “describing what can be seen without interpreting it.” In their study of AD guidelines from France, Germany, Greece,

\* ORCID 0009-0008-6896-6924, e-mail: zi.ye@bristol.ac.uk

\*\* ORCID 0000-0001-6334-6525, e-mail: xiaochun.zhang@ucl.ac.uk

\*\*\* ORCID 0000-0002-1377-7527, e-mail: carol.osullivan@bristol.ac.uk

\*\*\*\* ORCID 0000-0002-5877-280X, e-mail: MarinettiC@cardiff.ac.uk



Spain, the UK and the US, Rai et al. (2010) show that these established AD guidelines generally follow the objective principle. To date, China lacks unified AD guidelines, and different AD groups have their own in-house principles. While some leading groups in China regard objectivity as a characteristic of quality AD, according to Tor-Carroggio (2020), another study has found that Chinese AD is more subjective in practice than corresponding Spanish versions (Liu et al., 2022).

Previous work on the AD of humour mostly focuses on whether or not source humour is retained, and shows that many instances of humour are omitted in AD, probably due to time constraints (Martínez-Sierra, 2009, 2010). Research in recent years has started more specifically to examine the ways in which humour is audio-described. Spanish AD in López Rubio's (2024, p. 331) corpus of three comedies is mostly based on objective principles, which may allow users to perceive that the content is humorous, but do not contribute to the humorous effect directly. Chinese AD for five comedy films in Ye's (2025) corpus is also dominated by the objective rule and emphasises narration over entertainment. Discussions on the r/Blind subreddit about the AD of humour (e.g., [deleted], 2024; theOriginalBlueNinja, 2024) suggest that an objective style cannot always satisfy AD users' expectation for humour enjoyment. Some users on these threads advocated for a more comedy-first AD style. Therefore, the first author decided to experiment with creative AD specifically for humour phenomena. 'Creative AD' or 'creativity' in the context of AD is by no means a new concept and can refer to different types of interventions (see section 2: e.g., Walczak & Fryer, 2017; Luque Colmenero & Soler Gallego, 2020). In this study, creative AD can be understood as making AD funny in its own verbal right, rather than following an ocular-centric idea and trying hard to duplicate the source visual humour. Notably, creative AD does not involve reducing the information load of the AD in favour of comic content; instead, it seeks to describe narrative clues more vividly and funnily, facilitating the audience's humorous experience above and beyond comprehension.

To better practise creative AD, we propose that a collaborative mode would be beneficial. Focusing on the conveyance of humorous effect in audiovisual translation, Mateo and Zabalbeascoa (2019, p. 141) highlight that the translator "must have a sense of humour, or substitutory skills to recognise and produce humour." When the creative AD pivots to put greater emphasis on comic impact, this will make demands on the comic skills of the AD creators. Based on Tor-Carroggio's (2020, pp. 55–58) survey, most Chinese audio describers are volunteers and do not receive systematic training in AD or in a particular media field. More recently, Liu et al.'s survey (2025) shows that Chinese AD provision remains volunteer-driven, though most respondents have received some training. Liu et al. admit that these results may be impacted by sample bias since the respondents are mostly younger age groups from the Beijing area, possibly participating in university-led initiatives. Nevertheless, it is clear that there are training constraints among Chinese audio describers, and they are less likely to possess specialised expertise in humour creation. To compensate for this, a collaborative, 'user-centred' approach in media accessibility can be used. This approach highlights that all relevant agents should engage in the creation of accessibility services to address common gaps (Greco, 2018, p. 206). This user-centred approach has led to the flourishing of a collaborative working mode in the AD field (e.g., Patiniotaki, 2022; Wang, 2024), incorporating different insights and expertise for quality products and a friendly, artistic user experience.

An exploratory study was conducted to practise a creative and collaborative AD approach for conveying humour. In 2024, I<sup>2</sup>, as the researcher and audio describer, brought together an AD co-creation team consisting of two persons with visual impairments and three student comedy creators. Over a period of 40 days, the team worked together to co-create AD for an eleven-minute fragment from *Never Say Die* (羞羞的铁拳, *Xiuxiu de Tiequan*) (dir. Yang Song & Chiyu Zhang, 2017), a popular Chinese comedy film publicly available online. During the co-creation process, I took notes as a self-reflection, focusing on the AD product. After the process, the co-creators also reflected on the collaboration process through a focus group. Section 2 of this article reviews relevant scholarly work. Section 3 introduces the study and situates it as a piece of action research. Sections 4 and 5 discuss key insights about creative AD and collaboration which emerged from our study. Section 6 makes some final remarks and suggests future directions for research.

## 2. Literature review

This section reviews previous creative interventions in the field of AD and the complexities involved in audio describing humour. It then reflects on prior work on co-creation with AD users or creative teams, before introducing Greco's collaborative, user-centred approach and the action research framework, within which this exploratory study falls.

### 2.1. Creativity in AD

While objectivity is the norm in AD, its dominance has long been challenged. The ADLAB Guidelines (Remael et al., 2015) state that AD is based on the interpretation of audio describers, and a balance between subjectivity (personal interpretation) and objectivity should be achieved. Schaeffer-Lacroix et al. (2023, p. 2) argue that “a strict application of the objectivity rule might lead to factual and too technical descriptions [...] that prioritise comprehension over engagement.” Alternative AD approaches have emerged, which depart from the orthodox objective AD. Fels et al. (2006b) invited the original creative team of the animated comedy *Odd Job Jack* to create a first-person AD style. A follow-up reception study (Fels et al., 2006a) shows that this innovative AD style is more engaging and entertaining than the conventional third-person, objective version. Walczak and Fryer (2017) develop ‘creative description’ to include film terminology and subjective descriptions of the characters, their actions, and scenes crucial to the plot, promoting a stronger sense of presence. A metaphorical AD style is adopted by Luque Colmenero and Soler Gallego (2020) to design guided tours for art museums, where they use metaphors for both representational and abstract art to create subjective connections. Such creative interventions pursue “an artistic, imaginative or creative contribution to the audiovisual text that can elicit a new audience experience” (Romero-Fresco & Chaume, 2022, p. 77). These previous experiments were a source of inspiration for the creative AD in this project, which seeks to exploit opportunities to bring to the fore the comedic dimension of the source material.

The account above should not be taken as presenting comic AD in a simplistic, binary relationship with the objective approach. This relationship is complicated by a number of factors. Martínez-Sierra (2009) proposes that AD should provide contextual information to ensure the process of inference and humour production, but

acknowledges that time constraints may be an obstacle. His case study of *I want Candy* shows that, due to the lack of time, two-fifths of visual humorous elements are omitted in its English AD version (Martínez-Sierra, 2010). But this calculation may overstate the loss of humour in AD. More recently, Martínez-Sierra (2020) has argued that in a multimodal context, missing certain visual elements may not severely impact the overall experience, as other elements can compensate. This view is supported by Helal Birjandi et al.'s (2025) recent analysis of the Persian AD for three dialogue-free animated features. Two reception studies (Martínez-Sierra, 2021; López Rubio, 2024) designed to evaluate the effectiveness of audio-described humour found that sighted viewers and viewers with visual impairments rate their experience of the comedies similarly, even if their ratings are based on different multimodal combinations of meaning-making resources. Based on the above, the AD of humour is part of a complex multimodal negotiation and may not be dependent on a single AD approach.

## 2.2. AD co-creation

A co-creative working mode has been adopted in several studies involving people with visual impairments, thus empowering them in the AD field. Based on what she terms 'participatory accessibility,' Di Giovanni (2018) reports on an AD co-creation with blind, partially sighted, and non-blind children for an opera event. She found that children with visual impairments are more active in collaboration and contribute through multiple senses (p. 167). Hirvonen et al. (2023, p. 26) highlight 'user as maker' in AD co-creation. Observing team AD practices for nine productions for films or TV, they found that blind user representatives went beyond providing feedback to actively participate in solving AD problems and making final decisions (p. 44). Embracing the concept of 'Blindness Gain,' Eardley et al. (2024) challenge the ocular-centric opinion that museum AD needs to be created by sighted professionals. They developed the Workshop for Inclusive Co-created Audio Description model to engage blind, partially blind, and sighted individuals for their different ways of understanding and experiencing museum collections.

Another approach is to co-create AD with the original creative teams behind the content, ensuring the AD serves as a creative extension of the original show. Udo and Fels (2009) led an AD co-creation involving the directors and audio describers for a Toronto production of *Hamlet*. According to them, the engagement of the creative team makes the AD cater more to the feel of the scene. Fryer (2018) advocates for the 'Integrated AD Approach,' which includes AD provision in the creative process and requires collaboration with the artistic team. Working with blind or partially blind actors from Unscene Suffolk, Fryer found that the AD they created was more creative and less neutral. Patiniotaki (2022) also employed this integrated AD approach when producing AD for the contemporary dance performance *A Clear Midnight*. Working with nine dancers, the choreographer and the dramaturgy consultant, a creative AD version was generated featuring "extremely expressive, yet informative vocabulary and expressions" (p. 155).

Greco (2018) introduces a collaborative, user-centred approach to fill the maker-expert-user gap. According to his approach, the design process should be a co-construction where all relevant agents contribute in an equal manner. Drawing on this framework, Wang (2024) established a collaborative team to co-create a voice-driven interactive smart guide for museum accessibility. The team consisted of three

functional roles: blind and partially sighted visitors, audio descriptive guide, and internal stakeholders (e.g., museum management, smart guide editor). The final product was effective in improving the experience of visitors with visual impairments, supporting museum management and reducing unnecessary investment. This project also follows Greco's triangular model (Figure 1), involving comedy creators, users with visual impairments, and an AD volunteer in the co-production process.

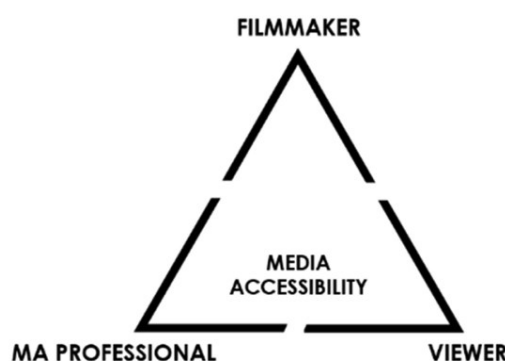


Figure 1. Greco's triangular model, adapted by Branson (2018) (Romero-Fresco, 2019, p. 5).

### 2.3. Action research approach

This research falls under the heading of action research, which is “a democratic process concerned with developing practical knowing in the pursuit of worthwhile human purposes” (Reason & Bradbury, 2006, p. 1). Kurt Lewin (1946, p. 38) describes this approach as a cyclical process involving “planning, action, and fact-finding about the result of the action.” Planning starts with examining the general idea carefully; action is about executing the overall plan; fact-finding mainly serves as the evaluation of action. Kemmis and McTaggart (1982, p. 8) develop this sequence of steps into the self-reflective circle of ‘plan–action–observe–reflect,’ identifying ‘observe’ as a separate activity followed by reflecting on the purpose of the plan and the consequences of the action. This circle is in fact a spiral, and, in reality, the process can be more flexible, with stages overlapping.

Lewin (1946, p. 46) believes that action research can be a remedy to address discrimination faced by minority groups and facilitate “independence, equality, and cooperation.” Action research has been used in disability-related studies (e.g., Bessaha et al., 2020; Reviers & Hanouille, 2023). It has also been invoked in audiovisual translation studies. According to Bogucki (2013, p. 115), “the routes of theoreticians and practitioners exploring the nature of translation tend to fork too often,” and action research can bridge such a gap by addressing practical problems. Di Giovanni (2018) steeps her participatory accessibility project in the action research framework, developing materials use and workflow plans, inviting blind, partially sighted, and non-blind children to take part in AD creation, observing the collaboration process, and collecting reflections for further steps, an approach that the present study emulates.

### 3. Project outline

This study took place within the UKRI-AHRC International Placement Scheme (IPS) from October to December 2024. It was hosted by the Shanghai Theatre Academy

(STA), a prestigious cultural institution in China, which plays a leading role in Chinese visual and performing arts and in training creatives. The main phases of this project are in Figure 2 below.

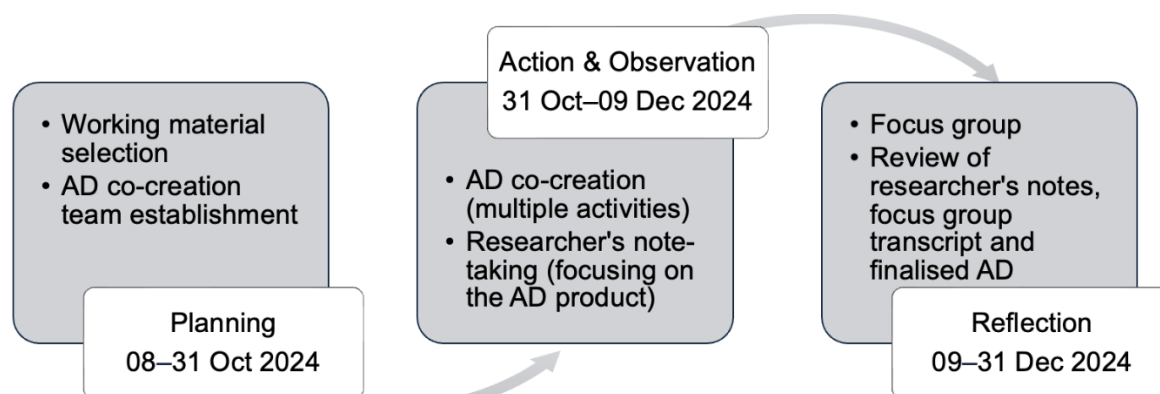


Figure 2. Phases of the study.

The first phase involves material selection and team establishment. We initially expected to approach a comedy filmmaking team from STA or beyond, so that we could request working material from them for AD practice, and invite them to co-create the AD. However, applying for materials use turned out to involve complicated administrative procedures and copyright problems, which could not be achieved within the three-month project period. In the end, we decided to use an eleven-minute segment (00:45:33–00:57:10) from a Chinese comedy film, *Never Say Die*, which is publicly available online. In this film, the male and female protagonists swap bodies accidentally. In our chosen segment, the two body-swapped characters learn boxing in a remote martial arts school. This material suits the project as it is a broad popular comedy (the second most popular domestic film at the 2017 Chinese box office) and contains extensive visual humour elements, many of which are presented without dialogue, thus requiring AD.

The next step was to establish a co-creation team. Drawing on Greco's (2018) user-centred approach and similar studies (e.g., Wang, 2024), the team needed to consist of three roles: 1) the audio describer, 2) comedy artists, and 3) user representatives with visual impairments. All the roles were 'co-creators,' and they should fully engage in the whole process of producing the AD. Given my hands-on experience and knowledge, I took the role of the audio describer. Since approaching the original creative team behind this big-budget film would have involved additional uncertainties and complexities, we opted to invite external creative artists. Via my Shanghai Theatre Academy advisors' networks, I contacted STA's Ju Guang Drama Caring Public Welfare Service Team, which has pioneered audio-described theatre performances in China in collaboration with accessibility groups. Three student comedy artists from Ju Guang joined the project, and they all had practical expertise in comedy creation. Through personal contacts, two user representatives with acquired blindness also joined our team, one of whom had congenital low vision prior to becoming blind. They are passionate about promoting Chinese AD development and have created many ADs. The profiles of team members are shown in Table 1, in an order based on their roles and alphabetical order of surnames:

| <b>Roles</b>         | <b>Name/Pseudonym</b>  | <b>Gender</b> | <b>Age</b> | <b>Sight condition</b>                  |
|----------------------|------------------------|---------------|------------|-----------------------------------------|
| Audio describer      | [Zi Ye]                | [Female]      | [18–30]    | [Sighted]                               |
| Comedy creators      | Jinglan Su             | Female        | 18–30      | [Sighted]                               |
|                      | Yingyi Tang            | Female        | 18–30      | [Sighted]                               |
|                      | Xiaoqi Yu              | Female        | 18–30      | [Sighted]                               |
| User representatives | Qinghan Qi (Pseudonym) | Male          | 18–30      | Blind, Acquired (congenital low vision) |
|                      | Junyi Zhu              | Male          | 31–40      | Blind, Acquired                         |

Table 1. Profiles of the co-creation team.

The action phase to co-create the AD lasted from 31 October to 9 December 2024. After understanding the concept and working direction of creative AD, co-creators were not required to follow specific AD guidelines, but were allowed more flexibility for creative solutions. The main collaboration activities during the action phase are in Table 2, together with corresponding objectives, locations and dates. For each part of our AD script, I, as the audio describer, also voiced it and generated corresponding audio files through Movavi Video Editor software. For reasons of copyright, the generated audio files were used only within the team to facilitate the co-creation process. The final AD script can be found in the Appendix.

| <b>Activities</b>                                             | <b>Objectives</b>                                                                                                                                                                          | <b>Locations</b>       | <b>Dates</b>                  |
|---------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|-------------------------------|
| Communication through WeChat group                            | <ul style="list-style-type: none"> <li>Familiarisation with the project and the source comedy fragment (The fragment was divided into three events)</li> </ul>                             | Wechat group           | Flexible (31 Oct–04 Nov 2024) |
| An online session                                             | <ul style="list-style-type: none"> <li>Icebreaking</li> <li>Discussion on creative AD and working flow</li> <li>Source fragment analysis</li> </ul>                                        | Wechat group           | 5 Nov 2024                    |
| Frequent communication through WeChat group/in-person meeting | <ul style="list-style-type: none"> <li>Co-creation of AD for event 1 (Analysis of AD problems, initial AD draft by the audio describer, rounds of discussion and revisions)</li> </ul>     | Wechat group /Flexible | Flexible (06–14 Nov 2024)     |
| An in-person workshop                                         | <ul style="list-style-type: none"> <li>A focused working period and the establishment of mutual trust</li> <li>Co-creation for events 2 &amp; 3</li> </ul>                                 | STA campus             | 24 Nov 2024                   |
| Frequent communication through WeChat group/in-person meeting | <ul style="list-style-type: none"> <li>Continued co-creation for event 2 &amp; 3;</li> <li>Further revision for event 1</li> <li>Integration and finalisation of the AD product</li> </ul> | Wechat group /Flexible | Flexible (28 Nov–09 Dec 2024) |

Table 2. Main co-creation activities.

During the project, one co-creator shared a link to an existing AD version of this fragment created by another AD group. To avoid influencing our creative process, I asked the team not to consult this version.

The process of making AD decisions was observed and included in the researcher's notes, with my reflections. Other co-creators also reflected on the collaborative process through an online focus group, which I moderated. The 90-minute focus group took place on Zoom on 12 December 2024 and was audio recorded.

I took a dual researcher-practitioner role throughout the project, adopting an “insider perspective” (Reviere & Hanouille, 2023, p. 102). Methodological challenges of this kind of research may include difficulties in distinguishing the researcher's subjective experience from the object of the study (Marinetti & Rose, 2013, p. 170) or lack of naturalness of data (Hirvonen et al., 2023, p. 31). During the co-creation process, I alternated between roles: organising activities as the investigator, while joining in activities as the audio describer. To mitigate methodological challenges, decisions could only be made through team discussion and agreement by all co-creators. In my notes, I also reflected on different opinions about specific problems, and on how the final AD product was shaped, with or without my intervention. For the focus group, I only worked as the moderator, refraining from expressing any opinions. Nevertheless, the findings of this project can only be interpreted within the specific context of these interactions, constraints, and relationships.

After compiling the researcher's notes and the focus group transcript, the first author analysed the data inductively, drawing where necessary on quotations from the final AD product. The researcher's notes mainly cover the evolution of the AD script from draft to final version as well as the researcher's reflections about the co-creation process. The focus group transcript contains co-creators' reflections on the collaboration process. Free text coding was conducted to identify a range of themes, which we do not have space to discuss in full here. The key insights are discussed in sections 4 and 5 respectively.

#### **4. Creative interventions for comic effect**

Three approaches to comic impact in AD emerge from the data: verbalising original humour mechanisms, adding new verbal humour cues, and avoiding stereotypes and sensitive expressions. The first approach occurs 17 times in the researcher's notes, reflecting 16 uses in the final AD script. The other two approaches are less frequent, occurring five times each in the final AD script, but we argue that the two are still significant and worthy of being independent categories, as they offer new perspectives which can enrich our understanding of the functioning of AD in relation to humour.

##### **4.1. Verbalising original humour mechanisms**

Visual humour is structured through various icons, signs, and symbols which cannot be retained in AD. However, the mechanisms underlying visual humour, such as twist and incongruity, are not restricted to visual repertoires and can be represented in AD through language resources.

For example, in *Never Say Die*, the male boxer Disheng Ai and the female journalist Xiao Ma swap bodies in a freak lightning incident. In a later scene, Disheng gets too drunk to hide that ‘he’ is really Xiao and lies seductively on the stone steps, a behaviour more commonly associated with females in the Chinese cultural context. Such a gender swap leads to an incongruity between the character’s male identity and the stereotypically feminine behaviour, deviating from conventional gender norms and creating comedic value. Martínez-Sierra (2020, p. 189) highlights that being denied access to visual features of a joke does not necessarily result in a diminished overall humour experience, but in this example, the incongruity is primarily visual and the loss of these elements may significantly reduce the effectiveness of the underlying humour mechanism. Based on the co-creation team’s analysis, this incongruity should be verbalised in the AD to activate the comedic effect.

To show the feminine characteristics of this action, I used an adverb 妖娆地 ‘seductively’ in the initial draft, but it did not satisfy the co-creators. User representatives felt it was not clear exactly how Disheng was being ‘seductive.’ As pointed out by Reviere et al. (2015, pp. 16–17), a verb plus an adverb may still fall short of clarity and conciseness. However, due to time constraints, it is difficult to describe this action in detail to highlight the stereotypically feminine behaviours. Even with a full explanation, the comic action is likely to lose the dramatic effect within a long string of words. During the discussion, one user representative suggested a neat reference, 贵妃醉酒 ‘The Drunken Concubine,’ to highlight Disheng’s feminine behaviour in a metaphorical way (Table 3). This reference is to a well-known story in which Emperor Xuanzong of Tang’s beloved beauty, Yang Guifei, gets adorably drunk at a banquet. Co-creators agreed that using such a metaphorical reference to compare Disheng, a male character, to this famous beauty, would accentuate the incongruity for humorous effect. This concise linguistic metaphor also mitigates the time constraint, a significant barrier for audio describing humour (Martínez-Sierra, 2010).

| Initial draft                                                                              | Final version                                                                                        |
|--------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------|
| AD: 喝醉的艾迪生妖娆地半躺在石阶上。<br>Translation: Drunk Disheng Ai lies seductively on the stone steps. | AD: 艾迪生贵妃醉酒般卧在石阶。<br>Translation: Disheng Ai reclines on the stone steps like the Drunken Concubine. |

**Table 3. An example of verbalising the incongruity.**

Other humour mechanisms in the clip were verbalised in a similar way, such as twists and silliness. To illustrate briefly, when the character Zhuyu Zhang makes his first appearance, he looks like a martial arts master. But when he shows *Qinggong*, an airy movement like flying, he falls heavily to the ground. This twist was verbalised using the contrastive linking words 谁知 ‘who would have thought’ and 竟 ‘even, unbelievable,’ combined with a tone change in the AD voicing from admiration to mockery. In another scene, when Zhuyu lies completely still on a recliner, a person nearby puts fingers in front of Zhuyu’s nose to test if he is alive or not, making Zhuyu look sillier. A commonly recognised northern dialect expression, 还有气儿没 ‘still have breath or not,’ is then used to mock the silliness of Zhuyu’s posture. Dore (2020, p. 277) also notes that the Italian AD of the comedy *The Big Sick* is sometimes evocative in its attempt to suggest certain mechanisms, such as suspense, providing further support for this approach.

## 4.2. Adding new verbal humour cues

The co-creation team added five new verbal humour cues in the AD. One involved singing a line of the AD. Two examples involved retaining laughter by the audio describer. Two examples involved an exaggerated description of movement.

For example, in one scene, the two characters have a staring competition with a hawk. The hawk flaps its wings and teeters frantically on a branch. While the two characters wait for the hawk to fall and lose the game, the hawk regains its balance and settles. Based on the team's analysis, such a twist made the two characters look sillier, possibly amusing the audience. As Table 4 shows, the initial AD was 然而雄鹰只是一个踉跄而已 'However, the hawk only stumbles and regains its balance.' The comedy artists pointed out that this sounded a bit bland. After trying many unsatisfactory factual descriptions for the action, the comedy artists suggested slight exaggerations, describing this image as 雄鹰只是虚晃一枪 'the hawk just makes a feint one shot.' The Chinese four-character phrase 虚晃一枪 refers to confusing the opponent through dummy moves. The source visual information does not make clear what has caused the hawk to act like that, but the co-creators believed that adding this subtly exaggerated cue could make the hawk look more strategic and the two characters sillier, thereby amplifying the humorous effect. When applying this creative description, the co-creators agreed that the sound effect of the hawk flapping its wings first rapidly, and then to a steady rhythm, should be left audible. This helps to communicate that the hawk gradually regains its balance, a detail not included in the proposed 'exaggerated' AD. This interaction between the AD and the sound effects illustrates the expressive flexibility of the meaning-making repertoires which the compound nature of audiovisual humour (Martínez-Sierra, 2020) offers to the audio describer. A similar point has been made by O'Sullivan (2013), that multimodal texts offer expressive opportunities as well as challenges to the translator. If the narrative details are also carried by other elements, the AD has more room for creativity and emphasis on the dramatic value.

| Initial draft                                                                                                                                                                                                                                              | Final version                                                                                                                                                                                                                                                                   |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| AD: 突然！雄鹰疯狂地抖动翅膀，摇摇欲坠。两人一激灵，又在克制自己的兴奋之情，然而雄鹰只是一个踉跄而已。<br>Translation: Suddenly, the hawk flaps its wings, tottering frantically! The two people feel alert and have to restrain their excitement. However, the hawk only stumbles and regains its balance. | AD: 突然！雄鹰疯狂地抖动翅膀，摇摇欲坠。两人精神一振，然而雄鹰只是（停顿）虚晃一枪（憋住笑）。<br>Translation: Suddenly, the hawk flaps its wings, tottering frantically! The two people feel alert. However, the hawk (short pause for the sound effect of wings flapping) just makes a feint attack (suppressed laughter). |

Table 4. An example of adding new verbal humour cues in AD.

New verbal cues can also be introduced to the vocal delivery. During the co-creation process, user representatives suggested that when voicing funny scenes, I could occasionally retain my laughter in the AD as a natural response. In the example above, when I voiced the last line of the AD, the team agreed to maintain my suppressed laughter as a subtle way to mock the characters' silliness. According to user representatives, the laughter could make them feel more engaged. Helal Birjandi et al.

(2025) note a similar phenomenon in the Persian audio-described version of *Shaun the Sheep*, where one scene is accompanied by the audio describer's laughter. Helal Birjandi et al. believe the laughter could be an emotive touch for the audience (p. 70). However, López Rubio (2024, p. 455) shows that AD users hold different opinions about the idea of using a jocular tone for audio describing humour, and most object to it. This is in line with the neutral voicing conventions which have long been recognised as the norm in the AD field (Fryer, 2016, p. 88). Previous scholars (e.g., Cabeza-Cáceres, 2013) have found that emphatic, emotional AD delivery generated lower enjoyment in users. Jankowska et al. (2023) argue, on the other hand, that film genre, voice type and emotional valence can influence the voicing style, and that an emphatic voice may be beneficial for comedies. The choice of AD voicing seems not limited to an expressive vs. neutral binary frame but may be influenced by a number of factors. Although the practice of adding the audio describer's laughter seemed a promising expressive resource in this study, it still needs further examination regarding its applicable contexts, audience reception and so on.

These more or less subtle new verbal humour cues can compensate for inevitable losses of some of the visual humour. As Martínez-Sierra points out (2021, p. 35), when consuming audio-described comedy products, audiences with visual impairments can experience their humorous pleasure from different meaning-making resources without the need to rely on exactly the same ones used by sighted people. This also provides a rationale for complementing the comedic effect with new verbal humour cues.

### 4.3. Avoiding stereotypes and sensitive expressions

Humour is a multifaceted phenomenon, and when associated with social or cultural topics inappropriately, humour can become stereotypical or offensive. Our co-creation team encountered this five times in the clip: once in relation to a gendered expression; once in relation to a vulgar word, and three times in relation to the theme of death. For example, in one scene, Disheng Ai, with Xiao Ma's consciousness in his body, bursts into tears when recalling his/her previous tough life. The appearance of this big male character significantly conflicts with the act of wailing, which here is associated more with Xiao's femininity. In the initial AD draft, I used a humorous expression to highlight the incongruity: 艾迪生娘们似的哭哭啼啼 'Disheng Ai is crying like a baby girl.' 娘们, a casual colloquial term in Chinese for women/girls, has a negative meaning, which disparages women, and has been frequently used to mock people for being overly delicate. This AD caused dissatisfaction among the comedy artists. They considered that this 'humour' might have been accepted in the past, but today, it was unfunny and made women feel disrespected. User representatives further added that the feminine characteristics could also be deduced from the squeaky sound of Disheng's crying, thus enabling them to perceive the incongruity mechanism and get the joke. As Martínez-Sierra (2020) highlights and as we mentioned in the section above, the humour in audiovisual products is a multimodal negotiation and the missing of certain elements does not inevitably result in the missing of the overall humorous experience. Here, unlike the previous example in section 4.1, the incongruity mechanism still works without the funny visuals. After discussion, the final version reverted to a plain description of the visual details without the inappropriately gendered term.

Death proved to be another sensitive topic. When one supporting character, Xiunian, shows his martial arts skills on the highway, he is run over by a series of cars. The

initial draft used an internet slang phrase, 死透透 ‘hopelessly dead.’ However, one user representative believed that it was not appropriate to make fun of someone’s tragedy, and such humorous descriptions should be avoided. The other user representative added that he was unfamiliar with this internet slang. They both favoured simply describing the cars crashing into the character. Reminiscent of the cinematic technique of slow motion, comedy artists suggested relying on a slow and prolonged tone to hint at the imminence of death.

In the examples above, the audio describer’s proposed interventions were rejected, and the objective versions were agreed to be more appropriate. This supports the finding of previous research which shows that “vague and short descriptions” tend to be used by Chinese audio describers for sensitive visual content (Liu, 2023, pp. 187–188). As the ADLAB guidelines highlight, AD wording and style are always context-dependent and can be determined by many characteristics of both the source and the target (Taylor, 2015, pp. 46–47). Our findings therefore suggest that creative AD is not a one-size-fits-all solution for humour, but must maintain a balance with objectivity, depending on context. For sensitive topics, audio describers should carefully consider discourse norms and ethical issues. Only in this way can the products be truly accessible, funny, and friendly for all.

## 5. Lessons on collaborative working

Collaborative working mode refers to the engagement of different agents, perspectives and experiences in the meaning-making process. Reflecting on the co-creation team’s work, and combined with the analysis of the co-creators’ comments, we identify three aspects of the collaboration as particularly salient: engaging with creatives, engaging with diverse users with visual impairments, and power relations.

### 5.1. Engaging with creatives

Collaboration in creative AD has usually involved the creative team behind the original media (see section 2.2). It is argued that this ensures that a voice representing the source material is included (Fels et al., 2006b, p. 74) and that the overall AD team will have the authority to adapt the source material where necessary (Fryer, 2018). Not only did these studies involve the original creators, but they were initiated, or at least co-initiated, by them. As Patiniotaki (2022) observes, such a scenario requires that “creators [of the arts/media] [...] be aware [of] and in complete agreement with the provision of AD” (p. 147). In the absence of buy-in from the original creative team, an audio describer must seek either “alternative routes of integration or a more traditional approach” (p. 148). This is the case with our project. I did not find media makers interested in collaborating on an AD of their own work, which may reflect a certain lack of AD/accessibility awareness in the Chinese media industry.

When reflecting on the current relationship among the three sets of stakeholders involved in AD production (media makers, accessibility groups, and users with visual impairments), the co-creators in this project pointed out the same problem. According to them, accessibility groups represent the main facilitators of AD development in China. They are responsible for event organisation, material selection, AD standard-setting and exhibition. Users with visual impairments also contribute by making

requests and providing advice. By comparison, the engagement of media makers is limited:

Tang: From my experience as a student [in comedy creation], when we create student work, we rarely approach diverse user groups. However, some student clubs have been focusing on this area and organising relevant [accessible] events.

[...]

Zhu: One result [of filmmakers not caring much about this issue] is the delay, always delay [in producing accessible versions of these products by the third party].

While the participation of external creatives was beneficial to our AD script, the co-creators explicitly advocated for the inclusion of the original creative team as very desirable for future AD projects. As noted by Ye and Jie (2023, p. 229), the current engagement of media makers in China's AD development seems mostly limited to authorising AD groups to use their completed work. As our co-creators observed, the fact that AD is still an afterthought by media makers tends to cause problems, including delays in producing the accessible version of a film until after its theatrical run or even later. The 2025 Spring Festival Gala, organised by China Media Group, launched its sign language and audio-described versions simultaneously with its original version. To the best of our knowledge, this was the first attempt at the official level where the original creative team initiated and fully engaged in producing the accessible version together with media accessibility experts and user representatives. This approach was repeated in the 2026 Gala, which might set a new trend in China to engage media makers more in the accessibility process.

## 5.2. Engaging with diverse users with visual impairments

Viewers with visual impairments may have different sight conditions. According to the national Classification and Grading Criteria of Disability, visual impairment in China is divided into blindness and low vision based on visual acuity or visual field radius (Yantian District Civil Affairs Bureau, 2025). Different types of visual impairments vary in visual memory and residual vision, which may impact consumption of audio-described products. Chmiel and Mazur (2021, p. 564) found that types of visual impairment (congenital blindness, non-congenital blindness, and low vision) can modulate AD preferences, including in relation to objective descriptions of characters; audiences with low vision are more likely to accept such subjective judgements as they can see some of the source images (p. 560). The user representatives in our project felt that their views could not represent all the experiences of members of their community:

Qi (pseudonym): As an individual who became blind [several] years ago, I don't know the situation of those congenitally blind, and the way they think or understand the world.

[...]

Qi (pseudonym): I can only represent a very small group of audience with visual impairment [...] In the future, can we invite more individuals with visual impairments to join the collaborative working process?

Along similar lines, the recent Workshop for Inclusive Co-created Audio Description (W-ICAD) model also suggests that each co-creation group should be made up of

“partially blind, blind, and sighted co-creators to embrace, acknowledge, and integrate their diverse voices within co-created descriptions” (Eardley et al., 2024, p. 223).

### 5.3. Power relations: authority and dialogue

The co-creators applied their expertise and exerted power in different aspects of our project (see section 4). They shared that they enjoyed dialogue and deep interaction with members from different communities, which not only improved the AD quality, but also gained them first-hand insights into each other’s lives and opinions, as shown in the quotes below. This suggests that, through close dialogue, team AD can facilitate the interaction “that surpasses the cognitive capacities” (Hirvonen et al., 2023, p. 37) and can allow access to be a “social process” (Reviers & Hanoulle, 2023, p. 117).

Zhu: This working pace enables people with different views to get to know each other better and then to work together for one shared goal.

[...]

Qi (pseudonym): As persons with visual impairment, we can engage in the production process behind the scenes, knowing what creative teams’ work looks like and their workflow.

Though rare, some minor challenges or frictions cropped up. Comedy creators seemed to experience some uncertainty in how to interact appropriately with user representatives. As Tang said, “I have not interacted with persons with impairments before [...] I was afraid that I made some mistakes [...] I have indeed realised that there are some aspects we should consider.” When using what turned out to be inappropriate humorous descriptions to mock one character’s death, the comedy creators and I did not realise there was a problem until one user representative approached me after the team discussion and expressed his discomfort with mocking someone’s tragedy. Only after this did the team revisit this point further and change to a more tactful description.

These subtle frictions were not escalated in this project, mainly because much of the authority still rested with me (the first author), as I was the initiator, convenor, designer of the research and audio describer in this situated research practice. Because the research was designed to listen to the voices of co-creators and lean into the collaborative approach, I was quite happy for the co-creators to push back against some of my initial suggestions. This would not necessarily be the case in a study with multiple creative stakeholders invested in their individual creative input. For example, in Marinetti and Rose’s (2013) collaborative translation project for a theatre performance, the team experienced tense conflicts and a temporary halt to the collaborative activities. This is because Marinetti’s and Rose’s expectations and professional dispositions as researcher-translators were different from those of the artistic team, and their authority was weakened by the director’s “own authority and his professional legitimacy” (p. 175). We should also bear in mind that if the original creative team are involved in the production of AD, this may create different dynamics again.

## 6. Conclusion

Visual humour requires nuanced AD treatment in order to be still funny. This exploratory study employed a creative collaborative AD approach to address a comedy fragment and prioritise its comedic value. Visual humour mechanisms can be verbally refracted or recreated and new verbal humour cues can be added. Solutions such as the inclusion of laughter by the audio describer were well-received by the co-creation team. At the same time, creative solutions were not always found to be appropriate. They were sometimes rejected, and a more objective approach reinstated, especially, though by no means only, when dealing with sensitive topics.

Our study also sheds light on practising a collaborative working mode for AD creation in the Chinese context. External creatives can make significant contributions to AD, though our co-creators still felt that the participation of the original creative team would be the ideal. A co-creation team should consider user representatives with different sight conditions, not assuming that any individual represents the whole community. While power relations did not prove a significant problem in our study, this is likely to be because the audio describer was also the lead researcher and heavily invested in making the collaboration work. Scaling up collaborative approaches to AD will require each co-creator being open to trusting other members and their expertise, and to familiarising themselves with the norms and needs of other communities involved.

As an exploratory study, this project involved a small team of student comedy creators and user representatives in a university setting. In the future, AD co-creation could invite artistic teams behind the working material and start from the pre-production stage in a fully authentic practical environment. It could also invite external creatives, not involved with the original material, but in a more authentic industry setting. The existence of a previous AD script for the film, which we were unable to consider in detail in our study, could also allow for further research, e.g., a comparative reception study.

On the basis of our study, we argue that co-creation is a promising approach to the creation of AD scripts for humorous content and that creative approaches to the AD of humour have great entertainment potential. It remains to be seen whether ‘creative AD’ becomes more widely used as an alternative to the conventional objective approach, but our study also shows that hard and fast distinctions cannot necessarily be drawn between the ‘creative’ and the ‘objective’ approach, because both types of solution are represented in our AD script based on the specific context, the nature of the material, and the views of a variety of co-creators.

## Acknowledgements/Funding

This project was funded by the UKRI-AHRC International Placement Scheme hosted at the Shanghai Theatre Academy (Ref: APP29773). Zi Ye’s doctoral study is funded by the AHRC via the South, West and Wales Doctoral Training Partnership [AH/R012776/1].

We would like to thank Dr Jie Yi and Ms Yu Zong at Shanghai Theatre Academy for their invaluable support throughout this project. We also extend our gratitude to Mr

Qinghan Qi (pseudonym), Ms Jinglan Su, Ms Yingyi Tang, Ms Xiaoqi Yu, and Mr Junyi Zhu for joining this project.

## References

- [deleted]. (2024). *Any shows or movies make jokes with the audio description?* [Online forum post]. Reddit. [https://www.reddit.com/r/Blind/comments/19buemn/any\\_shows\\_or\\_movies\\_make\\_jokes\\_with\\_the\\_audio/](https://www.reddit.com/r/Blind/comments/19buemn/any_shows_or_movies_make_jokes_with_the_audio/)
- Academy of Accessibility Research. (n.d.). 首页 [Home]. <https://aar.org.cn/>
- Bessaha, M., Reed, R., Donlon, A., Mathews, W., Bell, A., & Merolla, D. (2020). Creating a more inclusive environment for students with disabilities: Findings from participatory action research. *Disability Studies Quarterly*, 40(3). <https://doi.org/10.18061/dsq.v40i3.7094>
- Bogucki, Ł. (2013). Action research as a methodology for audiovisual translation. In M. Thelen, & B. Lewandowska-Tomaszczyk (Eds.), *Translation and meaning part 9* (pp. 111–118). Hogeschool Zuyd/Maastricht School of Translation and Interpreting.
- Cabeza-Cáceres, C. (2013). *Audiodescripció i recepció. Efecte de la velocitat de narració, l'entonació i l'explicitació en la comprensió filmica* [Doctoral dissertation, Universitat Autònoma de Barcelona]. Dipòsit Digital de Documents de la UAB. <https://ddd.uab.cat/record/111648?ln=en>
- Chmiel, A., & Mazur, I. (2021). A homogenous or heterogeneous audience? Audio description preferences of persons with congenital blindness, non-congenital blindness and low vision. *Perspectives*, 30(3), 552–567. <https://doi.org/10.1080/0907676X.2021.1913198>
- Di Giovanni, E. (2018). Participatory accessibility: Creating audio description with blind and non-blind children. *Journal of Audiovisual Translation*, 1(1), 155–169. <https://doi.org/10.47476/jat.v1i1.50>
- Díaz Cintas, J. (2005). Audiovisual translation today: A question of accessibility for all. *Translating Today*, 4, 3–5.
- Dore, M. (2020). *Humour in audiovisual translation. Theories and applications*. Routledge. <https://doi.org/10.4324/9781003001928>
- Eardley, A. F., Jones, V. E., Bywood, L., Thompson, H., & Husbands, D. (2024). The W-ICAD model: Redefining museum access through the Workshop for Inclusive Co-created Audio Description. *Curator*, 68(1), 219–242. <https://doi.org/10.1111/cura.12649>
- Fels, D. I., Udo, J. P., Diamond, J. E., & Diamond, J. I. (2006a). A comparison of alternative narrative approaches to video description for animated comedy. *Journal of Visual Impairment & Blindness*, 100(5), 295–305. <https://doi.org/10.1177/0145482X0610000507>
- Fels, D. I., Udo, J. P., Ting, P., Diamond, J. E., & Diamond, J. I. (2006b). *Odd Job Jack* described: A universal design approach to described video. *Universal Access in the Information Society*, 5(1), 73–81. <https://doi.org/10.1007/s10209-006-0025-0>
- Fryer, L. (2016). *An introduction to audio description: A practical guide*. Routledge. <https://doi.org/10.4324/9781315707228>
- Fryer, L. (2018). Staging the audio describer: An exploration of integrated audio description. *Disability Studies Quarterly*, 38(3). <https://doi.org/10.18061/dsq.v38i3.6490>
- Greco, G. M. (2018). The nature of accessibility studies. *Journal of Audiovisual Translation*, 1(1), 205–232. <https://doi.org/10.47476/jat.v1i1.51>



- Helal Birjandi, M., Emami, H., & Ameri, S. (2025). Audio describing humour in Persian: The case of *Shaun the Sheep*, *Road Runner* and *The Pink Panther*. *The Journal of Specialised Translation*, 44, 64–84. <https://doi.org/10.26034/cm.jostrans.2025.8469>
- Hirvonen, M., Hakola, M., & Klade, M. (2023). Co-translation, consultancy and joint authorship: User-centred translation and editing in collaborative audio description. *The Journal of Specialised Translation*, 39, 26–51. <https://doi.org/10.26034/cm.jostrans.2023.065>
- Jankowska, A., Pilarczyk, J., Wołoszyn, K., & Kuniecki, M. (2023). Enough is enough: How much intonation is needed in the vocal delivery of audio description? *Perspectives*, 31(4), 705–723. <https://doi.org/10.1080/0907676X.2022.2026423>
- Kemmis, S., & McTaggart, R. (1982). *The action research planner*. Deakin University Press.
- Lewin, K. (1946). Action research and minority problems. In G. W. Lewin (Ed.), *Resolving social conflicts. Selected papers on group dynamics* (pp. 143–152). American Psychological Association.
- Liu, H. (2025). 2024 年中国电影产业发展分析报告 [2024 analysis report on Chinese film industry development]. *Contemporary Cinema*, 2, 14–23.
- Liu, Q., Jia, H., Pang, C. L., & Jankowska, A. (2025). Audio describers in Mainland China: Insights from a survey of emerging professionals and unique practices. *Między Oryginałem a Przekładem*, 31(3(69)), 53–77. <https://doi.org/10.12797/MOaP.31.2025.69.04>
- Liu, Y. (2023). Journey to the East. What is happening in the Chinese audio description (AD) world? An interview-based study. *Hermēneus*, 25, 169–194. <https://doi.org/10.24197/her.25.2023.169-194>
- Liu, Y., Casas-Tost, H., & Rovira-Esteva, S. (2022). Made in China versus Made in Spain. A corpus-based study comparing AD in Chinese and Spanish. *The Journal of Specialised Translation*, 38, 180–207. <https://doi.org/10.26034/cm.jostrans.2022.088>
- López Rubio, M. (2024). *Descripción y recepción de la audiodescripción de comedias en Netflix España para personas con discapacidad visual. Una propuesta de mejora* [Doctoral dissertation, Universitat de València]. ResearchGate. 10.13140/RG.2.2.12456.38403
- Luque Colmenero, M. O., & Soler Gallego, S. (2020). Metaphor as creativity in audio descriptive tours for art museums: From description to practice. *Journal of Audiovisual Translation*, 3(1), 64–78. <https://doi.org/10.47476/jat.v3i2.2020.128>
- Marinetti, C., & Rose, M. (2013). Process, practice and landscapes of reception: An ethnographic study of theatre translation. *Translation Studies*, 6(2), 166–182. <https://doi.org/10.1080/14781700.2013.777258>
- Martínez-Sierra, J. J. (2009). The relevance of humour in audio description. *Intralinea*, 11. <https://www.intralinea.org/index.php/archive/article/1653>
- Martínez-Sierra, J. J. (2010). Approaching the audio description of humour. *Entreculturas*, 2, 87–103. <https://doi.org/10.24310/Entreculturasertci.vi2.11762>
- Martínez-Sierra, J. J. (2020). Audio describe humour: Seeking laughter when images do not suffice. In M. Dore (Ed.), *Humour translation in the age of multimedia* (pp. 177–195). Routledge.
- Martínez-Sierra, J. J. (2021). Testing the reception of audio described humor: An exploratory approach. *International Studies in Humour*, 10(1), 22–39.
- Mateo, M., & Zabalbeascoa, P. (2019). Translation and humour. In R. Valdeón, & A. Vidal (Eds.), *Routledge Handbook of Spanish translation studies* (pp. 139–156). Routledge.
- O’Sullivan, C. (2013). Introduction: Multimodality as challenge and resource for translation. *The Journal of Specialised Translation*, 20, 2–14. <https://doi.org/10.26034/cm.jostrans.2013.398>



Patiniotaki, E. (2022). Audio description for dance performances: An artistic and collaborative approach. *Status quaestionis*, 23, 143–165. <https://doi.org/10.13133/2239-1983/18228>

Rai, S., Greening, J., & Petré, L. (2010). *A comparative study of audio description guidelines prevalent in different countries*. Royal National Institute of Blind People. <https://unidescription.org/storage/app/uploads/public/5f1/a3e/bb1/5f1a3ebb17896460620035.pdf>

Reason, P., & Bradbury, H. (Eds.). (2006). *Handbook of action research* (Concise paperback edition). SAGE.

Remael, A., Reviers, N., & Vercauteren, G. (Eds.). (2015). *Pictures painted in words: ADLAB audio description guidelines*. EUT. <https://repository.uantwerpen.be/docman/irua/99e3cf/130865.pdf>

Reviers, N., Remael, A., & Daelemans, W. (2015). The language of audio description in Dutch: Results of a corpus study. In A. Jankowska, & A. Szarkowska (Eds.), *New points of view on audiovisual translation and accessibility* (pp. 167–189). Peter Lang.

Reviers, N., & Hanouille, S. (2023). Aesthetics and participation in accessible art experiences: Reflections on an action research project of an audio guide. *Journal of Audiovisual Translation*, 6(2), 99–121. <https://doi.org/10.47476/jat.v6i2.2023.277>

Romero-Fresco, P. (2019). *Accessible filmmaking: Integrating translation and accessibility into the filmmaking process*. Routledge. <https://doi.org/10.4324/9780429053771>

Romero-Fresco, P., & Chaume, F. (2022). Creativity in audiovisual translation and media accessibility. *The Journal of Specialised Translation*, 38, 75–101. <https://doi.org/10.26034/cm.jostrans.2022.084>

Salway, A. (2007). A corpus-based analysis of audio description. In J. Díaz Cintas, P. Orero, & A. Remael (Eds.), *Media for all: Subtitling for the deaf, audio description, and sign language* (pp. 151–174). Rodopi. [https://doi.org/10.1163/9789401209564\\_012](https://doi.org/10.1163/9789401209564_012)

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>

Soler Gallego, S., & Luque Colmenero, M. O. (2023). Increased subjectivity in audio description of visual art: A focus group reception study of content minimalism and interpretive voicing. *Journal of Audiovisual Translation*, 6(2), 55–76. <https://doi.org/10.47476/jat.v6i2.2023.248>

Taylor, C. (2015). The language of AD. In A. Remael, N. Reviers, & G. Vercauteren (Eds.), *Pictures painted in words: ADLAB Audio description guidelines* (pp. 46–49). EUT.

theOriginalBlueNinja. (2024). *Audio description humour?* [Online forum post]. Reddit. [https://www.reddit.com/r/Blind/comments/1bwsx0n/audio\\_description\\_humor/](https://www.reddit.com/r/Blind/comments/1bwsx0n/audio_description_humor/)

Tor-Carroggio, I. (2020). *Audio description in China: Past, present and future* [Doctoral dissertation, Universitat Autònoma de Barcelona]. Dipòsit Digital de Documents de la UAB. <https://ddd.uab.cat/record/241107?ln=en>

Udo, J. P., & Fels, D. I. (2009). Suit the action to the word, the word to the action: An unconventional approach to describing Shakespeare's *Hamlet*. *Journal of Visual Impairment & Blindness*, 103(3), 178–183. <https://doi.org/10.32920/14637930>

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

Wang, X. (2024). Co-design of a voice-driven interactive smart guide for museum accessibility and management. *Journal of Audiovisual Translation*, 7(1), 1–24. <https://doi.org/10.47476/jat.v7i1.2024.267>



Yantian District Civil Affairs Bureau. (2025, October 27). 深圳市残疾人等级评定工作 [Shenzhen disability grading assessment work]. The People's Government of Yantian District, Shenzhen. [https://www.yantian.gov.cn/cn/service/shbz/shfl/cjrl/zjpdzbz/content/post\\_12458608.html](https://www.yantian.gov.cn/cn/service/shbz/shfl/cjrl/zjpdzbz/content/post_12458608.html)

Ye, Z. (2025, March 19–21). *Exploring the audio description of humour: A multimodal corpus study in the Chinese context* [Conference presentation]. Advanced Research Seminar on Audio Description, Barcelona, Spain.

Ye, Z., & Jie, H. (2023). Enhancing cinema experience for visually impaired persons in China: A case study on Shenzhen. *Journal of Chinese Film Studies*, 3(2), 221–240. <https://doi.org/10.1515/jcfs-2022-0030>

## Data availability statement

The data supporting this manuscript will be made available — with registration required — via the University of Bristol repository within one year following Zi Ye's PhD completion expected by the end of 2026. A formal link to the data will be provided once it becomes available.

## Appendix: Chinese AD script

Co-created by Qinghan Qi (pseudonym), Jinglan Su, Yingyi Tang, Zi Ye, Xiaoqi Yu, Junyi Zhu (listed in alphabetical order).

马小、艾迪生抵达卷帘门。门内出现一个胖墩墩的白衣背影，长发飘逸，正弯腰舀水。

【“大婶”之后】

这人一脸懵地抬头。

【“请问”之后】

原来是一个男性，胡子拉碴、长相油腻。

【“在吗”之后】

长发男子生无可恋地盯着马小，喝了一口瓢中的水，然后叹了口气儿。

【“我在”之后】

众人抬头，循声望向二楼连廊，只见一人双手背后，迈着四方步走来，他身着黑衣，胸前用金线绣着一个巨大的‘卷’字。黑衣人浅浅一笑，振臂一挥，直接从二楼飞了下来（崇拜语气），谁知一个踉跄，最后竟跪倒在大家面前（笑出声音）。

【“是吧”之后】

长发男反应过来，忙慌地将张茱萸搀扶起来。

【“我就是掌门”之后】

长发男吐槽。

【“怎么了”之后】

马小失望地转身要走。

【“等等”之后】

马小回头看着张茱萸。

【“说出来吗”之后】

张茱萸摆出了骂街的姿势，艾迪生说。

【“凭我”之后】

马东撑着大门、出现在卷帘门口，张茱萸吃惊地张开了嘴。（停顿）转眼到晚上。

【“吃饭”之后】

四个杀马特造型的弟子瞬间全挤到饭桌上，狼吞虎咽。

【“大白鹅”之后】

一群大白鹅在前开路。众人来到高速公路。

【“车辆上”之后】

马小愣住。

【“什么”之后】

众人站在路旁，扫视着疾驰而过的车辆。

【“秀念”之后】

秀念就是之前身着白衣的长发男子。听到召唤，他从众人身后一个翻身，飘逸的长发在空中直直竖起，然后轻轻落地，十分潇洒。他斜眼看着艾迪生，拿过他手里的广告，迈着不屑的步伐，面无表情地来到高速公路中间。两辆车呼啸而过，又将他飘逸的发丝吹起。他将广告一把撒向空中，胖墩墩的身体做好冲刺状，踮起脚尖，凌波微步，轻盈地跑了起来（快）。一辆红色大货车迎面而来。

**【歌曲开始时】**

秀念腾空而起，抓过空中一广告，叠成三角状，看招！（留白音效）啊哦，被创飞倒在地上。他的脚抽动了一下，看来还有救。众人呆住，马小慢慢走向秀念，刚要伸手（快）（留白音效）一辆轿车又顺着秀念碾身而过……（“过”字稍微拖长）。灵堂内摆放着秀念露齿笑的黑白照片，音容宛在。

**【“零失误”之后】**

秀念灵位立在台前，其余三位弟子头缠白布，跪在地上烧纸钱，夸张地在那儿哭丧叫魂。马小艾迪生马东站在一旁。

**【“带走吧”之后】**

张茱萸端着一篓子黄色烧纸哭啼啼地站在一边。

**【“活不下去了”之后】**

他把纸全扣在火盆里，直接把火给灭了（无语语气）。

**【画面转换之后】**

天黑，马小在屋里收拾衣服。

**【“可能呢”之后】**

艾迪生站在背后，眉头微皱。

**【“你赢了呢”之后】**

刚拉上背包的马小呆愣在原地。随后来到室外，马小大大咧咧岔着腿坐在石阶上，艾迪生呢，双膝并拢。

**【“喝酒”之后】**

马小十分潇洒地与艾迪生碰杯。一眨眼，地上多了一大堆空啤酒罐。

**【“干死他”之后】**

她哪里是不在意呀（调侃语气）。

**【“放过”之后】**

喝醉的马小骑在扶手上。

【“他们”之后】

艾迪生贵妃醉酒般卧在石阶上。

【“打到我了”之后】

大老爷们艾迪生还嘟嘴撒娇（嫌弃语气）。

【“气质”之后】

艾迪生搔首弄姿起来。

【“我爸”之后】

画面泛黄，呈老照片式回忆（压台词）。

【“父亲”之后】

带着红领巾的马小上台敬了队礼。

【“哪去了”之后】

画面闪回到现实。

【“我错了”之后】

艾迪生放开了大哭起来，马小赶紧从扶手上下来。

【“不一样”之后】

艾迪生举起啤酒罐一饮而尽。

【“好记者”之后】

马小坐到艾迪生身边，左手轻拍着他后背，自己却也只能仰天叹气。

【“我，我”之后】

马小用手不断戳着自己。

【“放弃”之后】

她眼神低垂，神情平静却又落寞。

【“不该当记者”之后】

马小眉头紧皱，满脸的无语。

【“又哭了呢”之后】

好话不听只能来硬的了呗。

【“憋回去”之后】

果然奏效，艾迪生捂住嘴巴。

【“我不能再喝了”之后】

艾迪生马小站起来，走路都走不稳当，摇摇晃晃地互相打闹着，天蒙蒙亮，深邃的山谷还是一片祥和宁静。（停顿）白天，张茱萸闭眼躺在椅子上，马东用手试试他还有气儿没。

【“要走”之后】

张茱萸一下睁开眼睛。

【“设计广告吧”之后】

张茱萸意味深长地摇摇头（轻），突然手向上一指。

【“往上看”之后】

一片普通的树荫而已，没啥特别的。

【“在干什么啊”之后】

张茱萸想吹个帅气的口哨（音效保留）。诶，没出响。（鸟鸣）但雄鹰还是从天而降了，稳稳落在一个树干上，眼神十分犀利。众人惊讶。

【“太难了吧”之后】

马小转向艾迪生，温柔地看着他。

【“熬死它”之后】

艾迪生信心满满地点头。之后，马小艾迪生两人并坐一排，眼睛一眨不眨地盯着雄鹰，志在必得。雄鹰此时也是悠然自在。夜晚，马东在后面开始搭帐篷，打算在这看家了。

雄鹰的眼神依旧凌厉，马小艾迪生也是眼睛瞪得像铜铃（唱起来），死死地盯着雄鹰。两人从刷牙到吃饭、从白天到黑夜，从坐着到站着，无论做什么，眼神都不能离开雄鹰。马东拿着电蚊拍都能拍死艾迪生头上的虱子。两位女技师在给两人做足疗。

### 【“怕痒”之后】

有只青蛙趴在艾迪生头上，二人已经完全变成熊猫眼，鸡窝头。突然！雄鹰疯狂地抖动翅膀，摇摇欲坠。两人精神一振，然而雄鹰只是（停顿）虚晃一枪（憋住笑）。

### 【加速音乐时】

又过了好多天，两人或是用胶带把眼皮粘起来、或是直接用手扒拉着眼皮。雄鹰时而摇晃一下，时而又稳若泰山。艾迪生马小咬牙切齿，已经快要疯了！又是一个白天，二人还坐在雄鹰面前，眼神呆滞、不成人样，困到一直翻白眼。马东在后面和一条巨大的蟒蛇厮杀搏斗，完全的无厘头。

### 【“能熬呢”之后】

就在这时！雄鹰直挺挺地！倒地了！

---

## Notes

<sup>1</sup> Throughout this article, we refer to Mainland China as China.

<sup>2</sup> All instances of 'I' in this paper refer to the first author, and 'we' refers to all the authors, except when these pronouns appear in quotations.