Delivery approaches in audio description for the scenic arts[1]

Irene Hermosa-Ramírez

Universitat Autònoma de Barcelona, Spain

Abstract

Audio description (AD) is becoming an increasingly mature modality within Audiovisual Translation Studies (AVTS) and Media Accessibility Studies. Concurrently, technological advances are steadily being put at the forefront of its practice. The aim of this article is to define the current status and development of AD for the scenic arts from a technical perspective. First, an overview of guidelines that specifically include recommendations on delivering AD for the scenic arts is presented. The emphasis is then placed on the implications of the delivery approaches currently applied to this modality. In this context, theatre venues can offer AD – along with other access services – in a live, semi-live or automated manner. The advantages and challenges for each approach are thus analysed and compared by presenting examples and applications in practice. Ultimately, the present descriptive study concludes that the live, on-site delivery approach is no longer the default in Spanish venues. This conclusion opens up new research paths on the reception of innovative practices and software solutions. It is tentatively suggested that involving the creative team and the blind and visually impaired patrons would be key to choosing the most suitable delivery approach for each production.

Keywords

Audio description, scenic arts, state of the art, live delivery, text-to-speech

1. Introduction

Accessibility measures have a somewhat longstanding tradition within the scenic arts, with theatre being the first audiovisual medium to adopt audio description (AD) in the 1980s (Pfanstiehl & Pfanstiehl, 1985) and surtitles for opera emerging that very same decade (Mateo, 2002). Today, access services for the scenic arts are understood in a broader sense. On the one hand, (multisemiotic) translation modalities encompass audio subtitles, surtitles, subtitles for the deaf and hard of hearing, AD, sign language interpreting and braille notes. On the other hand, complementary services such as touch tours, scale models, easy-to-read programmes and relaxed performances are also offered at some theatre venues. In recent times, universal systems for media access that provide several of these services, including linguistic access, are being developed (Oncins, Lopes, Orero, Serrano & Carrabina, 2013).

AD – a "verbal commentary providing visual information for those who are unable to perceive it themselves" (Fryer, 2016, p. 1) – falls within the scope of AVT as an intersemiotic translation modality. In addition, as illustrated by the array of services mentioned above, AD also belongs to Media Accessibility Studies, in the universalist sense that it concerns "access to media products, services, and environments for all persons who cannot, or cannot completely, access them in their original form" (Greco, 2018, p. 211). Out of all the dynamic modalities of AD – film, television, the scenic arts and other live events – (Matamala, 2007; Braun, 2008), the present study specifically targets the state of the art of AD delivery for theatre and opera with an emphasis on the Spanish context.

Earlier academic literature on AD devoted specifically to theatre (Pfanstiehl & Pfanstiehl, 1985; Navarrete, 1997) and opera (Matamala, 2007; Orero & Matamala, 2007; Puigdomènech, Matamala & Orero, 2008; Cabeza-Cáceres, 2010) largely took a descriptive approach. In turn, reception studies have certainly gained ground in recent years, with user-centered approaches becoming increasingly prevalent in researching AD for the scenic arts (Eardley-Weaver, 2014; Di Giovanni, 2018). Similarly, academic interest has seen a significant trend towards inclusive design, non-traditional AD and an advocacy for access to creation (see Udo & Fels, 2010; Whitfield & Fels, 2013; Fryer, 2018; Roofthooft, Remael & Van den Dries, 2018). These studies illustrate inclusive design within theatre, and they have advocated for a switch from conventional, ad-hoc and neutral AD to AD as an integral part of the creative process. This is, however, still not the case in most productions, regardless of the technical approach to AD which is applied.

Dealing with a promising but largely unexplored research topic (Remael, Reviers & Vandekerckhove, 2016), it was fitting to carry out a descriptive study to explore the state of the art of the prevalent technical approaches to delivering AD at live events. This article is thus organised as follows. The point of departure are some preliminary remarks on AD guidelines that specifically deal with delivery recommendations for the scenic arts (section 2). Then an overview of the describers' tasks and the necessary technical equipment is briefly presented (section 3). The core of the paper later disseminates the most common technical approaches in AD for the scenic arts in recent years, with an emphasis on the Spanish context: live (section 4), semi-live (section 5) and automated (section 6) delivery of AD. In section 7, a comparison of the advantages and challenges of the different approaches is proposed so as to tentatively define the different scenarios in which they would be most useful. To conclude the article, we outline the need to further tackle the matters discussed from a user-centered approach.

2. Delivery in AD guidelines

Starting with the subject of recommendations of AD, standards and guidelines – developed by public bodies, associations and non-profit organisations (Matamala & Orero, 2013) – have been published in countries that offer AD. Some of these documents specifically provide rec-

ommendations on the scenic arts, though most standards and guidelines focus on the filmic AD modality. Here, we look at three documents: the Spanish official standard UNE 153020 (AENOR, 2005, pp. 9-10), the set of recommendations proposed by the American Council of the Blind (ACB) (Snyder, 2010, pp. 21-42) and the European ADLAB guidelines (Remael *et al.*, 2015, pp. 58-61, 64-68). This section briefly summarises specific recommendations on (live vs. semi-live) delivery approaches, voice differentiation and gender, the widespread combination of audio introductions with audio descriptions and guidelines proposed by academic experts.

All three cited guidelines mainly presuppose a strictly live delivery, with no mention of semi-live possibilities (either pre-recorded or following a text-to-speech [TTS] approach, both of which will be elaborated on below). The exception is the ACB document, which illustrates an interesting outlook on repeatable plays: "Increasingly, certain long-running and/or touring productions have recorded description keyed to lighting cues and accessed via PDAs attached to seatbacks" (Snyder, 2010, p. 24). The question of repeatability will be key in the delivery approaches that aim for a more automated workflow, as detailed in sections 5, 6 and 7.

Regarding the number or differentiation of voices that deliver the AD script, the ACB and the ADLAB guidelines suggest that two describers take on the task and deliver the AD jointly (Snyder, 2010, p. 21). Or, in the case of combining AD with audio subtitles – "the media accessible mode of reading aloud, or voicing, subtitles" (Orero, 2007, p. 141) – to enable linguistic access, the ADLAB guidelines propose that "two or more voices may be used for the audio subtitles to help the target audience differentiate between speakers" (Remael *et al.*, 2015, pp. 61-62). For its part, the Spanish standard does not mention the possibility of distinguishing between voices. As for the gender of the voice or voices specifically, the ADLAB guidelines briefly touch on the relevance of choosing a voice talent "whose voice contrasts with the voices of the dialogues (e.g. in a film with many male roles you may want a female AD voice)" (Remael *et al.*, 2015, p. 55). Within the scope of this article, the ACB guidelines suggest that, in the case of opera, two describers, a man and a woman, take on the task jointly: one reads the surtitles and the other describes (Snyder, 2010, p. 38). By having two differentiated voices, blind and visually impaired patrons can discriminate their functions easily.

All three reported guidelines do agree that AD for the scenic arts is often preceded by an audio introduction (AI). AIs are "pieces of continuous prose, spoken by a single voice or a combination of voices lasting between five and 15 minutes" that provide information such as "running time, cast and production credits" and "descriptions of the set, costumes and characters" (Fryer & Romero-Fresco, 2014, p. 11). On a similar note, Cabeza-Cáceres (2010, p. 234) specifies that an AI for opera includes "broad facts about the opera, general features of the scenography, a summary of the plot, an introduction to the characters and some general notes on the costumes." According to York (2007), in the opera houses of London, a standalone, non-intrusive AI has traditionally been the common practice for dance and opera performances. The author goes on to provide some recommendations regarding aspects to be considered when drafting the script: avoid the sense of theatricality, keep a certain sympathy between the period setting of the play and the language used in the introduction, etc. Delivery-wise, AIs are the most flexible segment of an AD. For the most part, they do not have to deal with the time constraints of the AD. While the audience takes a seat, AIs can be voiced live, or a pre-recorded version can be rendered. Additionally, several opera houses in Europe offer to send out CDs with the AI recording for the patrons to listen to in advance, such as the Scottish Opera. Alternatively, some venues have chosen to upload these files to online platforms such as Soundcloud, as is the case of the Royal Opera House in London. For the purposes of this article, AD is broadly understood as the combination of AI and AD.

Lastly, besides the discussed official guidelines, several scholars have drafted their own recommendations that address the idiosyncrasies of the scenic arts (Orero, 2005; Matamala, 2007; Matamala & Orero, 2007; York, 2007; Puigdomènech *et al.*, 2008; Holland, 2009; Cabeza-Cáceres, 2010; Eardley-Weaver, 2010). Regarding delivery, Fryer (2016, p. 113) also acknowledges that it is common practice for the scenic arts AD in the UK to split the AD task between two describers, who, having made sure that their terminology is consistent, take one act each. Technical possibilities for delivery in practice will be discussed in the next section.

3. AD production and delivery, step by step

Nowadays, the two most widespread approaches to scenic arts AD are live AD (Matamala, 2007) and semi-live AD, an approach that has as yet received little attention in AD research. For the sake of terminological clarity, a live AD approach implies that the audio describer is present at the venue, voicing and tailoring their script as the performance takes place. Meanwhile, in the scenic arts context, a semi-live strategy refers to an AD fragmented in extracts and later synchronised live, though not necessarily by the audio describer. These extracts can either be pre-recorded in separate files – by one or two voice talents – or they can be a fragmented written text read aloud by a speech synthesizer (TTS AD). As for the largely unexplored automated approach, its rationale is that AD is delivered without the need for a person to cue the AD fragments manually. That is to say, without human interaction.

Overall, the agents involved, as well as their workflow, differ when applying one strategy or the other. Nevertheless, the preparation and scriptwriting tasks are similar for audio describers working either way. Following common practice in the Spanish context, the describer should have access to the script or libretto at least two weeks prior to the show, together with a recent video recording of the play, and a copy of the programme (Matamala & Orero, 2007, p. 209). A few days before the show, the audio describer attends a rehearsal to adapt the script or make any necessary modifications (Hernández-Bartolomé & Mendiluce-Cabrera, 2004, p. 272; Matamala, 2007).

In a live AD context, the audio describer who has written the script is also generally the one who voices it live (Matamala, 2007). Conversely, in a semi-live AD, the audio describer is not necessarily the one that synchronises or launches the description the day of the show. Semi-live ADs can therefore be used in different geographical locations, some examples being touring productions or street performances. In this case, the same (fragmented) AD can be repurposed as many times as the play is performed.

Focusing on the technical requirements for this service – be it live or semi-live –, AD has traditionally been transmitted wirelessly via infra-red or FM radio systems (Snyder, 2010, p. 24). These systems usually allow for several channels to provide different access services, an example being one channel for AD and another for amplified sound. Either the venues themselves invest in this equipment, or they can loan it from a third party. In recent years, however, playback via mobile apps is becoming an increasingly plausible option (Oncins *et al.*, 2013), particularly for semi-live and automated initiatives.

Figure 1. Multi-channel equipment for accessibility. Source: The John F. Kennedy Center for the Performing Arts, ©2013

At present, the application of technologies to AD is an emerging topic being discussed in terms of speech recognition, machine translation and speech synthesis (Matamala, 2016), for instance. In the following section, challenges and applications of live, semi-live and automated AD for the scenic arts will be placed at the forefront. The focus will be the suitability of TTS AD, mobile apps, software and efforts towards automation.
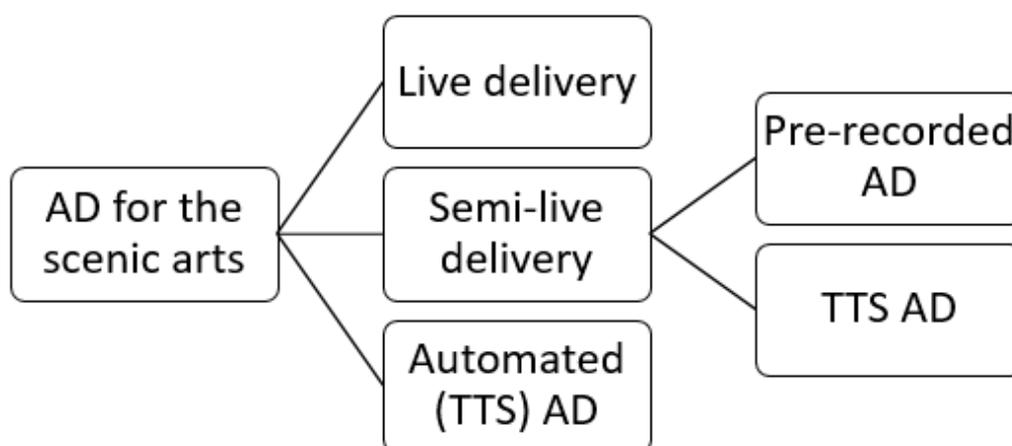

Figure 2. Delivery approaches for the scenic arts

4. Live approach

Live AD implies a synchronous and on-site description that is often applied to theatre, opera, sports and other live events. This approach has been illustrated in most of the previous opera AD studies (Cabeza-Cáceres & Matamala, 2008; Cabeza-Cáceres, 2010; Orero & Matamala, 2007; Eardley-Weaver, 2010; Eardley-Weaver, 2014), as well as theatre AD studies (Holland, 2009; Udo & Fels, 2010). Particularly, live AD for opera has been extensively documented in the Grand Teatre del Liceu opera house context. Several members of the Transmedia Catalonia research group have conducted studies in order to define the method applied in the Barcelona opera house since the 2007-2008 season (Orero, 2007; Orero & Matamala, 2007; Cabeza-Cáceres & Matamala, 2008; Puigdomènech *et al.*, 2008; Cabeza-Cáceres, 2010; Orero *et al.*, 2019). Here, the main advantages and challenges inherent to the live approach will be highlighted.

First, live AD – often preceded by an AI – can adapt to unforeseen events that may occur both inside and outside the stage space. Particularly, amongst the several defining factors of a stage performance (Törnqvist, 1991), some can only be addressed if the AD is voiced live: namely, the two-way nature of communication between the audience and the actors in the theatre,

the inherent uniqueness of every performance and the varying space conditions (shape and size of the stage, etc.). Equally, the live audio describer can suitably react to omissions and changes in the action.

Second, a trained voice talent will be able to skilfully express a certain sympathy with the content of the play. According to Fryer (2016, p. 87), the "supra-linguistic aspects of speech convey meaning through stress, pitch, tempo, dynamic range and, especially, the way the words are segmented, i.e. those momentary pauses and intakes of breath, often indicated by, but not restricted to, those indicated by punctuation". In accordance with related publications (Szarkowska, 2011; Fernández-Torné & Matamala, 2015; Walczak & Fryer, 2018), human voices, as opposed to synthetic voices are still favoured by AD users, although TTS AD has also been deemed acceptable. More precisely, Fryer and Freeman (2014) conducted a test indicating that human-voice AD was more able to elicit specific emotions, such as fear, than TTS AD. Indeed, "emotion was effectively conveyed via the paralinguistic content of the describer's voice rather than the semantic content of the AD script" (p. 105). Moreover, according to Walczak and Fryer (2018), genre also seems to have an impact on acceptability, with drama favouring human-voice AD while documentaries score similar results for natural and TTS AD. Still, the natural effect can be equally achieved by cueing pre-recorded AD fragments. Additionally, synthetic voices could soon reach a point in their development where prosodic features could be (manually) integrated into synthetic voices: Besides the "clean" AD text, further information can be added to enhance the prosodic features of the voice output, e.g. speed, pitch, tone, emphasis, intonation of individual letters, words and sentence segments (my translation, Kurch, 2018, p. 445). For the time being, nuances in supra-linguistic aspects are nevertheless still exclusive to voice talents and trained audio describers.

Following the theme of quality and acceptability, Fryer (2019b) develops a quality assessment proposal for professionals and students alike by exploring overlapping aspects of AD and simultaneous interpreting. The macrocriteria on quality assurance that can be applied to both modalities are accuracy, language, delivery, and synchrony. Fryer (2019a) has further addressed the most frequent mistakes when delivering an AD live, namely omission, action, inaccuracy, reaction, facial expression, vocal delivery, word choice, excess, timing and character identification. The author has also alluded to common recording faults (Fryer, 2016, p. 98): stumbling, breathing, clarity, clipped start, cut off, emphasis, hesitation, lip-smacking, irritating noises, pace, popping (distortion on plosive consonants) and mispronunciation. All of said faults can indeed occur when following a live approach. In turn, pre-recorded AD fragments offer an advantage in the sense that they can be corrected unrestrictedly.

In addition to the listed mistakes that an audio describer or voice talent can make in a live setting, perhaps the biggest drawback for the live approach, also highlighted by Eardley-Weaver (2014, p. 34), is that a live AD is restricted to a limited number of performances. In most theatre venues and opera houses, AD is usually offered in one or two sessions per run of the show. In any case, this would also apply to some semi-live initiatives, since they too are usually offered once or twice per run, and not for every session. Precisely, this widespread restriction goes against principles one and two of universal design (Connell *et al.*, 1997): equitable use and flexibility in use.

5. Semi-live approach

5.1. Pre-recorded AD vs. TTS AD

As Szarkowska (2011, p. 143) points out, the lengthy preparation process and the high cost of producing an AD hinder the wider availability of the modality. In this regard, a semi-live delivery may speed up the process. As previously mentioned, this approach can either apply TTS technology or pre-recorded fragments synchronised live. The present section will put forward some pre-recorded notions and review existing studies on TTS AD. Then the focus will be shifted to the advantages and challenges of TTS and pre-recorded AD. Lastly, software solutions and their functionalities will be introduced and examples in practice will be examined.

Concerning the pre-recorded (human-voice) approach, the AD script is taped in several audio clips and later cued live by a technician. This method is currently being applied at the Teatro Real opera house in Madrid and it is also the prevalent approach in France, where the association Accès Culture provides AD for opera, theatre and dance shows following this very same premise (Resche, 2015, p. 214). Furthermore, AD is recorded by two different voices, one that reads the synthesised surtitles (audio subtitles) aloud and the other that presents the visual descriptions as such. This voice distinction is further explained in section 5.1.

As for academic papers that report on this approach, Di Giovanni (2018, p. 202) exemplifies the application of pre-recorded AD within the framework of the Macerata opera festival in Italy, where AD "was delivered live at the Sferisterio arena: an operator manually launched pre-recorded audio clips and delivered them to the B&PS's cabled seats with headsets". Furthermore, associations such as the Royal National Institute of Blind People in the UK have also raised the possibility of this approach: "for productions that do not vary too much in timings, pre-recording audio description would be a great way of allowing blind and partially sighted customers to attend performances in spaces that don't have the technology"[2]. Overall, while this approach is gaining popularity in practice, becoming the most extended approach in Spanish theatres[3], dedicated research on semi-live AD for the scenic arts is still scarce, some exceptions being Oncins *et al.* (2013) and Di Giovanni (2018).

As anticipated in the previous section, when it comes to proposing TTS solutions – also referred to as synthetic voices or electronic speech –, quality and user acceptability are necessarily the key factors. Preference studies conducted by Szarkowska, (2011), Walczak and Szarkowska (2012), Fernández-Torné and Matamala (2015), and Walczak and Fryer (2018) have obtained similar results in this regard: although human-voice AD was preferred, TTS AD was accepted as an alternative and even a lasting solution.

Understandably, the main advantage of a semi-live approach is its ability to offer AD – and other accessible services all at once – for touring productions at a lower price, as the materials can be used repeatedly. Furthermore, considering the question of prosodic features, Cabeza-Cáceres (2013, p. 331) advises to "avoid a uniform intonation and try to adapt AD intonation to the tone, the context and the genre of the audiovisual product without letting it develop into an emphatic intonation". Nuances in intonation are very much possible following a live or pre-recorded approach, but not always feasible when applying synthetic voices.

Another of the potential drawbacks of semi-live AD is its inability to respond to setbacks or changes in the performance. The ACB guidelines mention other occurrences such as a delay in the start of the play or an emergency in the audience (Snyder, 2010, p. 30), which would be

---

[2]   https://www.rnib.org.uk/practical-help/home-and-leisure/television-radio-and-film/news/accessible-opera. Retrieved September 28, 2019.

[3]   http://www.teatroaccesible.com/es/about/theaters. Retrieved March 30, 2020.

explained by the describer in a live setting. In this regard, the notion of inclusive AD or AD as a part of the creative process (Fryer, 2018; Udo & Fels, 2010) may prove particularly useful. Holland (2009, pp. 176-177) exemplifies the live, unfixed nature of theatre in that even though theatre plays are usually rehearsed and repeated, this is not always true. Admittedly, some directors deliberately allow for some creative freedom. Taking this argument further, if AD is conceived within the creative process, the stakeholders involved can conclude whether a live or semi-live approach makes more sense. For instance, for those productions more prone to improvisation, live AD would, understandably, be a more suitable option. Yet another quandary would be whether the production is going on tour or it is a one-off show. In short, communication between the different stakeholders could prove beneficial when deciding to opt for one delivery approach or the other.

Particularly in the case of opera, "different musical performances have different tempo, so it is possible that the day of the performance when the AD is delivered the conductor decides to alter the tempo and the AD comments do not fit the already timed gaps" (Orero & Matamala, 2007, p. 272). On balance, the challenge of synchrony can be extrapolated to virtually all of the scenic arts.

In addition, although this article focuses on AD, the possibility of live audio subtitling – or the combination of AD with audio subtitling – has also been raised in the performing arts context (see Orero, 2007; Braun & Orero, 2010; Oncins *et al.*, 2013). For semi-live audio subtitling, Eardley-Weaver (2014) proposes yet another cueing possibility:

> The advantage of pre-prepared audio subtitles, similarly to pre-recorded AI, is that they could be made available at all performances. In this case, the audio subtitles could be synchronised with the written surtitles, so that the surtitler could cue both simultaneously (Eardley-Weaver, 2014, p. 60).

As a final theoretical reflexion, a parallel can be drawn between semi-live AD and translation as forms of human-computer interaction (O'Brien, 2012). As Translation Memories and Machine Translation have significantly changed the translator's profession, semi-live AD solutions, as well as other tools for automation can also be expected to flourish in the scenic arts context and beyond. It is therefore a topic that merits further research and perhaps for it to be included in Media Accessibility curricula. In the next section, we look at specific examples of software and mobile apps, with an emphasis on the Spanish context.

5.2. Practical applications of semi-live AD

Moving on to specific semi-live tools and, again, focusing on the Spanish context, Startit is a software solution developed by the company Aptent currently being applied primarily to theatre[4]. This software supports the creation of both AD and subtitles. Indeed, Startit follows the same concept behind Oncins *et al.*'s Universal Accessibility System (2013). In both cases, a dependency of AD on the subtitles is created on the basis that: "AD is usually never delivered when meaningful audio can be heard; in short, the AD is complementary to the subtitles" (Oncins *et al.*, 2013, p. 156). To put it simply, the AD fragments are launched when there are no surtitles or subtitles being displayed. This avoids the overlapping of AD with dialogue, one of the "golden" rules of this accessible modality (ITC, 2000, p. 9).

---

[4]   This software solution is primarily being applied in venues adhered to the Teatro Accesible initiative. As well as renting technical equipment, Teatro Accesible provides several accessibility services such as AD, surtitles, magnetic loops, amplified sound, etc.

Figure 3. Software interface. Source: Startit

The day of the show, there are two possibilities when applying the Startit solution. AD can be delivered through a wireless (infra-red or FM radio) system, if it is available in the venue. Simultaneously, surtitles are projected, usually over the proscenium (see Dewolf, 2001; Oncins, 2015). Alternatively, the audience can access both services from their phones via the Startit app. In both cases, the broadcasted AD is not pre-recorded in segments, but it is voiced through a TTS system. The TTS AD fragments are launched during those gaps when there are no dialogues and thus no subtitles.

Though this is often not the case with Startit, it is worth noting that services synchronously made available by mobile apps are not necessarily limited to AD and (written) subtitles or surtitles. For instance, linguistic accessibility can also be achieved by providing (TTS) audio subtitles (Oncins *et al.*, 2013; Eardley-Weaver, 2014). This may be the case at international festivals or touring theatre productions that are in a language other than the audience's own.

Yet another software solution is the internally available Accuseq from the French association Accès Culture (Resche, 2015). For opera productions, the describers create an Excel file with two main columns: namely one for the AD script and another intended for the translated chant. Once the scripts have been completed, the file is imported to Accuseq. From there, one person records the AD fragments and another voices the libretto translation. The result is a combination of AD fragments and audio subtitles (Braun & Orero, 2010) that, once again, are delivered via infra-red or FM radio systems the day of the show (see Figure 1). In contrast, though the Teatro Real opera house in Madrid also distinguishes the AD and a summarised account of the chant by two different voices, the audio fragments are broadcasted through the mobile app Teatro Real Accesible. To the best of our knowledge, TTS solutions are not being applied to opera AD in any European country as of yet.

In summary, several accessible services can be offered simultaneously by applying the dependency principle, where the same person can launch the (TTS) AD (as well as audio subtitles and pre-recorded sign language interpreting) by having it linked to the subtitles. As for now, many theatres in Spain have adhered to the Teatro Accesible initiative and thus apply the Startit software and app. This implies that TTS AD is the preferred approach for touring theatre productions in the Spanish context. As for opera houses in France and the Teatro Real in Madrid, they also follow a semi-live approach, but they have opted for pre-recorded AD with human voices.

6. Automation

As in other fields within Translation Studies, the possibility of automating AD has been met with great interest. Proposals to automate AD range from the potential of machine translating and post-editing AD scripts (Fernández-Torné & Matamala, 2016) to creating machine-generated AD (Rohrbach *et al.*, 2017; Braun & Starr, 2019). In this section, three exploratory proposals towards AD automation will be synthesised and advantages and challenges for this approach will be put forward.

For the purposes of this article, automation in our scenic arts context implies the launch of AD fragments – be they pre-recorded or TTS – without human interaction. That is to say, there is no need for a technician to manually cue these fragments for every performance. In the Spanish context, the most recent proposal which has taken a step further towards automation has been developed by the research group SoftLab from the Carlos III University in Madrid[5]. Their app Stage-sync combines deep-learning strategies with audio and image-processing techniques in order to synchronise the pre-prepared AD, subtitles and sign language interpreting to the pace of the play. The foundation of the system is that it automatically keeps with the rhythm of the performance and is therefore able to launch the fragments of the accessibility elements synchronously, without human interaction. As of now, Stage-Sync has been applied to the touring musical show The Addams Family.

Yet a different pre-recorded method leaning towards automation is the one proposed by the ACB guidelines (Snyder, 2010, p. 24), with AD fragments being cued not by a technician or the audio describer, but through the lighting of the show. According to this document, this approach has only been put into practice in pilot cases and it does not foresee full automation: "Ideally, a describer monitors the use of these systems so that variations from the original described performance can be incorporated at any particular performance" (Snyder, 2010, p. 24). In short, the lighting of the show is a plausible option to trigger the AD and other accessible services, though further testing regarding user acceptability would be advisable.

More recently, also in the United States context, researchers at New York University are working on a system that utilises reference audio recording from previous performances and an online time warping algorithm to automatically synchronise the AD with the pace of the production (Vander Wilt & Farbood, 2019). As in the case of the SoftLab research group, they have tested their system in a musical theatre production and their results so far show an accuracy of 79-86% within two seconds of the occurring event (Vander Wilt & Farbood, 2019, p. 108). The researchers have also opted for a mobile app for end users to broadcast the AD and thus avoid installation of wireless technical equipment in theatre venues.

Moving on to the advantages and challenges for this approach, the main assets of automated AD are similar to those featured in the semi-live approach: repeatability, reduction in costs, and therefore wider availability. Besides them, the automated approach can be applied in virtually any venue and for every performance as long as it operates through an accessible mobile application, or the venue provides a wireless system. In the first case, there is no need for venues to purchase or install any equipment, as the audience is the one equipped with the app. However, as it is a novel approach, it is again unclear how this technology would respond to highly varying productions. Further runs would have to be conducted to test quality issues such as synchrony in order to boost or replicate this system. Furthermore, both the Stage-Sync app in the Spanish context and the system developed by Vander Wilt and Farbood have

---

[5] http://softlabweb.softlab.uc3m.es/softlab/What.html. Retrieved September 23, 2019.

been applied specifically to musical theatre productions, other genres still need to be further tested. For the time being, this approach is in its infancy and further validation by end users is required.

7. Taking stock: a tentative comparison

Throughout sections 2, 3 and 4, the most salient advantages and challenges for each approach have been discussed. In Table 1, these arguments are gathered and synthesised in order for them to be discussed with blind and visually impaired users in a future preference study.

		Live AD	Semi-live AD		Automated (TTS) AD
			Pre-recorded AD	TTS AD	
Advantages	Ability to adapt to unforeseen events	x			Sometimes
	Vocal sympathy with the tone of the play and prosody	x	x		
	Repeatability		x	x	x
	Cost-effectiveness		x	x	x
Challenges	Recording and vocal faults	x			
	Necessary equipment in theatre venues	x	Sometimes	Sometimes	

Table 1. Advantages and challenges of the discussed delivery approaches

Considering the advantages first, the ability to adapt the AD script to unforeseen events both inside and outside the stage was – and in most productions and venues still is – exclusive to the live delivery approach. Unforeseen events can range from an actor forgetting their lines or choosing to ad-lib to an emergency in the audience. With automated TTS AD, it may also be possible to recognise skipped segments and for the system to resume the action, and alert messages can be pre-recorded and broadcasted. More specifically, the question of whether automated AD could work for highly varying productions does remain open to more trials. As for the second major advantage of live AD: vocal sympathy with the tone of the play (Cabeza-Cáceres, 2013), it can equally be achieved or even improved by pre-recording AD, as the same fragment can be taped several times and the audio enhanced until satisfactory.

In the case of semi-live and automated AD, the advantages are in line with the quest for wider availability. Pre-recorded or TTS AD (and automatically synchronised AD, should this approach become widespread), can ensure repeatability and thus wider accessibility for touring productions. Yet, to date, AD is usually limited to one or two performances per run, no matter whether it is live or follows a pre-recorded approach. Repeatable approaches could potentially enable accessibility any day of the performance, though, at this stage, this is a mere theoretical statement and does not reflect today's reality, namely in Spanish venues. In the same line, cost reduction is a compelling argument for venues and funding institutions alike.

As a drawback for live AD – perhaps to be countered by the ability to adapt to unforeseen events – audio describers and live voice talents might incur in the vocal faults proposed by Fryer (2019a), summarised in section 2. This is not the case in pre-recorded AD and well-made TTS AD: the AD script should phonetically adapt certain words in their written form so that the synthetic voices can produce them, an example being a foreign proper noun. As of today, it must be noted that nuances in intonation are still not as refined neither in TTS AD nor in automated TTS AD.

As for the necessary wireless equipment to be distributed, live AD is often delivered through a radio or infrared system, which can be too costly for some venues. Alternatively, apps such as Startit, Teatro Real Accesible or Oncins *et al*.'s Universal Accessibility System (2013) allow users to control and interact with the accessible services autonomously through their own mobile phones. In this case, venues still have to make sure that the Internet connection at their premises is stable enough for a large audience, or they have to provide a reliable and robust Wi-Fi signal (Orero *et al*., 2019, p. 253).

This classification is, in itself, a starting point that is yet to be tested with audiences with sight loss. It would therefore be worth conducting a user-centered study in order to identify and amend inherent quality, comprehension and engagement issues (Fryer & Freeman, 2014) with each approach. For instance, a reception study devoted to differences in quality between live and pre-recorded AD would provide tangible proof of which approach is more precise.

8. Conclusions

This paper has tackled the scenic arts AD framework from a largely technical perspective. Noting that the illustrated guidelines mainly presuppose that AD for the performing arts is voiced live, an updated framework has been provided on current delivery approaches. In this sense, even though semi-live AD has only recently started being proposed as an alternative by scholars (Snyder, 2010; Oncins *et al.*, 2013; Eardley-Weaver, 2014; Di Giovanni, 2018), both pre-recorded and TTS AD are already increasingly widespread approaches for the scenic arts in countries like France, Spain and the UK.

Advantages and challenges of live, semi-live, and automated AD for the scenic arts have thus been outlined. Further research would have to be conducted in terms of which genre amongst the scenic arts would benefit the most from a live approach, a priori the less cost-effective approach. Again, it is reasonable to assume that unique performances, such as an opera production that does not go on tour, or festival productions would be the most ideal candidates for this approach. Nonetheless, our intention here is not to advocate for one option or the other (Szarkowska, 2011). Rather, different productions have contrasting characteristics that would be best suited to one of the delivery approaches. Understandably, the aim here is to foster accessibility measures and improve AD availability. In summary, our argument is to promote closer communication between all stakeholders.

Specifically, the outlook for TTS AD has proved to be positive so far (Szarkowska, 2011; Fernández-Torné & Matamala, 2015; Matamala, 2016; Walczak & Fryer, 2018). Yet, even though TTS AD has been tested with younger users (Walczak & Szarkowska, 2012), other audiences and genres are yet to be targeted. Older adults and the elderly constitute the principal audience of certain scenic arts such as opera, and their preferences need to be equally addressed. More specifically, TTS AD has been tested specifically within the filmic genre and not so much in the scenic arts context, where we might see some differences in acceptability and preferences.

On another note, universal software solutions that combine accessibility measures such as audio subtitles, subtitles and AD – like Startit, Stage-sync or Oncins *et al*.'s proposal (2013) –

are already a reality for many venues that otherwise may not be able to afford a live delivery approach, or that are unable to install wireless (infra-red or FM radio) systems. Nonetheless, it cannot be overlooked that all the delivery approaches that this article has covered are still often an afterthought, deviating from the universal design principle (Udo & Fels 2010; Romero-Fresco, 2013). The alternative is that AD ought to be part of the production from the beginning of the creative process. From a delivery perspective, the patrons' preferences and the directors' standpoints are expected to be significant when deciding the most suitable approach for each production. Involving all stakeholders in the technological advances of AD for the scenic arts means taking a step forward towards wider availability and quality.

9. References

AENOR. (2005). *Audiodescripción para personas con discapacidad visual. Requisitos para la audiodescripción y elaboración de audioguías*. (UNE 153020). AENOR.

Braun, S. (2008). Audiodescription research: State of the art and beyond. *Translation Studies in the New Millenium, 6*, 14-30.

Braun, S. & Orero, P. (2010). Audio description with audio subtitling – an emergent modality of audiovisual localization. *Perspectives: Studies in Translatology, 18*(3), 173-188.

Braun, S. & Starr, K. (2019). Finding the right words: Investigating machine-generated video description quality using a corpus-based approach. *Journal of Audiovisual Translation*, *2*(2), 11-35.

Cabeza-Cáceres, C. (2010). Opera audio description at Barcelona's Liceu Theatre. In J. Díaz-Cintas, A. Matamala & J. Neves (Eds.), *New insights into audiovisual translation and media accessibility* (pp. 227-237). Rodopi.

Cabeza-Cáceres, C. (2013). *Audiodescripció y recepció. Efecte de la velocitat de narració, l'entonació y l'explicitació en la comprensió fílmica* (Unpublished doctoral dissertation). Universitat Autònoma de Barcelona, Barcelona.

Cabeza-Cáceres, C. & Matamala, A. (2008). La audiodescripción de ópera: una nueva propuesta. In A. Pérez-Ugena & R. Vizcaíno-Laorga (eds.), *ULISES. Hacia el desarrollo de tecnologías comunicativas para la igualdad de oportunidades* (pp. 95-106). Madrid: Observatorio de las Realidades Sociales y la Comunicación.

Connell, B. R., Jones, M., Mace, R., Mullick, A., Ostroff, E., Sanford, J., Steinfeld, E., Story, M., & Vanderheiden, G. G. (1997, April 1). *The principles of universal design*. Retrieved from https://projects.ncsu.edu/design/cud/about_ud/udprinciplestext.htm.

Dewolf, L. (2001). Surtitling operas. With examples of translations from German into French and Dutch. In Y. Gambier & H. Gottlieb (Eds.), *(Multi)Media Translation. Concept, Practices, and Research* (pp. 179-188). John Benjamins.

Di Giovanni, E. (2018). Audio description for live performances and audience participation. *The Journal of Specialised Translation, 29*, 189-211.

Eardley-Weaver, S. (2010). Opening doors to opera: the strategies, challenges and general role of the translator. *InTRAlinea, 12*.

Eardley-Weaver, S. (2014). *Lifting the curtain on opera translation and accessibility. Translating opera for audiences with varying sensory ability* (Unpublished doctoral dissertation). Durham University, Durham.

Fernández-Torné, A. & Matamala, A. (2015). Text-to-speech versus human voiced audio description: A reception study in films dubbed into Catalan. *The Journal of Specialised Translation, 24*, 61-88.

Fernández-Torné, A. & Matamala, A. (2016). Machine translation in audio description? Comparing creation, translation and post-editing efforts. *Skase*, *9*(1), 64-85.

Fryer, L. (2016). *An introduction to audio description: A practical guide*. Routledge.

Fryer, L. (2018). The independent audio describer is dead: long live audio description! *Journal of Audiovisual Translation, 1*(1), 170-186.

Fryer, L. (2019a, March). A catalogue of errors: Lessons for students from professional audio description of live performances. Paper presented at ARSAD (Advanced Research Seminar on Audio Description).

Fryer, L. (2019b). Quality assessment in audio description: lessons learned from interpreting. In E. Huertas-Barros, S. Vandepitte & E. Iglesias-Fernández (Eds.), *Quality assurance and assessment practices in translation and interpreting* (pp. 155-177). IGI Global.

Fryer, L. & Freeman, J. (2014). Can you feel what I'm saying? The impact of verbal information on emotion elicitation and presence in people with a visual impairment. In A. Felnhofer & O. D. Kothgassner (Eds.), *Challenging Presence: Proceedings of the 15th International Conference on Presence* (pp. 99-107). Facultas.

Fryer, L. & Romero-Fresco, P. (2014). Audiointroductions. In A. Maszerowska, A. Matamala & P. Orero (Eds.), *Audio description: New perspectives illustrated* (pp. 11-28). John Benjamins.

Greco, G. M. (2018). The nature of accessibility studies. *Journal of Audiovisual Translation, 1*(1), 205-232.

Hernández-Bartolomé, A. I. & Mendiluce-Cabrera, G. (2004). Audesc: Translating images into words for Spanish visually impaired people. *Meta, 49*(2), 264-277.

Holland, A. (2009). Audio description in the theatre and the visual arts: Images into words. In J. Díaz Cintas & G. Anderman (Eds.), *Audiovisual Translation. Language Transfer on Screen* (pp. 170-185). Palgrave Macmillan.

ITC. (2000). *ITC guidance on standards for audio description*. Retrieved March 31, 2020 from http://audiodescription.co.uk/uploads/general/itcguide_sds_audio_desc_word3.pdf.

Kurch, A. (2018). Produktionsprozesse der Hörgeschädigten-Untertitelungen und Audiodeskription: Potenziale teilautomatisierter Prozessbeschleunigung mittels (Sprach-)Technologien. In C. Maaß & I. Rink (Eds.), *Handbuch Barrierefreie Kommunikation* (pp. 437-454). Frank & Timme.

Matamala, A. (2007). La audiodescripción en directo. In C. Jiménez (Ed.), *Traducción y accesibilidad: la subtitulación para sordos y la audiodescripción para ciegos* (pp. 121-132). Peter Lang.

Matamala, A. (2016). The ALST project: Technologies for audio description. In A. Matamala & P. Orero (Eds.), *Researching audio description: New approaches* (pp. 269-284). Palgrave MacMillan.

Matamala, A. & Orero, P. (2007). Accessible opera in Catalan: Opera for all. In J. Díaz Cintas, P. Orero & A. Remael (Eds.), *Media for all. Subtitling for the deaf, audio description and sign language* (pp. 201-214). Rodopi.

Matamala, A. & Orero, P. (2013). Standarising audio description. *Italian Journal of Special Education for Inclusion, 1*, 149-155.

Mateo, M. (2002). Los sobretítulos de ópera: dimensión técnica, textual, social e ideológica. In J. Sanderson (Ed.), *Traductores para todo. Actas de las III Jornadas de doblaje y subtitulación* (pp. 51-73). Alicante: Universidad de Alicante.

Navarrete, F. J. (1997). Aplicación al teatro del sistema AUDESC. *Integración: Revista sobre ceguera y deficiencia visual, 24,* 26-29.

O'Brien, S. (2012). Translation as human–computer interaction. *Translation Spaces*, *1*, 101-122.

Oncins, E. (2015). The tyranny of the tool: Surtitling live performances. *Perspectives: Studies in Translatology, 23*(1), 42-61.

Oncins, E., Lopes, O., Orero, P., Serrano, P. & Carrabina, J. (2013). All together now: A multi-language and multi-system mobile application to make living performing arts accessible. *The Journal of Specialised Translation, 20*, 147-164.

Orero, P. (2005). Audio description: Professional recognition, practice and standards in Spain. *Translation Watch Quarterly, 1*, 7-18.

Orero, P. (2007). Audio subtitling: A possible solution for opera accessibility in Catalonia. *TradTerm, 13*, 135-149.

Orero, P., Bestard, J., Edo, M., Iturregui-Gallardo, G., Matamala, A. & Permuy Hércules de Solás, I. (2019). La ópera accesible del siglo XXI: nuevos servicios, nuevas posibilidades. *TRANS. Revista de Traductología*, *23*, 245-256.

Orero, P. & Matamala, A. (2007). Accessible opera: Overcoming linguistic and sensorial barriers. *Perspectives: Studies in Translatology, 15*(4), 262-277.

Pfanstiehl, M. & Pfanstiehl, C. (1985). The play's the thing. *British Journal of Visual Impairment, 3,* 91-92.

Puigdomènech, L., Matamala, A. & Orero, P. (2008). The making of a protocol for opera audio description. In L. Pegenaute, J. DeCesaris, M. Tricás & E. Bernal (Eds.), *Actas del III Congreso AIETI. La traducción del futuro. Mediación lingüística y cultura en el siglo XXI* (pp. 381-392)*.* Universidad Pompeu Fabra.

Remael, A., Reviers, N. & Vandekerckhove, R. (2016). From Translation Studies and Audiovisual Translation to Media Accessibility: Some research trends. *Target, 28*(2), 248-260.

Remael, A., Reviers, N. & Vercauteren, G. (2015). *Pictures painted in words. ADLAB audio description guidelines*. EUT.

Resche, S. (2015). Un système d'audiodescription d'opéra pour public de mal- et non-voyants. *Ligeia, 141-144*(2), 212-221.

Rohrbach, A., Torabi, A., Rohrbach, M., Tandon, N., Pal, C., Larochelle, H., Courville, A. & Schiele, B. (2017). Movie description. *International Journal of Computer Vision*, *123*, 94-120.

Romero-Fresco, P. (2013). Accessible Filmmaking: Joining the dots between audiovisual translation, accessibility and filmmaking. *The Journal of Specialised Translation, 20*, 201-223.

Roofthooft, H., Remael, A. & Van den Dries, L. (2018). Audio description for (postdramatic) theatre. Preparing the stage. *The Journal of Specialised Translation, 30*, 232-248.

Snyder, J. (2010). *Audio description guidelines and best practices*. Retrieved September 1, 2019, from http://www.acb.org/adp/ad.html.

Szarkowska, A. (2011). Text-to-speech audio description: towards wider availability of AD. *The Journal of Specialised Translation, 15*, 142-162.

Törnqvist, E. (1991). *El teatro en otra lengua y otro medio* (M. Mateo, Trans.). Arco.

Udo, J. P. & Fels, D. I. (2010). Universal design on stage: live audio description for theatrical performances. *Perspectives: Studies in Translatology, 18*(3), 189-203.

Vander Wilt, D. & Farbood M. M. (2019). Method and system for aligning audio description to a live musical theater performance. In *Proceedings of the 14th International Symposium on Computer Music Multidisciplinary Research,* (pp. 103-111). PRISM.

Walczak, A. & Fryer, L. (2018). Vocal delivery of audio description by genre: Measuring users' presence. *Perspectives: Studies in Translatology, 26*(1), 69-83.

Walczak, A. & Szarkowska, A. (2012). Text-to-speech audio description to educational materials for visually-impaired children. In S. Bruti & E. di Giovanni (Eds.), *Audio visual translation across Europe* (pp. 209-233). Peter Lang.

Whitfield, M. & Fels, D. I. (2013). Inclusive design, audio description and diversity of theatre experiences. *Design Journal, 16*(2), 219-238.

York, G. (2007). Verdi made visible. Audio introduction for opera and ballet. In J. Díaz Cintas, P. Orero & A. Remael (Eds.), *Media for all. Subtitling for the deaf, audio description and sign language* (pp. 215-229). Rodopi.

ⓘD Irene Hermosa-Ramírez

Universitat Autònoma de Barcelona
Edifici MRA 020, Plaça del Coneixement
Campus de la UAB, 08193 Bellaterra
(Cerdanyola del Vallès)
Barcelona
Spain

Irene.hermosa@uab.cat

Biography: Irene Hermosa Ramírez is a PhD student in Translation and Intercultural Studies at the the Department of Translation and Interpreting at Universitat Autònoma de Barcelona (UAB). Her research interests include Media Accessibility, Multimodality and accessibility for the scenic arts, with her doctoral thesis focusing on opera audio description. She holds a B.A. in Translation and Interpreting from the University of the Basque Country and an M.A. in Audiovisual translation from UAB. She is a member of the TransMedia Catalonia research group (2017SGR113), where she collaborates in the RAD project (Researching Audio Description: Translation, Delivery and New Scenarios).