A profile of audio description end-users
Linguistic needs and inclusivity

Elisa Perego and Christopher Taylor

1. Introduction

The majority of end-users of audio description (AD) are referred to variously as people with sight loss (PSL), the visually impaired (VIP) or the blind and partially sighted (BPS), and indeed AD was first envisaged as a service for the benefit of this section of society (Starr in this volume for an account of other end-users of AD and Perego, 2016). This chapter will thus focus on these end-users. Already, terms such as “service” and “assistance” have been questioned, as the “rights” of all people with disabilities are being gradually extended through local and national legislation and public awareness. The notion of “universal design” in all its facets is leading to greater social inclusion and, in the case of AD, to the inclusion of audio description as an integral part of a visual product, be it a film, a television programme or a sculpture in a museum and not an addition, or an afterthought, in a limited number of cases. The chapter will therefore begin by assessing the real and current needs of various sectors of the PSL population, estimated to be in the range of 5% of the populous in Western countries and even more elsewhere (even these figures are optimistic as many PSL are not covered by official registration). We will concentrate on the optimum kind of language to use in description, bearing in mind the need to adapt this language to create the highest degree of “inclusion” while at the same time not pandering to the PSL audience, who have shown their dislike for any kind of compromise or condescension. Inclusion must mean that PSL experience a visual product in as similar a way as possible to a sighted audience and enjoy it, or otherwise, to the same extent. In this sense, a further aspect that must be investigated is the preferences professed by end-users. We will illustrate these preferences based on the most relevant results of an extensive online survey (ADLAB PRO, 2017) conducted within the framework of the European ADLAB PRO project (Audio description: A new laboratory for the development of a new professional profile). The project’s initial stage was devoted to the consultation of both working audio describers and AD end-users through a series of semi-structured interviews that enabled us to gain fundamental information on their background and preferences. Because recruiting end-users is always challenging, within the framework of the aforementioned project we were able to work cross-nationally and rely on a niche group in each of the countries involved as partners (Italy, Poland, Slovenia, Spain, the UK). Thus we managed to shed light on the shared needs,
wishes and dissatisfactions that we consider should be tackled in order to start ameliorating the service, or to make us more aware of what should be done about it. Furthermore, we posit that gaining focused feedback from end-users could be crucial for improvement in a comprehensive and good quality course design.

2. Basic information: what is blindness?


Partial sight and blindness can be broadly defined as a limitation in one or more functions of the eye or visual system, most commonly impairment of visual acuity (sharpness or clarity of vision), visual fields (the ability to detect objects on either side or above or below the direction in which the person is looking), contrast sensitivity and colour vision. Normal vision is recorded as 20/20 in Imperial measures (6/6 in metric), which means that a person can see at 20 feet (six metres) what a person with normal vision can see at 20 feet. Degrees of partial sight and blindness are measured similarly, where the first number in the measure is the furthermost distance at which the person can clearly see an object and the second number is the distance at which a person with normal vision could see the same object. For example, 20/40 vision means that the person can clearly see at 20 feet (but not more) an object that a person with normal vision could see at 40 feet (but not more).

Thus the different groups of blind and partially sighted people are defined as follows (ADLAB, 2014: 9) (see also Bestard and Arias-Badia in this volume).

<table>
<thead>
<tr>
<th>Table 3.1 Definitions of blindness</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Blindness</strong> (severe sight loss) is defined as best-corrected visual acuity of &lt;6/60 in the better-seeing eye.</td>
</tr>
<tr>
<td><strong>Partial sight</strong> is defined as best-corrected visual acuity of &lt;6/12 to 6/60 in the better-seeing eye and is categorised as: mild sight loss – best-corrected visual acuity of &lt;6/12 but better than or equal to 6/18 and moderate sight loss – best-corrected visual acuity of &lt;6/18 but better than or equal to 6/60.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3.2 The WHO data classification of blindness</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Light visual defect</strong>: the upper level is 3/10; vision is at least 20/200 and it is possible to read big print.</td>
</tr>
<tr>
<td><strong>Severe visual defect</strong>: visual sharpness lies between 1/20 and 1/10.</td>
</tr>
<tr>
<td><strong>Partial blindness</strong>: when one has binocular visual sharpness from 1/5 to 1/20 and a range of vision between 50° and 10°. The counting of fingers can be done from a distance of three metres and as far as reading is concerned one can read letters the size of headlines.</td>
</tr>
<tr>
<td><strong>Near total blindness</strong>: one can perceive light; sometimes masses, volumes or shapes can be perceived, fingers can be counted at one metre or less, hand movements at five metres, binocular visual sharpness is 1/50 or less and the range of vision is five degrees.</td>
</tr>
<tr>
<td><strong>Total blindness</strong>: There is no visual perception whatsoever.</td>
</tr>
</tbody>
</table>
However, there are no definitive guidelines adopted everywhere. For example, the World Health Organisation provides the following classification and this would seem appropriate for the scope of this chapter:

What emerges, however, from all sources is that there are various degrees of blindness. PSL range from ten-year-olds and sixty-year-olds who have been blind from birth, to twenty-year-olds and eighty-year-olds who have become blind through illness or accident or old age, to the many who have slight or severe partial blindness. In addition, it is important to stress that more than 80 percent of people suffering sight loss are over 50, with a large proportion being over 65. This group of people with sight loss due to old age are not always included in official figures. Furthermore, their numbers are likely to increase given the ageing of the population. Projections in the UK, for instance, suggest that the number of PSL in the UK will rise to 2,250,000 by 2020 and to 4,000,000 by 2050. The needs of these groups are different and until recently it was not possible to provide different ADs to suit all users, and describers had to provide, consciously or unconsciously, a “harmonised” version (see Poli, 2010 regarding a similar approach to subtitles for the deaf and hard of hearing) which would hopefully be useful to all the varied sectors. With the current digital revolution and technological advances (Orero in this volume), it is now at least theoretically possible to provide, on separate channels, alternative AD versions tailored to the needs of different groups.

3. Current issues

3.1 The importance of AD

Data gathered through the ADLAB PRO (2017) online survey can show us the stance of end-users on several crucial AD-related issues. To start with, how important is AD for blind and visually impaired people? Previous discussion (ADLAB, 2012: 53) showed that at least up to a decade ago there was still some diversity of opinion – and there still is. For many people today, audio description is paramount when it comes to watching television, going to the cinema, attending plays, visiting exhibitions, and so on. Returning to the survey, when asked to rate AD on an importance scale of one to five, with one equaling “not at all important”, institutions in Germany and Italy, for example, assigned a high score (five out of five). However, some do not opt for this accessibility service (or do so to a lesser extent). In Poland, for instance, “some people claim that the blind have more important things to do and AD is a waste of time and effort”, though this is not, of course, a universal opinion. The only Portuguese informant in the survey claimed that jobs and better education should be prioritized over AD. Finally, there is general agreement – coming from a Belgian respondent – that “five is what we should aim for”, but this essentially depends on the individual, and a reliable general consensus does not exist at the moment. What we cannot forget, however, is that its introduction in the 70s in the US generated an overall positive reaction to the service and provided increased television access and increased cultural independence to visually impaired people, thus improving their quality of life by making programmes more interesting and informative, as well as by increasing blind and visually impaired users’ comprehension and sense of immersion (Perego, 2016).

3.2 AD-related language

It has already been established that audio description constitutes a linguistic genre in its own right (Mazur in this volume; Piety, 2004; Arma, 2011). This is particularly the case with screen translation where the use of language is dictated partly by the needs of a particular
A profile of audio description end-users

41

audience and partly by the technical constraints of having to fit descriptions into the gaps between dialogue and other parts of the soundtrack. The first aspect to strike the uninitiated reader of an AD script is the almost exclusive use of present tenses in declarative sentences, given that the describer is working in real time and providing an essentially objective picture of events. For the same reason events are related solely in the third person in simple, paratactically connected sentences unadorned by complicated subordinate clauses (ADLAB, 2014). Attention to the choice of vocabulary should result in the use of a varied and vividly colourful lexis aimed at rendering the description interesting and easy to listen to. Non-finite phrases at the beginning of sentences, in theme position according to Hallidayan functional linguistics (Halliday, 1994), are used with more than usual frequency. Thus we find past and present participles beginning sentences (Sitting at a desk, . . .; Wearing only a dressing gown, . . .; Stopped by a red light, . . .). Isolated noun phrases are also common in providing concise information as to location (the desert), an object (a gun on the table) or time (midnight). A more controversial element is that of the use of degrees of appraisal, that is subjective judgements, opinions and assessments. Different schools of thought exist regarding this question and will be discussed later.

With other areas of audio description, for example, museum AD, the features mentioned previously are much less relevant, but others come into play (Perego, 2018). The time constraints are of a different nature, namely those of providing the necessary detail, for example in describing a Roman statue, without moving into information overload. The use of tenses other than the present is of course required in dealing with historical artefacts. Simple sentences are still recommended but not obligatory, especially if the AD is given live by, for example, a trained museum guide and interaction with the visitors is possible in order to clarify any unclear explanations. The choice of lexis is still important in order to make the description acceptable: the use of technical terminology may be necessary and the describer needs to walk that fine line between essential detail and comprehensible delivery.

3.3 The end-user perspective

The previous preamble on the observable language of AD leads on to the wider question of why such measures were initially adopted, whether they work and how they are received by the end-user. The slogan “nothing about us without us”, first used in a political context, for example, by Black South Africans and later as a clarion call for all disabled communities (see Charlton, 2000), expresses the desire on the part of the blind and sight impaired community to have a say and a decisive say, in how audio description is conceived and prepared. From research carried out during the ADLAB PRO project (2017), it transpires that most of the end-users contacted in the survey (N = 100, 55 male, 74% totally blind, 67% with declared visual memory) use AD for watching films and television programmes or on visits to museums. Consequently, most of the feedback refers to these areas. In fact, AD services do not normally cover other areas (theatre, opera and live events) especially in some countries, because some cultural products and contents are too complex to be made accessible and they require specific expertise, experience and strategies. The comments of some anonymous VIP respondents to the ADLAB PRO survey also point out that providing the service is often not enough:

The biggest problem is that we don’t have information on the availability of audio description. It can’t be used if its existence is unknown. This information should be improved and favour (positive discrimination) accessibility to all kinds of audio descriptions.
Or again, on the same lines:

Often the AD is not signalled and can’t be activated without the help of a sighted person. It’s hard to find AD for a lot of the situations mentioned here. In my opinion, the best one is museum AD.

In the worst case scenario, “AD is poorly publicised and completely non-existent”. These comments raise a thorny issue for the blind community, who are often not able to exploit what is in place for lack of advertisement and dissemination, a theme that should be urgently tackled internationally, in spite of the many efforts that have been made to bolster the service and its usability.

We can now move to AD appreciation. In order to discover what end-users from various countries like and dislike about current AD practice and thus to identify the areas where we could concentrate our focus in the training of future describers, PSL respondents in the ADLAB PRO survey were asked to express their appreciation of a list of varied aspects of AD. The most appreciated aspect was the narrative effectiveness of the AD, the way the description helped the users to follow the story. The way the AD helped in understanding the visual product and the clear and easy to follow sentence structure were also stressed. When asked about negative aspects in relation to the quality of the information in the AD the most disliked items included poor synchronisation with the dialogue and sound effects, the lack of cohesion between images and dialogue and the presence of significant omissions. Narrative effectiveness implies the ability to convey the event and yet maintain distance between the event and the reader, as well as to position the characters successfully in the spatio-temporal settings they inhabit.

As regards the language and style adopted in the descriptions, end-users showed themselves to be particularly sensitive to certain aspects and unwilling to accept, for example, complex lexicon and syntax. They pointed to lack of comprehensible vocabulary, too much repetition and the fact that the descriptions were not engaging. The lesson to learn would thus seem to be that words need to be chosen accurately to be able to describe the material in question in a precise way. Participants found it disturbing when the AD language was not suited to the product being described. This could be linked to the social and educational functions of AD, in the sense that end-users can augment their knowledge about the reality they cannot see.

Complex vocabulary and syntax are usually an effective way out of the lack of time and space to fit them between dialogues. In English, a noun sequence is shorter, but it could also be more ambiguous and cognitively taxing (e.g., is a “wine glass” a glass of wine, or is it a type of glass that is used to drink and taste wine?). The same applies to jargon, for example a word such as “jemmy” which saves the describer from having to use a longer definition (a long metal bar with a curved end, often used by thieves), but could well result unclear for the uninitiated. Repetition might not be so engaging, but it surely contributes to preventing confusion and it might even have a didactic function.

As for complex syntax, it might be more disrupting in film AD than in art AD. In the former case, it is usually not listened to more than once: it is consumed quickly, on the spot and might be quite challenging if the sentence is too long. In the following excerpt taken from the AD of The Shipping News by Di Langford (September 2002), the initial sequence of simple coordinate clauses eases comprehension of a long message squeezed between dense dialogues. However, the second part of the AD adds complexity through a reduced relative,
A profile of audio description end-users

itself further expanded by a relative clause and an adjunct – to mention just some possible difficulties:

A cabin window is shattered with a heavy wooden beam. A woman is thrown brutally aside, a man falls, his throat cut. The wreckers work quickly unloading the cargo from the sinking ship, which lies out in the bay, half under water. Flames leap from a stone cairn as they unload their booty on the shore.

The 56-word text consists of four sentences whose average length is 13 words (SD=4.3): the lexical density (i.e., the ratio of grammatical words to content words) equals 57%. Nouns are the dominant word class (34%), adjectives are 4% of the whole text. The readability score based on the Gunning Fog index is 7.03. Listeners are not spared a very low frequency item, “stone cairn”, a noun sequence referring to human-made stacks, mounds or piles of rocks which can take different forms and that are found in cultures around the world and used for many different purposes, including serving as monuments, burial sites, navigational aids and so forth.

The survey also asked about technical aspects of the ADs the participants had followed. If the understanding and enjoyment criteria of the ADs were deemed positive, then it seems the technical competence was taken for granted, but when this was not the case, criticisms were made of poor timing, the defective mixing of sound and poor vocal skills on the part of the describer.

Not all audio describers, in not all countries, are in fact trained to be both script writers and voice talents, which remain separate competences. A good script must be fluent and “playable”. Good vocal delivery is in turn crucial, especially to ensure the comprehension and listenability of end-users (e.g., Glenn et al., 1995). Listenability is a top quality in all products that include a spoken channel – it is of course paramount in radio broadcasting. Listenability is a measure of the ease or pleasure in listening to something, such as music, or even an oral text. Audio descriptions should always display this quality in order to be enjoyed and understood as widely as possible. It is difficult to list all the features a listenable AD should contain. However, we will here mention the need to deliver it at a slow pace and volume, with a clear and engaging tone of voice, as well as the need to remove all interfering sounds that might impede the production of a clear final soundtrack. To produce a listenable text, a voice talent should use an oral-based style, not read too fast, convey engagement, adapting if necessary to the style and genre of the source text. The selection of an appropriate voice for aural reproduction, the calibration of the volume and the quality of the recording and the adjustment of the rate of delivery are further key elements (for a practical overview, see ADLAB PRO, 2019).

Readability and listenability are different concepts. Because a film AD cannot be re-read, the describer should make sure that the message is “instantly intelligible”, stimulating, attention-grabbing. So, professionals should always take an audience-centered perspective when audio describing, in order to convey an engaging message, which is what is required by end-users.

Finally, in the survey, participants were questioned on the textual aspects of the ADs. Among those aspects that gained a positive response, particular reticence was reserved for the lack of effective narrative, the failure of the AD to tell a story and maintain the interest of the listener. There was also mention of a lack of an appropriate level of detail and of background and contextual information.

When dealing with film AD, maintaining coherence and an effective narrative is in fact challenging, sometimes impossible. Long dense dialogues might cause the AD to be perceived
as non-linear. For this reason, some end-users maintain that they prefer to watch films without AD. This can in fact spare the user the task of construing a whole meaning from the AD and the original soundtrack. A task that might not be so easy for every listener – especially for those who need simplified linguistic input to ensure understanding.

The final picture that emerges from the ADLAB PRO survey is that, according to the perspective of the end-users, textual and linguistic skills seem to be the critical features, indicated by their explicit appreciation of AD aspects that have to do with the way language is used to deliver information. As mentioned earlier, end-users favour a clear, coherent and engaging AD narrative that helps them follow and enjoy the product thanks to simple syntax and well-chosen comprehensible vocabulary. As we will see in the next paragraph, Plain Language or Easy Language could contribute to making a difference in this respect.

4. Plain and/or Easy Language in AD

Broadly speaking, Easy Language is a minimal language variety with maximally enhanced comprehensibility for an audience with cognitive and intellectual disabilities (Maaß, 2020; Perego, 2020). Resorting to paralanguage, too, Easy Language reduces the message to the minimum (content- and language-wise) in order to enhance its comprehensibility and increase its usability. On the other hand, Plain Language covers the area between Easy Language and standard language and it is meant to make specialised messages readily available for all, including non-specialist users. Both Easy and Plain Language are user-centred, reader- versus text-oriented varieties.

Interestingly, the responses and the direct comments of the respondents to the ADLAB PRO survey (2017) pointed to the preference for a use of language and style that seem to remind us of Plain Language (e.g., Cutts, 2013; Maaß, 2020) and of those language choices that are preferred any time accessible communication is in question (Perego, 2020).

As mentioned earlier, Plain Language is known to be a simplified language variety that enables non-experts to access a text and understand it. Up to very recently, the idea of simplifying AD has been unexplored: AD is supposed to be vivid, descriptive, semantically rich (Snyder, 2008; Remael et al., 2015) – in other words, difficult to access for some people. Thanks to the EASIT project,3 the simple though cutting-edge effort of simplifying the language of audiovisual translation and of other audiovisual products and services to enhance their usability started to become real and consistently researched (Bernabé & Orero, 2019; Matamala & Orero, 2019; Perego, 2020).

At the time of writing, the EASIT project is nearing completion, but the idea of introducing language that is easier to understand into AD, but also in audio subtitling (AST) and audio introductions (AI), seems to be gaining traction and it could certainly meet the needs of at least some blind end-users. Furthermore, it may well be useful for a wider category of users, ranging from those who are cognitively impaired to those who wish to learn new words and structures, or are too young to appreciate vivid and close to literary descriptions. Relying on shared recommendations to produce material that is easy to understand (e.g., Inclusion Europe, 2009; ILSMH, 1998; IFLA, 2010) can be helpful in producing or adapting usable texts. We will observe some examples here.

As mentioned earlier, depending on what we audio describe, an AD can be more or less complex. In films, it normally consists of short simple sentences that exploit the semantic richness of the chosen words to convey as much information as possible in few seconds – those available between dialogues. In the sentence “The brunette rushed after him”, for instance, the single-word noun “brunette” refers to a woman or girl with dark brown hair: it conveys much
information in just eight characters. The same applies to the verb “to rush”, which specifies the manner of carrying out the movement – with urgent haste. Syntax-wise, the sentence is simple: it consists of three constituents, a subject, a verb and an adjunct (or adverbial element). Screen ADs are not always so simple in terms of number of words. Furthermore, if we analyse this sentence, we realise it can present some difficulties. The noun “brunette” and the verb “to rush” might be too semantically rich for some users. So, even if the syntactic structure of the sentence is simple and the length of its constituents is limited, the words used might not be easy to understand (or usable) for some. Choosing common words is a helpful solution when we need to produce an easy-to-understand (E2U) AD or to adapt an already existing one. The inevitable compromise will be losing semantic nuances which might not be so subtle. In an E2U AD, “brunette” could be translated into the simpler “girl” or “woman”, thus however failing to convey the fact that she has brown and not blond hair. This is a detail that has a connotative meaning but that can however be delivered at a different point in time in the film, if relevant, thus resorting to a sort of “compensation technique” – to use the terminology of Translation Studies.

As far as ASTs are concerned, very similar strategies can be applied. However, to enhance their comprehensibility one might also consider briefly introducing their presence in order to offer the necessary background to the user and to prevent confusion. A short text in E2U language – written on screen and read aloud – could suffice:

This film includes audio subtitles.
Audio subtitles are subtitles read aloud.

Furthermore, in fully audio subtitled films, using multiple, slightly nuanced voices for different characters might help users to enjoy the product more while understanding and remembering it better, while on a more technical level, one might consider enabling the user to manage the volume of audio subtitles separately from the volume of the original sound.

Moving to audio introductions, an extensive selection of information is often necessary to produce or translate E2U texts. To exemplify this, we will focus on some excerpts taken from Louise Fryer’s audio introduction to the film *Slumdog Millionaire*. The introductory excerpt is very rich in details:

Welcome to this audio introduction to *Slumdog Millionaire* directed by Danny Boyle in 2008. It’s based on the novel Q&A by Vikas Swarup with screenplay by Simon Beaufoy. It has a running time of two hours and stars Dev Patel and Frieda Pinto.

We might in fact wonder whether such details are needed in an E2U scenario, where we should prioritise what is really central and easier to remember. With this general principle in mind, we could rewrite the text following some principles, such as adding the word “film” to give a clearer context for some users and then repeating it or using an anaphoric pronoun (“it”) to give coherence and linearity to the text without repeating the film title. Furthermore, if maintaining the film director’s name and production date seem crucial pieces of information, keeping the name of the author of the novel on which the film is based, or the screenwriter’s name, is possibly too taxing. Additionally, we can also substitute the verb “to star” by a longer but clearer sentence type (with a subject, a copular or linking verb and a complement) and introduce the straightforward noun “actors”. Finally, starting each sentence on a new line is a good practice, especially if we think of using audio introductions also in the form of printed documents that can be read rather than listened to, based on the situation or the preference of
the user. If we do so, adding non-linguistic elements – such as a picture of the film poster, or of the actors – would improve the usability of the written document and would comply almost fully with most Easy Language principles (e.g., Inclusion Europe, 2009; IFLA, 2010; ILSMH, 1998; Cutts, 2013). Not all users like, can or wish to consume texts aurally, especially if these users have cognitive or intellectual disabilities, but they are not blind. For this reason and for these audiences, when working with audio introductions, creating a written version and making it available online – both as an accessible PDF or as a printable brochure – might be a good idea.

Based on the previous discussion, a possible easy-to-understand audio introduction version could be the following:

Welcome to this audio introduction to the film *Slumdog Millionaire.*
The film was directed by Danny Boyle in 2008.
The film is based on a novel.
It lasts two hours.
Its main actors are Dev Patel and Frieda Pinto.

More descriptive and detailed passages could require further analysis and modifications. We can focus on the following:

Their home is a shack – one of thousands cobbled together from sheets of corrugated iron, painted wooden planks and scavenged cardboard in the teeming Juhu slums – sandwiched between an airport runway, a vast rubbish tip and the railway tracks at the edge of the city. Rubbish litters the slum’s narrow passageways. It’s a patchwork of poverty but rich and throbbing with life – and full of colour – walls are daubed in blue and yellow, the women wear patterned saris of every hue, there are fruit and vegetables for sale at outdoor stalls and painted statuettes of gods set up in home-made shrines.

Though beautifully written (and read by Fryer) and close to literary style, this text needs to undergo major simplification processes to become easy to understand. First, the text is rich in vivid yet uncommon words: *shack,* which is a primitive shelter; *cobbled* meaning paved with small stones. But also *corrugated,* *planks,* *scavenged,* *teeming,* *Juhu* (which is an upmarket neighbourhood of Mumbai), *slums* – or poor areas in a city; *rubbish tip,* *tracks,* *edge,* *litters,* *passageways,* *patchwork,* *throbbing,* *daubed* (meaning carelessly painted), *patterned,* the cultural borrowing *saris,* but also *hue* for colour or shade, *stalls,* *statuettes* – a derived word exploiting a diminutive suffix and to conclude *shrines.*

The text is long (101 words, but only three sentences). Its readability is too low – with a Gunning Fog index telling us that a reader would need over 16 years of schooling to be able to understand the text properly. Translating it into an E2U text will require an intensive compression work that will also cause the loss of the most colourful details. The following E2U version is just a proposal:

Their home is small and made of iron sheets and cardboard.
Their home is in a very crowded area of the city.
It is located between an airport, a rubbish tip and the railway.
Rubbish is all over.
Poverty is all over.
But, a lively atmosphere comes
from the colours of the walls;
from the women’s traditional dresses;
from the fruit and vegetables for sale;
from the painted statues of gods.

The text is still informative, but shorter (71 words and six sentences) and the level of instruction needed to understand it is now approximately half: seven years – not 16 any longer. The idea of the slum, of the poverty and its liveliness are not completely lost and an engaging tone of voice will help to compensate for what had to be eliminated.

5. Non-discriminatory language

The first area of discrimination lies not so much in the actual language of AD as in what it conveys. It has been ascertained, through both personal communication and through the opinions of staff in blind organisations or institutions, that persons with sight loss hate being patronised. This is the case in all situations but with regard to AD, they do not want what they consider to be unnecessary help in understanding or following a story-line or documentary information. This can at times put the describer in a difficult situation. Having to decide how much information to put in the description requires intuitive skill and experience. Following to the letter the “say what you see” philosophy when describing film is the surest way to avoid offending the discerning end-user (see Snyder, 2008, 197). However there are times when the describer might feel it necessary to give the blind person some assistance. On these occasions the mere description of the characters, the location and the action may not provide all the information needed for a complete understanding. There may be subtleties hidden within the dialogue and the images that are visible to the sighted viewer, but are inevitably not available to PSL.

Taking as an example the BBC’s television adaptation of Charles Dickens’ novel Little Dorritt (Dickens, 2008 [1857], p. 78), the AD begins with Amy (Little) Dorritt leaving the Marshalsea Prison in London where her father is incarcerated for debt (see Taylor, 2017: 159). The door to the prison is opened for her by John, a prison worker. The official AD describes the young man as follows: “a neatly dressed man in his 20s with short brown hair unlocks a small wooden door”. The decision to restrict the description to three elements (neatly dressed, in his 20s, short brown hair) is as acceptable as it is open to discussion (see Taylor, 2020: 93). No mention is made of the actual clothes he is wearing, his gait, his height or his facial expressions. Given that time constraints dictate the amount of detail it is possible to describe and that too much attention to detail leads to information overload, the AD is concise and well-worded and from this point of view can be considered a perfectly good description, as in the rest of the series.

However, in terms of the subtleties already mentioned, the point that the sighted audience might perceive and the PSL would miss, is that John’s body language and gaze indicate that he has a more than passing interest in Amy, a fact that will be confirmed later in the story. The dilemma facing the describer is precisely whether to mention this fact somehow, or leave it to the blind viewer to work this out from the simple, though nuanced, verbal exchange “Good morning, Miss Amy” – “Good morning, John”.

Film technique can also provide the describer with choices to make regarding subjective or imaginative intervention. For example, all films are to some extent dependent on well-defined structures and conventions. More or less explicitly, they tell a story and the audience need to be able to reconstruct the story being told. They do this by creating mental models (see Remael et al., 2015: 13) of who is involved, what is going on, where it is happening and in what time frame. This information is constantly updated as the action progresses through a series of
clues that explain when and why spatio-temporal settings change, when new characters are introduced and how events are linked. These form the basis for the comprehension of the story on the part of sighted viewers. Audio describers need to make sure that their AD contains all the necessary clues to allow the blind and partially sighted audience to create a context for every event taking place in the story (op. cit. 14). Other elements may assist in this process of contextualisation. There can be reference to other films made either explicitly or implicitly, both verbal and visual, and this may be important to understanding the plot. The level of world knowledge on the part of the describer and what the describer considers to be the level of world knowledge on the part of the audience, will influence the AD choices. There is a fine line between judicious intervention and unwanted interpretation, but sometimes a little extra help could be considered justified. This remains a controversial issue among different AD schools. It comes down to whether interventions of this type, however subtle, can be considered helpful or discriminatory.

Returning to the way John looks at Amy in the aforementioned scene from *Little Dorritt*, much can be gleaned by a sighted audience from gaze alone. A typical cinematic convention concerning gaze, that is used in many filmic texts, is that of a character seeing something surprising, or threatening or even pleasing. The close-up technique shows the gaze vector honing in on an object or a scene outside the audience’s field of vision. In a typical case, a character enters a room and the eyes fix on something off-screen, for example a dead body, which is then brought into focus. At the beginning of this sequence, the describer can focus on the facial expression of the character and make a descriptive reference to the eyes “her eyes dilate and remain wide open” or opt for a more subjective approach “she looks with horror at something in the room”. As the process of aural information takes more mental effort than that of visual information (Fresno, 2014), a little extra assistance for the blind audience can be justified. This is especially the case when what is seen by the character (the dead body) is not actually seen by the audience until much later in the action, for example in a flashback as the character recalls the incident. In this case a little explicitation for the blind audience is doubly justified. In the flashback, the extra clue “she sees again in her memory the dead body she saw earlier”, is of more assistance than “the eyes dilate and remain open as she sees a dead body”, as the “say what you see” philosophy would recommend. The sighted viewer can take in both the visual and the semantic elements simultaneously. The blind audience would have to process the visual information (in aural form) to then extract the semantic information, which would take longer. Experiments by Fresno et al. (2016) showed that end-users generally prefer the former approach (see also Vercauteren in this volume).

Another common cinematographic device designed to create suspense and tension in the viewer is the use of the “anchor”. In some ways this creates the opposite effect to that described earlier. In this case an object or a wording is left hanging in the air, as it were, apparently having nothing to do with the story as it unfolds. This object or utterance will prove to be a crucial element much later in the plot. It will reveal who the perpetrator of a crime is or point to evidence that will only be uncovered further down the line. For example, a torn necklace found at a crime scene at the beginning of a film only becomes significant when, much later, a detective sees the same object in a photograph worn by a woman apparently above suspicion. In an episode from the seventh series of the American television drama “Homeland” (2018) an Iranian man being interrogated takes a cigarette from an unusually patterned packet, presumably an Iranian brand. He is then released but several scenes (and days) later the interrogator finds a similar cigarette packet in a waste bin, providing an important clue towards the man’s arrest. In cases like these it is presumed that the sighted viewer will, consciously or unconsciously, store this visual information and activate it at the later juncture. The question
facing the audio describer is whether to just describe this detail at point A and point B, thereby providing the same experience as the sighted person has, or give a little assistance to the blind person in order that they connect the two scenes. Two schools of thought at opposite ends of the objective/subjective spectrum would suggest respectively that no help should be given or that assistance is required to make sure the PSL perceives these important clues. The reaction of the PSL themselves is also mixed, some preferring the “no help” stance, others appreciating a little assistance. Consider the following descriptions of the two scenes:

a) the Iranian takes out a yellow patterned cigarette packet and proceeds to smoke;
b) Saul spies a patterned cigarette packet in the waste bin by the desk.

a) the Iranian takes out an unusual yellow patterned cigarette packet and proceeds to smoke;
b) Saul spies a cigarette packet in the waste bin by the desk. It is the same unusual brand as the Iranian smoked in the first interview.

Which of the two versions would be preferred by PSL? As explained previously, opinions differ but according to the research carried out during the ADLAB PRO project, some enlightenment was forthcoming. One of the many questions put to the blind and partially sighted end-users was concerning the difficulties encountered and what they disliked about the ADs they listened to. Among the answers, 44% cited the fact that “the audio describer gives their own opinion and prevents you from drawing your own conclusions”. 32% expressed similar views with “the audio description does not give you independence”. This certainly ties in with the widespread view among end-users that they do not like to be influenced, spoon-fed or patronised. However, it would also seem that just over half of the sample were not disturbed by describers “giving their own opinion”, or simply did not notice it. Similarly, when asked what they liked about the ADs, no responses were found that expressed an appreciation of objective, opinionless and non-assistive descriptions, though this might have been taken for granted.

Involving users more in the AD process could at least partially overcome this problem. For instance, in the 2012 ADLAB Report on user needs analysis (p. 53) the Belgian institution Brailleliga reported that their “work is based on contacts with visually impaired people and the personal experience of our collaborators”. Indeed, users have been involved to some extent in the process in all project countries. In Germany blind persons normally work with the describer teams and give invaluable feedback on the effectiveness of the product (Perego & Benecke, 2014), as they do at Senza Barriere in Italy (Perego, 2017). In Portugal, a blind member normally assists the team doing research. However, the involvement of blind users in the AD process within each country is not fully established and is inconsistent: highly developed in some places, non-existent in others, but when involved, blind patrons are not usually implicated in the actual making or script-writing of ADs, but rather in the editing and evaluation processes.

When audio describers themselves were questioned on what they found most difficult, they included in their responses the following elements: “providing the listener with a way of ‘understanding’ what is described” and “using language that is suited to the audience”. The first quote here would suggest that describers, apart from in the general sense of needing to explain the visual content clearly, find difficulty in balancing objective and subjective approaches when “understanding” may require more than “say what you see”. The second quote would suggest that the objective/subjective dialectic may depend on the audience and their age, their background and their levels of world knowledge.
6. The AD profession seen by end-users

It is still not clear how end-users consider the AD profession. The work of audio describers is not necessarily known outside professional circles and if it is known, there may be some confusion about its nature. With the ADLAB PRO (2017) survey we set out to gain knowledge also on how this profession is seen both from the inside and from the outside world and we asked respondents whether in their opinion the work of the audio describer is an “art” (meaning an innate talent that can be refined on the job) or a “craft” (meaning an activity you learn, the result of specialised training) (cf. Haig, 2005). Most end-users agree it is both, or they tend to see it more as a practical type of job (a craft vs. art), thus demonstrating a good degree of awareness of this profession and its implications (ADLAB PRO, 2017: 38). In terms of status or professional role, audio describers are compared to audiovisual translators, a stance that demonstrates that end-users consider the translational and the technical aspects of the AD profession very relevant (ADLAB PRO, 2017: 39). The profession on the other hand is considered demanding but not prestigious (as is the case in other AVT areas, such as in dubbing, Perego & Pavesi, 2006) and it is thought to be unknown to the general public – which suggests that it should be promoted and given more value and visibility. The difficulties encountered in accepting this as a profession might well come from the fact that too many people still do not know the practice and its importance. On the other hand, data show that most respondents think that the work of the audio describer is well-known in the inner circle. All respondents have a similar stance when they are asked whether they consider the work of audio describers creative and satisfying: end-users believe it is a creative and a satisfying job for most professionals. Almost all the sample agrees on the social usefulness of AD, as expected.

7. Future expectations of the end-users

The final part of the survey focused on the future expectations of the end-users. Overall most respondents wish that more ADs will be offered in the future. At the time of responding to the survey (2017) the quality of AD seemed of secondary importance, at least for some, particularly in contexts where not enough AD was offered. The hope that the overall textual and linguistic quality of AD can improve was in fact selected only by 12 out of 100 respondents (with the odd comment pointing to the fact that quality needs to improve if the quantity increases), while ten respondents pointed out that they hope that the overall technical quality of AD will improve. The need for proper synchronisation was rightly highlighted.

However, attention should be focused on the free comments that we gathered, and that offer a more highly tuned idea of the real needs of this category of audience. A comment that struck us regards the importance of describing images when only images are on screen and nothing else is going on (e.g., music or speech). The anonymous commentator claims that AD “is not so essential when the images are accompanied by conversations but it is essential when only visuals are shown because in that case you miss everything”. A star is normally shown on screen in subtitling for the deaf and hard of hearing to signal that there is no technical problem, but there is just silence. A similar stratagem should be considered for AD, too.

On a similar note, more attention to genres should be shown. One comment points out that in the case of documentaries “it is very annoying that half of [them have] subtitles and half [are] read. Instead of subtitles there could be the description of animals, environment, buildings”. So much is taken for granted that it would be useful to draw the attention of describers to these aspects of some particular AV genres in order to make them really accessible and
useable. In this respect, a proper team including a VIP member would certainly make a quality difference.

A further interesting comment points to the fact that a “unified European audio description project, in which text scripts would be produced in English and then translated in all the other languages of the European union” is particularly interesting, though complicated. We agree with the respondent, especially because they might not be aware of how much the translation issue has been and still is debated (e.g., Jankowska, 2015; Remael & Vercauteren, 2010). The idea of saving work and of providing one good product and using that template to make others equally good certainly makes sense. However, the difference between languages and cultures seems to render this practice problematical, especially when some types of AD are involved. The very strict time constraints of screen AD, for instance, do not make it easy to translate, especially from a concise language such as English into languages which tend to expand the source text, such as Romance languages. On the other hand, the situation is more flexible with art AD, or AD with fewer constraints, which can tolerate expansions more freely.

Finally, some comments point to the workflow and to a more regular provision of quality material. As pointed out by some respondents (see also Perego & Benecke, 2014), AD should become a routine/standard provision in all or most forms of artistic expression. I would hope that the promotion/marketing of AD is increased/improved so that more people living with sight loss are aware of the availability of AD. I would hope that the training and professional development of describers continues and that such development continues to involve close consultation with AD end-users”. We certainly agree and the scope of ADLAB PRO was exactly this, along with the credo that “a good AD can only be achieved through good collaboration with the screenwriter of the production and someone who has received a specific audiovisual training.

An important step forward to promote the practice could certainly be made in terms of lobbying power. In the 2012 ADLAB Report (p. 54) it emerges that governments at all levels should be approached to support AD initiatives with, for example, financial subsidies. Authorities, broadcasters, production companies, distributors, organisers of media events should all be contacted and financial resources sought. It is suggested that the blind associations themselves should be involved in organising initiatives, though the following quote from a German institution that “unfortunately even the associations for the blind are absolutely inactive” is echoed elsewhere. Many responses to the survey were rather generic, such as “to spread good practices” and “to do our best to promote audio description”. Poland stresses the need to strive for legal regulations and makes the important point that “we have to work more with visually impaired children to convince them that going to the movies can be fun, to raise awareness and educate future audiences”.

8. Conclusion

Providing a comprehensive picture of the AD end-user is challenging. It is challenging because the elements to be taken into account are diverse and they not only pertain to the heterogeneity of this audience. Dealing with end-users means dealing with the language that suits them best; knowing and wielding the particular language ingredients that should guide audio describers, but at the same time, should prevent them from patronising their target audience when scripting their texts. This means taking care of their wishes, which eventually will determine the extent of enjoyment and understanding end-users will experience when accessing described
content. It means considering new ways of approaching the language of AD, including the
possibility of using simplified language varieties (Plain and Easy Language) to make the final
product even more inclusive to diverse audiences.

AD is not a new discipline any more. However, its nature renders it a versatile form of
audiovisual translation that can be exploited horizontally, with primary and secondary aims.
This chapter has focused on blind users and unveiled stances and results that might be of help
to future professionals and scholars. Much more can be done in this extensive area, exploit-
ing a wide range of applications and reaching out to a potentially huge range of audiences.
Until only recently, actual awareness of what audio description could offer was limited and to
a certain extent and in some places still is, though improvements in communication and tech-
nological advances, making AD more and more available on all modern appliances, will help
establish it in the eyes of the public at large. The main end-users remain the blind and partially
sighted and this chapter has concentrated principally on this sector, but its usefulness to other
users will also become more widely known (Starr in this volume). For the audio describer and
the AD “industry”, knowledge of the end-users’ requirements is crucial to the acceptance and
expansion of audio description use and it is this aspect that the current chapter has attempted
to address.

Notes
1 Both authors contributed to the full chapter, with Christopher Taylor dealing specifically with the
Introduction, The importance of AD, The end-user perspective, Non-discriminatory language, Conclu-
sion, and Elisa Perego dealing specifically with Basic information: what is blindness? AD-related lan-
guage, Plain and/or Easy Language in AD, The AD profession seen by end-users, Future expectations
of the end users.
2 Usability is the extent to which something is user-oriented, cognitively effective and satisfactory. Text
usability is linked to text complexity, which determines its readability and lack of ambiguity (Perego,
2020: 19).
3 EASIT (Easy Access for Social Inclusion Training), 2018–21, financed by the European Union under
the Erasmus+ Programme, Key Action 2 Cooperation for Innovation and the Exchange of Good Prac-
tices, Strategic Partnership. Project number. 2018–1-ES01-KA203–050275.

9. Further reading
Oxford: Oxford University Press.

10. References
WP1%20Report%20DEF
www.adlabproject.eu/home
https://videosdigitals.uab.cat/almacen/downloads/461/11104.mp4Arma
it/8740/


Elisa Perego and Christopher Taylor


11. Filmography