Audio description with audio subtitling - an emergent modality of audiovisual localisation

Sabine Braun[a]; Pilar Orero[b]
[a] University of Surrey, Centre for Translation Studies, Guildford, Guildford, UK [b] UAB, Translation, Bellaterra, Spain

PLEASE SCROLL DOWN FOR ARTICLE

Routledge
Taylor & Francis Group

Audio description with audio subtitling – an emergent modality of audiovisual localisation

Sabine Braun[a]* and Pilar Orero[b]

[a]*University of Surrey, Centre for Translation Studies, Guildford, Guildford GU2 7XH, UK;*
[b]*UAB, Translation, Bellaterra, Spain*

Audio description (AD) has established itself as a media access service for blind and partially sighted people across a range of countries, for different media and types of audiovisual performance (e.g. film, TV, theatre, opera). In countries such as the UK and Spain, legislation has been implemented for the provision of AD on TV, and the European Parliament has requested that AD for digital TV be monitored in projects such as DTV4ALL (www.psp-dtv4all.org) in order to be able to develop adequate European accessibility policies. One of the drawbacks is that in their current form, AD services largely leave the visually impaired community excluded from access to foreign-language audiovisual products when they are subtitled rather than dubbed. To overcome this problem, audio subtitling (AST) has emerged as a solution. This article will characterise audio subtitling as a modality of audiovisual localisation which is positioned at the interface between subtitling, audio description and voice-over. It will argue that audio subtitles need to be delivered in combination with audio description and will analyse, systematise and exemplify the current practice of audio description with audio subtitling using commercially available DVDs.

Keywords: audio description; audio subtitling; subtitling; voice-over; audiovisual translation

1. Introduction

Audio subtitling (AST) is a reality in some countries such as the Netherlands. Mildred Theunisz (2002) describes the implementation of a project for the creation of spoken subtitles on Dutch TV, how the project was born and its primary objective: to make foreign TV programmes more accessible for those who are visually impaired, and also for the elderly and for people with language impairments such as aphasia or dyslexia, or cognitive impairment such as mental retardation or decreased concentration.

The method used in this project to bring spoken subtitles to the homes of these user groups is speech synthesis. The broadcaster uses a speech-synthesis computer which is fitted with speech-synthesis software and converts the text into speech. This output is then converted into a signal and broadcast without disturbing the programme. Users can receive the signal using a special decoder. In a pilot test, this solution was met with a high degree of acceptance, and on 14 December 2001 the AST service was officially opened by the Dutch Secretary of State and the president

*Corresponding author. Email: s.braun@surrey.ac.uk

of one of the public broadcasters of the Netherlands, the NOS (Nederlandse Omroep Stichting). In the meantime, AST has been implemented as a permanent service on Dutch TV (van der Heijden, 2007). However, further development had to take place to correct and improve the quality of sound, since the synthetic voice was found monotonous.

The acceptability of synthetic speech was also one of the major concerns of a study conducted by SVT (Sveriges Television) in 2003 to test the usability of a speech synthesis system for AST. The study involved 16 visually impaired people aged 22 to 85. The main objective was to elicit the informants' reactions to synthetic speech. A secondary objective was to find out whether the spoken subtitles should be mixed with the programme sound at the source (broadcaster mix) or if the users prefer to control the mixing at home. The results of the test (de Jong, 2006, p. 133) show a unanimous response with regard to mixing the programme sound and the AST: as could have been expected, the informants wanted to control the mixing. In terms of usability, the results indicate that informants adapted fairly quickly to synthetic speech but they also emphasised the crucial importance of a uniform voice quality and a high degree of intelligibility of the subtitles. What must furthermore be borne in mind is that such studies tend to be influenced by the 'novelty factor' (i.e. any ASTs are better than no access to a film), which may partially account for the positive reactions.

In contrast to the outcomes of the test in the Netherlands and Sweden, which were conducted with TV programmes in mind, the UK *Guidelines on the provision of television access services* (Ofcom, 2006) largely discard a speech-synthesis solution for full-length drama or feature film:

> For a visually impaired audience with little knowledge of [the source language],these programmes will remain inaccessible. Equipment is available which can read teletext-delivered subtitles aloud, but the expressionless quality of a synthesised voice is not suitable for an entire drama or a film, and it is not feasible to recognise a variety of different speakers within the programme.

This points to some of the problems arising with AST. What is more, one aspect that was not considered at all in the two studies reported above was the inclusion of audio description (AD) along with the ASTs. In other words, it seems to have been taken for granted that the (automatically generated) ASTs would be sufficient for the visually impaired audience to follow the programme.

At the other side of the of the spectrum of AST lies an experiment which took place at the Barcelona Opera House – the Gran Teatre del Liceu – in 2005 in order to test the adequacy and comprehension of ASTs in an opera context. The opera chosen for this experiment was the concert opera of Donizetti's *Roberto Devereux* (Orero, 2007b). The departing point were the features of a concert opera: lack of dramatic action, minimal stage production and costumes, leading singers and choir all dressed in black and in a still position throughout the representation. In this situation, little AD was needed, allowing the experiment to focus on the realisation of the audio subtitles – or 'audio surtitles' to be precise. At the Liceu Opera House, surtitles are always on offer, so it was possible to set up a test case delivering spoken surtitles and monitoring their reception by a group of visually impaired people who had been attending regular opera performances (with AD only) throughout a season and with whom the research team had already established contact to assess the

quality of the AD. The news that *Roberto Devereux* was to be audio surtitled was met with much reticence and scepticism by the visually impaired audience. The audio describer who usually described the operas and who would be in charge of delivering the live ASTs raised numerous objections to the experiment.[1] However, in spite of all the initial rejection the AST experiment was a success. It was rated highly. The visually impaired recipients not only accepted the new format, they enjoyed it better, and since then audio descriptions at Liceu have been combined with spoken surtitles whenever possible.

The approach taken in this experiment is different in principle to the approach followed in the studies that were conducted in Sweden and the Netherlands (as reported above). The opera experiment combined ASTs with AD. The relatively reduced need for AD in operas may have facilitated the delivery of the ASTs, but it also means that the positive overall outcome of this experiment may not be transferable to other settings. Whilst some TV programmes (e.g. soaps and sitcoms) require equally little AD, other audiovisual products, especially feature films such as action-packed thrillers, drama or science fiction films, normally need a great amount of AD as they convey much information visually. This makes the combination of AST and AD in feature films more complex than in other settings.

Films with AD and AST such as *Volver* (Almodovar, 2006) and *The science of sleep* (Gondry, 2006) can easily sound 'cramped', making it difficult for the audience to distinguish voices and to process the information. Given the growing number of multilingual films and the trend towards subtitling in many countries, the combination of AD and AST is, however, becoming an important modality of 'audiovisual localisation'. The question of how to achieve a successful combination of AD with AST in subtitled feature films is therefore of utmost importance for film distributors and TV broadcasters alike, and it is one of the issues which are at stake within the DTV4ALL pilot project.[2] Given the risk that TV broadcasters currently run of focusing on the quantity of accessible programmes (in an attempt to comply with national and EU directives), the primary focus of the DTV4ALL project is on making access services more widely available without losing sight of quality.

In line with this, the aim of this article is to explore, systematise and assess the current practice of audio describing and audio subtitling feature films, to highlight the problems and to draw initial conclusions. Section 2 reviews the major challenges for AD with AST. Drawing on this review as a framework for analysis, section 3 then presents a case study examining current practice. Section 4 discusses current solutions for AD with AST in terms of quality and (financial) viability. Section 5 concludes the article.

2. Audio description with audio subtitling: the challenges

At face value, the most prominent and noticeable problem in audio subtitling may well be the spoken delivery of the subtitles. The way in which this is done is likely to have consequences for character identification, naturalness and comprehensibility, and eventually the enjoyment of the audiovisual product as a whole. On closer inspection, the difficulties of the combined method of AD with AST arise from a range of factors. These include the differences between visual and verbal meaning-making and between written and spoken language. In fact, AD with AST 'inherits' many of the problems that are familiar from each of the three modalities of audiovisual translation involved in it: subtitling, AD and also voice-over. A

comprehensive review of each of these modalities would be beyond the scope of this article but highlighting some of the major problems of each modality seems to be an appropriate starting point for the discussion of problems and solutions relating to AD with AST.

Subtitling is an often deceptively difficult form of interlingual translation in which traditional translation problems (such as the lack of simple interlingual equivalences or the language-specific and culture-bound nature of, for example, humour, wordplay, metaphors and allusions, all of which occur frequently in film dialogue) are compounded by subtitling-specific problems such as timing constraints (due to reading speed, shot and scene changes), space constraints (two-line rule and semantic segmentation) and the problems arising from differences between spoken language (in film dialogue) and written language (in subtitles) (Díaz Cintas & Remael, 2007; Gottlieb, 1998; Ivarsson & Carroll, 1998). One particular challenge for AST is that subtitles often greatly reduce the source-text message, relying on the recipients' ability to use visual input to compensate for condensations and omissions in the subtitles. This creates problems for an audience who have no access to the visual mode. Another point, raised, for example, by Assis Rosa (2001), is that subtitles often sound stilted or bookish because important elements of spoken language (e.g. discourse markers, hesitations, false starts or repetitions) are omitted in favour of condensation or are notoriously difficult to represent in writing (e.g. intonation and other features of paralanguage). While the general wisdom for subtitling is that the transfer from speech to writing is a matter of compromise (see e.g. Díaz Cintas & Remael, 2007, p. 61), audio subtitling offers the chance to restore some features of spoken language. This can make ASTs more lively but it is unlikely to solve the problem that visual information is often required to understand (audio)subtitles. The visual information has to be delivered in another way, which is, of course, one of the main reasons why ASTs in films should be accompanied by AD.

Audio description is, however, in itself a complex process of cross-modal mediation with a plethora of challenges (Braun, 2007, 2008). One of these is that the visual mode is normally rich in information and full of detail, while the verbal descriptions of it need to be succinct as they have to fit in silent moments of the film to avoid overlap with the film dialogue or essential sounds, and saturation of information by the user. Succinctness is therefore as important in AD as in subtitling, and careful selection of information is of the essence. To achieve this, audio describers normally rely on the film dialogue (and other input, including sounds and music). In other words, they draw on the ability of the visually impaired audience to combine the AD narrative with the dialogue (and other sound input) in order to identify characters, infer the message and create a coherent whole of the film in their minds (Braun, 2011). Against this backdrop, the challenge for AST becomes very clear: while non-subtitled films allow the AD to benefit to a great extent from what happens in the film dialogue, subtitled films provide only a reduced version of the dialogue and thus 'deprive' the AD of some of its anchorage. What is more, foreign (subtitled) films may use visual images and even sounds whose meaning is less familiar to the target audience, requiring a more elaborate AD than non-subtitled films; and the film dialogue may contain culture-specific references which are more difficult to grasp for the target audience. All of this necessitates adaptations either in the ASTs (compared with the written subtitles) or in the AD (compared with AD in non-subtitled films), or in both.

How this can be achieved and the precise relationship between AD narrative and ASTs has yet to be defined. Moreover, the relative weight of AD and AST is likely to vary across different films, depending on the film genre, the role of the dialogue in any particular film, the number of characters involved, on the pace of the dialogue and other factors. What is clear, however, is that ASTs and AD are the two key elements in any audio subtitled product, i.e. the main carriers of the message. An equally important question is therefore how these two elements should be delivered.

From a technical point of view, van der Heijden (2007) has highlighted the crucial role of sound mixing and volume adjustment to ensure that the different elements of an audio-described film can be clearly distinguished by the audience. But audio-subtitled products also raise more fundamental questions about the way in which the subtitles should be voiced, including the number of voice talents to be used, their age and gender in relation to the film characters, and the question of whether the original dialogue, albeit incomprehensible, should be audible. As was pointed out earlier, the answers to these questions have important consequences for the comprehensibility and naturalness of an audio-subtitled product. In particular, the chosen method of delivery has an impact on the audience's ability to identify the film's characters, to know who is who, which constitutes a central aspect of, if not the starting point for, comprehensibility. Many of the delivery-related questions are familiar from voice-over, which can therefore provide important clues for AST. The specifics of voice-over (Matamala, 2008; Orero, 2006) that are most relevant for the voicing of subtitles in AST include the different delivery strategies such as synchronisation with the original voice vs. delivery with a lag, and the enunciation of the subtitles, especially the question of whether they should simply be read or whether they should be acted out.

This brief overview may suffice to show that audio subtitling is positioned at the interface between subtitling, audio description and voice-over. The particular features and challenges of AST resulting from this position will serve as a framework for analysing current practice in the next section.

3. Systematising current practice

The case study presented here aims to give a systematic account of current practice in AD with AST, identify emerging patterns in the multitude of solutions and highlight challenges and questions that require further research, but also to make initial recommendations as they can be deduced from the current best practices. As was pointed out in section 1, the combined method of AD with AST has not yet received much attention in research, nor has it been covered comprehensively in current AD guidelines. A qualitative case-study approach therefore seemed to be an appropriate initial, explorative step.

The corpus of data on which this case study relies is the small corpus of (partially or fully) audio-subtitled and audio-described feature films available on DVD for an English-speaking audience at the time of analysis: *Be with me* (Khoo, 2005), *Borat* (Charles, 2006), *Brick Lane* (Gavron, 2007), *Bride and prejudice* (Chadha, 2004), *East is east* (O'Donnell, 1999), *Everything is illuminated* (Schreiber, 2005)*, Hero* (Yimou, 2004)*, Kung fu hustle* (Chow, 2004)*, Letters from Iwo Jima* (Eastwood, 2006), *Munich* (Spielberg, 2005), *The passion of the Christ* (Gibson, 2004)*, The science of sleep* (Gondry, 2006), *Syriana* (Gaham, 2005) and *Volver* (Almodóvar, 2006).

Drawing on the points highlighted in section 2 to derive appropriate categories of analysis, the first aspect we explored was the overall nature of the films in terms of genre, number of speakers, pace of dialogue and other basic parameters. This is summarised in section 3.1. The next aspect under investigation was the assignment of voices for the AD narrative and the ASTs, and associated implications for character identification (section 3.2). We then analysed the ways in which the ASTs are delivered. We distinguished between basic techniques such as the different voice-over strategies (section 3.3) and creativity in the delivery including, for instance, how the subtitles were enunciated and whether they were delivered verbatim or with modifications (section 3.4). A final point of analysis concerned potential adaptations in the AD narrative to accommodate the ASTs as opposed to 'traditional' AD without AST (section 3.5).

3.1. The nature of currently audio-subtitled films

The first point to be noted is that there is great variety in the small corpus in terms of filmic genres. The genres that can be found in the corpus range from political epics such as *Syriana* to comedies such as *Kung fu hustle*, historical dramas such as *The passion of the Christ* or *Letters from Iwo Jima* or *Hero* and modern dramas such as *Volver*. The current inventory of audio-subtitled films covers the whole gamut of possible cases, as illustrated in Table 1.

A further element of heterogeneity in the small film corpus is that it includes films with only a very small number of speakers (*Hero*) as well as films with a rather large number of speakers (*Kung fu hustle*), films with slow-paced dialogue (*The passion of the Christ*) as well as films which include very fast-paced dialogue scenes (*Volver*);

Table 1. Categories of audio-subtitled films.

Category	Film	Source language(s)
1. Individual subtitles for Occasional foreign-language utterances	*Munic*	English; some Italian
	Syriana	English; some Farsi
2. Multilingual films with partial subtitles	*Borat*	English with Hebrew and Armenian
	Bride and prejudice	English with Hindi and Punjabi
	Brick Lane	English with Bengali
	East is east	English with Urdu
	Everything is illuminated	English with Russian
	The science of sleep	English with Spanish and French
3. Monolingual, fully subtitled Films	*Hero*	Mandarin
	Letters from Iwo Jima	Japanese
	Kung fu hustle	Cantonese
	Volver	Spanish
4. Multilingual, fully subtitled films	*The passion of the Christ*	Aramaic, Latin and Hebrew*
	Be with me	Cantonese, English, Hokkien and Mandarin

*This film is probably a case in point. The dialogues are cast in Aramaic, Latin and Hebrew, but the trilingual nature of the film is lost in the (written) English subtitles.

and films in which a narrator is present on top of the dialogue (*Letters from Iwo Jima*). Other films such as *Hero* and *Kung fu hustle* also include a wide range of intertitles, captions and other written information on the screen. *Be with me* offers perhaps the widest variety of spoken languages (Cantonese, English, Hokkien and Mandarin) along with different written languages (mobile phone screens, internet chats, computer screen, typewriter) and with different calligraphies, and tactile languages such as Braille and tactile sign-language.

The wide variety among currently audio-subtitled films suggests that the current decisions in the industry regarding the films which are audio-subtitled are not related to the genre of a film or other basic characteristics such as pace of dialogue. Or in other words, it is not possible to infer any trends in the industry's decision about the films that are audio-subtitled. The variation in the corpus also means that it is too early to identify global, genre-specific AST practices. What can, however, be observed is a range of practices or decisions at micro-level, i.e. with regard to the realisation and delivery of the individual ASTs and their local integration with the accompanying audio description.

3.2. Assigning voices for AST and AD and implications for character identification

As was pointed out earlier, the perhaps most noticeable feature of ASTs is the way in which they are voiced. An important decision therefore relates to the number and choice of 'voices' for both the AD and the ASTs. In practice, decisions at this level will often be confined by the two rules which govern much of an audiovisual production: time and money. Most prominently, this will include, for example, the budget available for the production of an audio-described/audio-subtitled version of a film and the time constraints related to the release date of a film. However, as was pointed out in section 2, the choice of voices has major implications for character identification, naturalness of delivery and comprehensibility, i.e. for the quality of the ASTs. Therefore, this section will focus on advantages and drawbacks of individual solutions from the point of view of their quality. The important question of striking a balance between quality and financial viability of the solutions will be addressed in section 4.

One of the relevant aspects of voice assignment is whether the ASTs should be presented by the same person who delivers the AD narrative. In films such as *Munich*, which only contain very few subtitles to relay occasional foreign-language utterances, the subtitles are typically read out by the audio describer. This solution seems unproblematic as long as each character in a film only speaks in one language, and it seems to work particularly well when the original voice (i.e. the foreign-language utterance) can be heard underneath the spoken subtitle. A very specific problem arises, however, when the characters who speak in a foreign language also speak English in the same film. In *Syriana* – the other film in our category of films with only few foreign-language utterances – George Clooney mostly speaks English but occasionally Farsi. The effect for the visually impaired audience is that George Clooney's character has two different voices: Clooney's own voice in English and the audio subtitler's voice for the (subtitled) Farsi utterances. This effect is even stronger in multilingual films in which the amount of foreign-language utterances is higher. In *The science of sleep* and *Everything is illuminated*, some of the foreign-language utterances are made by the films' main protagonists, who speak English throughout the film as well. Hence, the audio-subtitled versions of these films come with a

stronger version of the 'George Clooney' effect. This can be difficult to process for an audience who do not have the visual input. It is particularly problematic given that voice recognition and character identification go hand in hand. The effect itself seems unavoidable (unless the actors lend their voices to the audio-subtitled version in such cases), but there is room for discussion of the way in which the ASTs could best be delivered in these cases so as to minimise the difficulties. We will return to this point in section 3.3.

For the remainder of the present section, we will return to the question of whether the ASTs should be delivered by the same person who delivers the AD narrative. This question becomes even more relevant in films which contain more than individual subtitles (i.e. all films other than *Munich* and *Syriana*). Further questions arise in such films with regard to the number and gender match of the voices. Our analysis reveals that there is a wide variety of solutions.

One way of assigning voices can be observed in the audio-subtitled versions of *Volver* and *The passion of the Christ*, which use a gender distinction between AD and AST. Both films use one male voice talent to present the AD and one female voice talent for all ASTs. In the case of *Volver*, this solution has turned out to be problematic.[3] The major problem is that some of the scenes include as many as five women, all of whom are involved in fast-paced interaction and overlapping speech. This is compounded by a complex plot which often goes against expectations. Hence the use of one female voice to read out all the subtitles makes character identification an issue and creates problems with following the plot. In *The passion of the Christ* the same solution seems to work better, and this is all the more noteworthy given that a female voice is used for all ASTs, while most characters are male. It appears that the solution works here because the film dialogue is slow-paced and the underlying (biblical) story is well-known, making information-processing easier. However, the film *Volver* illustrates the general limitations of this solution.

A more complex but perhaps more appropriate solution is to use not only a gender distinction between AD and AST voice, but to assign different voices to the ASTs, matching the sex and possibly age of the film characters. This solution has been adopted, for example, in *The science of sleep* and in *Kung fu hustle*. In the former, the audio describer is female while another female voice talent reads the subtitles when the actress Charlotte Gainsburg speaks in French, and a male voice delivers the subtitles when the actor Gael García Bernal speaks in French or Spanish. In this film special attention has been paid to reading subtitles with voices that match the sex of the actors (though not the age). A larger number of talents than in other films has been used to avoid confusion. While more than one voice is delivered by the same voice talent, this is done using different voice qualities (pitch, rhythm etc.) to distinguish them. The same is true for *Kung fu hustle*. It uses one male voice talent for the AD narrative, and the ASTs are read out by male and female voices according to the sex of the characters. This solution makes the audio-subtitled product very lively and natural. Unless a different voice can be assigned to each character (as was the case in the audio-subtitled version of *The lives of others* in UK cinemas), some creativity and acting talent is, however, required to distinguish the voices of different characters and facilitate character recognition (see also section 3.4).

In the film *Hero*, a combination of the two above methods is used. A male voice reads the subtitles of all male characters, while a female voice delivers the AD and the subtitles for all female characters. This method is useful for character identification but it requires ways of distinguishing the AD narrative from utterances

of the female characters. This raises the question of how the AD can be adapted to the AST situation, which will be discussed in section 3.4. However, we will first look at the basic techniques of delivering ASTs.

3.3. Delivery of the audio subtitles: basic techniques

In principle, ASTs can be delivered in two ways. A 'dubbing' effect would be achieved when the ASTs and the original dialogue are delivered synchronously without the original voices being heard. A 'voice-over' effect is achieved when the original dialogue can be heard while the audio subtitle is superimposed. In our film corpus, the 'voice-over' technique is used throughout. There is, however, once again some variation in the individual solutions.

Above all, the delivery varies with regard to timing. Most films use a slightly asynchronous delivery technique in which the original voices can be heard first, while the delivery of the ASTs is delayed by a few words or seconds. The one film that stands out is *Kung fu hustle*. In this film, some of the ASTs are presented with no or minimal delay. Since the ASTs are shorter than the original dialogue in most cases, the effect is that a considerable part of the original utterances can still be heard after the audio subtitle has been delivered. As long as this is done consistently, it seems a reasonable way of facilitating character identification. In other instances in the film, the ASTs are, however, delivered with a rather long delay, and in some cases they only begin after the original utterance has been completed. This leads to inconsistency. Moreover, in scenes with rapid speaker change, the long delay has the effect that too many voices can be heard (original and audio subtitler), making it difficult to know who says what.

The voice-over method seems useful for character identification in particular in multilingual films where AST entails that some characters are assigned two different voices (George Clooney speaking English and Farsi, as described in section 3.2). In other words, it may be useful for the visually impaired audience to hear the original voice of the film character. The usefulness of this may, however, be limited when the original language is unknown to the recipient (as Farsi would be for the majority of English-speaking recipients). At least, it would have to be investigated whether the recognition of a character's voice is successful when the language in which the character speaks is unknown to the recipient.

Another key issue that arises in connection with the delivery of ASTs falls in the field of sound engineering and has to do with the quality of recording, volume adjustment and distinction between the different parts of the soundtrack. In the film *Everything is illuminated*, for example, the audio subtitle volume is too loud, producing a rather dominating or imposing effect. In other films, including *Volver*, the original dialogue is too loud, which makes it even more difficult to focus on the ASTs in this film. With regard to AD without AST, van der Heijden (2007, p. 18) emphasises that: '[i]t is important that a description between the dialogues is clearly audible and not drowned out by other sounds. The balance between the original sound and the added AD should therefore be adjusted to each other'. By analogy, the same applies to the distinction between AD and ASTs (in lieu of original film dialogue) and – as the example of *Volver* shows – to the distinction between original dialogue and ASTs. Van der Heijden (2007, pp. 24–25) describes the many layers of sound present in a film and discusses different technical options for adjusting the

volume of the voice-over to that of dialogue in such a way that the 'voice-over sound [becomes] more involved in what [is] happening on screen'.

The careful choice of technique and adjustment of volume levels are likely to be crucial for ensuring the success of an audio-subtitled film. Apart from these more technical prerequisites, we have, however, also identified other ways of making the ASTs sound 'more involved'. The next section will discuss the more creative aspects of AST delivery.

3.4. Delivery of audio subtitles and creativity

Creativity can be observed in current AST practice at different levels. As was pointed out in section 2, AST offers the opportunity to restore some of the spoken features of dialogue that are normally lost in written subtitles. Some of the films that were analysed demonstrate that audio subtitlers made very creative use of the options at hand. Apart from that, our analysis also shows that creativity is used to make up for the loss of visual information and to save time in the audio descriptions.

In terms of adaptation to speech, it is first of all interesting to note that some of the films resort to non-standard English accents in the ASTs such as Cockney in *Kung fu hustle* and *Everything is illuminated* or Mancunian in *East is east*. In *Kung fu hustle* this works well because of the use of different voices and because it goes hand in hand with the related effort that is made in this film to act out the subtitles rather than reading them. In *Everything is illuminated*, by contrast, the use of a non-standard accent is more problematic, mainly because one voice is used for all ASTs (and reinforced by the problems with the volume described in the previous section). The use of one voice for all ASTs (see also section 3.2) creates a more narrative kind of voice-over – also known as Gavrilov translation[4] – which, in turn, seems to come with certain expectations regarding a more formal use of language.

A case in point is *The science of sleep*, which is not only multilingual, as mentioned earlier, but which also uses accents in each language. The actor Gael García Bernal speaks Spanish with a marked Mexican accent, French with a Hispanic accent and English, also with a Hispanic accent. His accents in French and Spanish are lost in the written English subtitles, and these are delivered in standard English by the audio subtitler. As a result, Bernal not only speaks 'with two voices' but also with two different English accents – a Hispanic accent in his own utterances, and a standard English accent in the ASTs. The question arising is whether anything could or should have been done to create more coherence between Bernal's own English utterances and his audio-subtitled French and Spanish utterances, and, if so, what the solution could be. Only reception studies involving visually impaired users will eventually be able to shed light on this problem.

Two other interesting adaptations of ASTs to the spoken modality that can be observed in the corpus are deviations of the ASTs from the written subtitles to introduce spoken turns of phrase and instances in which the ASTs are acted out rather than read. In general, most ASTs in the corpus are acted out to a certain extent, using appropriate intonation patterns, while differences between written and audio subtitles with regard to wording are minor. The film *Kung fu hustle*, however, is a good example of changes in the wording of individual subtitles to adapt them to speech. When one of the residents of Pig Sty Alley, the Shanghai neighbourhood in which much of the plot is situated, is washing his hair under an outdoor standpipe and suddenly finds the water being turned off on him, he calls for his landlady, who

lives on-site, and the argument shown in Table 2 begins to unfold between the two characters. Soon, other residents begin to gather around the pair and, just like the hairwasher, they become the target of the landlady's verbal attacks.

The changes in the wording of subtitles 5, 8 and 10 – however minor they might seem – help to make the audio subtitled version sound natural in the spoken mode, especially the addition of the greeting in 10, which has an interactive function here. At the same time, the rewordings in 5 and 8 highlight the haughty and aggressive behaviour of the domineering landlady, which is symptomatic of the tensions in the neighbourhood and an important part of the plot.

The next aspect of creativity we analysed was the compensation for any lack of visual information. Here again, the film *Kung fu hustle* is an interesting example of creativity. In one of the earlier scenes, when the film is still introducing the different characters, the landlord is seen walking through the neighbourhood, eying a young girl whom he had not seen for a while. As he is walking past her, the English subtitle has him say, '*Jill, how you've grown! Come let me examine you!*', while the audio-subtitled version turns this into '*Jill, how you've grown! Let me give you a physical exam!*', providing a good 'translation' of the landlord's lustful gaze at the girl. The application of such compensation strategies is, however, by no means systematic. To give a counter-example, the dialogue presented in Table 2 continues with the landlady's yelling at another onlooker, a coolie who carries a heavy lot of bags. The subtitle of the landlady's utterance reads: '*Serves you right, you are a coolie*'. There is no time for any AD to be inserted at this point, and the subtitle is read out verbatim, with the result that the visually impaired audience does not get to know what it is that 'serves him right'. Given the limited time available for voicing a subtitle, there are, of course, clear boundaries for creativity. What can be observed, however, is that the audio describers/audio subtitlers of the films in our corpus found other ways of including essential information, especially by adapting the AD itself. This is the final aspect of our analysis, to which we will turn in the following section.

Table 2. Example of dialogue from *Kung fu hustle*.

		Subtitles	Audio subtitles
1	Resident:	What happened to the water?	What happened to the water?
2	Landlady:	Water ain't free.	Water ain't free.
3	Landlady:	You talk a lot for someone who won't pay his rent.	You talk a lot for someone who won't pay his rent.
4	Resident:	But I'm in the middle of a shampoo.	But I'm in the middle of a shampoo.
5	Landlady:	You think that's a problem?	**So what?**
6	Landlady:	From now on, no water on Mondays, Wednesdays and Fridays.	From now on, no water on Mondays, Wednesdays and Fridays.
7	Landlady:	Water ration Tuesdays, Thursdays and Saturdays.	Water ration Tuesdays, Thursdays and Saturdays.
8	Landlady:	I just love hearing you fools complain again. Assholes!	**You fools dare** complain again! Assholes!
9	Taylor:	Good morning, landlady.	Good morning, landlady.
10	Landlady:	Just pay your rent, or I'll burn your shop down.	**Good Jul Raz**. Just pay your rent, or I'll **burn down your shop**.

3.5. Adaptation of AD

There is an ongoing discussion about the question whether AD is more of a narration with overall 'responsibility' for carrying along the story, or whether it is a description with no more than a supportive function (Fix, 2005; Yeung, 2007; see also Braun, 2008). While the general approach to AD may depend, in part, on the filmic genre, the adaptations that we observed in the audio-subtitled films mostly have the effect of giving the AD more overall 'responsibility' for carrying the story. Thus, adaptations were made to integrate ASTs into the AD narrative, especially where they were read out by the person who read the AD. The AD also indicated more extensively than in non-subtitled films who is going to speak and/or who is being addressed.

The integration of ASTs into the AD takes many forms in the films analysed. Sometimes it is simply done by a change in intonation. For example, when the Pig Sty Alley neighbourhood is introduced in *Kung fu hustle*, the entrance to the quarter is shown, and a caption reading 'Pig Sty Alley' appears on top of the screen. The audio describer says, '*Only in the poorest districts which hold no interest in the gangs can people live in peace. Pig Sty Alley [appears as caption] is just such a district*', and he puts emphasis on 'Pig Sty Alley' by slowing down and raising his voice. The subtitle thus becomes part of the AD narrative, but the audio describer ensures that it is sufficiently emphasised and noted, creating an effect which is similar (or equivalent) to the visual image with the written caption on the screen. Other ways of introducing written material on the screen are to say 'a subtitle', 'a caption reads' etc. In the material that was analysed for the present study, these strategies are mostly used to render individual subtitles or captions, or to introduce the very first subtitle in a full subtitled film (e.g. in *The passion of the Christ*).

Another way of integrating subtitles into the AD narrative is to quote them as direct speech in the AD narrative. When, for example, the troops in *Hero* get ready to attack, their battle cry, which can be heard in Mandarin, is translated in the written subtitle as 'Storm!'. This is rendered in the AD as '"*Storm" cry the troops*', with an emphasis on 'storm' that serves to imitate or act out the cry.

This method is mostly used to indicate who is speaking, and it can be found in films which use one and the same voice talent for the AD and at least some of the ASTs. The overall effect in this case is that of voice-over but, in contrast to the Gavrilov method (see section 3.4), the audio describers basically add deictic references, either before or after the utterance, denoting who the speaker is. In *Hero*, it is not uncommon for the female audio describer, who also reads the subtitles for the female characters, to offer descriptions such as '*I would like to ask you a favour, she says*'. When there is not much time, at least the name of the speaker is mentioned (e.g. '*Snow: The imperial guards are not worth a mention*', Snow being one of the main characters in *Hero*). Apart from that, the indication of who is speaking also relies on other, more indirect strategies. For example, the description of a character's gaze direction can help the audience to infer who is going to speak next (e.g. in *Hero* '*Broken Sword raises his eyes to Nameless*' [Broken Sword speaks next]; in *The passion of the Christ*, '*He turns to face X*' [he speaks next]).

All of these strategies help with the identification of the speaker and are one of the many strategies of going beyond the straight reading of the subtitles. Their overall effect is that the AD becomes more of a narrative, adopting more responsibility for relaying not only the scenery and actions but also the interactions

between the characters, which in 'traditional' AD are often left for the recipients to infer.

It could be argued that the distinction between AD and AST vanishes in such cases and consequently that the two modes are even more integrated than many of the other examples have suggested, i.e. that AD and AST are two sides of the same coin. Further exploration of this aspect might be beyond the scope of the present article, but it would clearly be a fruitful topic for further research and discussion.

4. Discussion: between quality, awareness and feasibility

The analysis in the previous section has shown that even in the small corpus of audio-subtitled films currently available in English, the resources deployed for producing AD with ASTs vary greatly across different films. *Kung fu hustle* and *The science of sleep*, for example, use multiple voice talents to act out the subtitles of different characters and to present the AD narrative. *Hero*, on the other hand, uses only two voice talents for the AD and ASTs, but it does so in a creative way, especially by integrating the ASTs into the AD and adapting the AD narrative to the situation at hand. In spite of these differences in the resources used, each of these films can be said to provide a valid solution for making the subtitles audible and accessible. What requires further analysis then are the factors that have an impact on the quality of audio-subtitled products.

The choice and number of voice talents is likely to be the aspect of AST that is most directly conditioned by available resources (i.e. budgets). As was pointed out in section 2, this is, however, also the aspect that has the most direct and 'visible' impact on character identification, naturalness of delivery and comprehensibility of the ASTs, i.e. on the reception and acceptance of an audio-subtitled product as a whole. A balance is therefore needed between quality and financial viability. To comply with emerging national and European directives for the provision of AD especially on TV, quantity may at present be the major issue, and this may define the resources that are made available for the production of any single audio-described/audio-subtitled programme or TV drama.

At one level, our analysis shows that budget constraints do not have to result in low quality. Strategies such as creativity, adaptation of the AD narrative, integration of AD and AST and acting out of the subtitles can be applied regardless of available resources. It is equally obvious, however, that factors such as timing constraints for the insertion of AD and ASTs into a film generate some limitations to that creativity, especially to the amount of adaptation that is possible in the AD narrative. To be sure, as in most other modes of audiovisual translation, including AD, the aim of AST cannot be to recreate to the full the experience of a sighted viewer when watching a subtitled film, but the aim of media access is by default to create an experience for the audience that is similar or equivalent in its effect. When resources are limited, requirements are high (and growing) and creativity options are constrained, the producers of AST products need to resort to other strategies in order to create a viable product.

One of these strategies would appear to be good planning to ensure, for example, that the realisation of the ASTs (e.g. the approach to voice assignment as discussed in section 3.2, the method of delivery as discussed in section 3.3 and potential modifications to the written subtitles as discussed in section 3.4) is consistent and conceived in conjunction with the preparation of the AD script. One interesting

point in this connection is that at present, it is difficult to see what the decisions about the chosen approaches and methods are based on, given that the solution in *Kung fu hustle* (gender matching in ASTs) does not necessarily require a greater number of voice talents than the method of gender distinction between audio describer and audio subtitler while coming across as more natural. The decisions or strategies may be a matter of (growing) experience in AST. At the same time, the authors of this article are wondering for both *Kung fu hustle* and *Hero* whether it might not be a more useful strategy to use the dubbed English versions that are available on DVD for these two films to insert AD without AST.

A related point is this: while the films analysed in our case study provide many good solutions for AD with AST in principle, the weaknesses and inconsistencies show that current practice is still largely based on individual intuition. It seems that more explicit and systematic research-led guidance on AD with AST could be a useful long-term strategy for avoiding pitfalls such as inconsistencies in the approach. Combined with carefully designed reception studies, research that is capable of explaining the reasons why some AD/AST solutions are more efficient than others is likely to make a major contribution towards improving and refining the guidelines which are currently emerging throughout Europe (and in other parts of the world).

More elaborate explanations and justifications of solutions will also have a key role to play in the training of aspiring and future audio describers/audio subtitlers, a task which is becoming more important with the increasing amount of AD and AST required. Along with further research and the creation of systematic research-led guidance, the use of trained describers and voice talents would therefore appear to be another crucial strategy for establishing and maintaining quality while increasing the quantity of programmes with AD and AST.

5. Conclusion

This article has aimed to account for the current practice of audio description in combination with audio subtitling as an emergent modality of audiovisual translation and localisation to make foreign (subtitled) films accessible to visually impaired communities. Audio subtitling was characterised as being situated at the interface of subtitling, audio description and voice-over, and clearly some features of dubbing could be added. It was argued that typical features and problems of each of these audiovisual modalities – such as the condensation characteristic of subtitles – have implications for ASTs, requiring strategies and solutions that go beyond a simple 'reading out' of the subtitles on the screen. A variety of strategies was identified in a small corpus of audio-subtitled films available on the British DVD market, ranging from different methods of assigning voices to ASTs (and the accompanying AD narrative), different techniques of delivery (various types of voiceover), an integration of ASTs with the AD narrative and an adaptation of the AD narrative itself in order to explicate, for example, who is speaking and/or addressed. It was also emphasised that the quality of AD with AST is likely to benefit from such distinctive factors as appropriate resource allocations and creativity as well as good planning, the development of more detailed, research-led guidance, training and the use of trained describers and voice talents.

Acknowledgements

Notes

1. We can only speculate on the reasons for the initial scepticism. It may have arisen from what could broadly be subsumed under resistance to change, including a perception that an audio describer knows best what the audience needs are, and that research interferes with this knowledge.
2. http://www.psp-dtv4all.org/.
3. In a viewing experiment conducted by the RNIB, the majority of visually impaired people reported difficulties in understanding the audio-subtitled version of this film (Joan Greening, personal communication), and this experience is matched by the authors' own attempt to follow the sound track of the audio-subtitled version without watching the images.
4. Franco, Matamala and Orero (2011) have described this type of voice-over translation as follows: 'The main characteristic of a Gavrilov translation of a fiction film ... is that, although the original might contain various characters, males and females, the voice-over is generally performed by a single voice talent, usually male, who depicts a very fast paced discourse, usually overlapping the original dialogue – which can be partially heard – and expressing no emotions'.

References

Assis Rosa, A. (2001). Features of oral and written communication in subtitling. In Y. Gambier & H. Gottlieb (Eds.), *(Multi)media translation: Concepts, practices, and research* (pp. 213–221). Amsterdam: Benjamins.

Braun, S. (2007). Audio description from a discourse perspective: A socially relevant framework for research and training. *Linguistica Antverpiensia NS, 6*, 357–369.

Braun, S. (2008). Audio description research: State of the art and beyond. *Translation Studies in the New Millennium, 6*, 14–30.

Braun, S. (2011). Creating coherence in audio description. *Meta*, forthcoming.

De Jong, F. (2006). Access services for digital television. In R. Perez-Amat & A. Pérez-Ugena (Eds.), *Algunas consideraciones sobre la audiodescripción comercial. Sociedad, integración y televisión en España* (pp. 331–344). Madrid: Laberinto.

Díaz Cintas, J., & Remael, A. (2007). *Audiovisual translation: Subtitling*. Manchester: St Jerome.

Fix, U. (Ed.). (2005). *Hörfilm. Bildkompensation durch Sprache*. Erich Schmidt: Berlin.

Gottlieb, H. (1998). Subtitling. In M. Baker (Ed.), *Encyclopedia of translation studies* (pp. 244–248). London: Routledge.

Ivarsson, J., & Carroll, M. (1998). *Subtitling*. Simrishamn: TransEdit.

Matamala, A. (2008). Teaching voice-over. In J. Díaz Cintas (Ed.), *The didactics of audiovisual translation* (pp. 115–127). Amsterdam: John Benjamins.

Ofcom (2006). [Revised version] *Guidelines on the provision of television access services*. Retrieved from http://www.ofcom.org.uk/tv/ifi/guidance/tv_access_serv/guidelines/guidelines.pdf.

Orero, P. (2006). Synchronization in voice-over. In J.M. Bravo Gozalo (Ed.), *Aspects of translation* (pp. 255–264). Valladolid: Universidad de Valladolid.

Orero, P. (2007a). Sampling audiodescription. In J. Díaz Cintas, P. Orero, & A. Remael (Eds.), *Media for all* (pp. 111–125). Amsterdam: Rodopi.

Orero, P. (2007b). Audio subtitling: A possible solution for opera accessibility in Catalonia. In E. Franco & V. Santiago Araújo (Eds.), *TradTerm* (pp. 135–149).

Orero, P. (2011). The audio description of spoken, tactile, and written languages in *Be with me*. In A. Serban, J.-M. Lavour, & A. Matamala (Eds.), *Audiovisual translation in close-up: Practical and theoretical approaches* (forthcoming) Frankfurt: Peter Lang.

Theunisz, M. (2002). Audio subtitling: A new service in Netherlands making subtitling programmes accessible. Retrieved April 22, 2005, from http://www.sb-belang.nl.

Van der Heijden, M. (2007). *Making film and television accessible to the blind and visually impaired*. MA dissertation. Utrecht School of the Arts. Retrieved October 27, 2009, from http://www.audiodescriptie.com/downloads/Making%20film%20and%20television%20 accessible%20to%20the%20blind%20and%20visually%20impaired%20-%20M.%20van%20 der%20Heijden.pdf.

Yeung, J. (2007). Audiodescription in the Chinese world. In J. Díaz Cintas, P. Orero, & A. Remael (Eds.), *Media for all. Accessibility in audiovisual translation* (pp. 231–244). Amsterdam: Rodopi.

Filmography

Be with me (2005). Eric Khoo.
Borat (2006). Larry Charles.
Brick Lane (2007). Sarah Gavron.
Bride and prejudice (2004). Gurinder Chadha.
East is east (1999). Damien O'Donnell.
Everything is illuminated (2005). Liev Schreiber.
Hero (2004). Zhang Yimou.
Kung fu hustle (2004). Stephen Chow.
Letters from Iwo Jima (2006). Clint Eastwood.
Munich (2005). Steven Spielberg.
The passion of the Christ (2004). Mel Gibson.
The science of sleep (2006). Michel Gondry.
Syriana (2005). Stephen Gaghan.
Volver (2006). Pedro Almodóvar.