



impaginazione Verena Papagno

© copyright Edizioni Università di Trieste, Trieste 2012.

Proprietà letteraria riservata.

I diritti di traduzione, memorizzazione elettronica, di riproduzione e di adattamento totale e parziale di questa pubblicazione, con qualsiasi mezzo (compresi i microfilm, le fotocopie e altro) sono riservati per tutti i paesi.

ISBN 978-88-8303-347-6

EUT - Edizioni Università di Trieste  
Via E. Weiss, 21 - 34128 Trieste  
<http://eut.units.it>

Emerging topics  
in translation:  
Audio description  
edited by  
Elisa Perego



# Table of contents

- Elisa Perego*  
7 Introduction
- Pilar Orero*  
13 Film reading for writing audio descriptions: A word is worth a thousand images?
- Nathalie Mälzer-Semlinger*  
29 Narration or description: What should audio description “look” like?
- Saveria Arma*  
37 “Why can’t you wear black shoes like the other mothers?” Preliminary investigation on the Italian language of audio description
- Agnieszka Chmiel, Iwona Mazur*  
57 AD reception research: Some methodological considerations
- Agnieszka Szarkowska, Anna Jankowska*  
81 Text-to-speech audio description of voiced-over films. A case study of audio described *Volver* in Polish
- Bernd Benecke*  
99 Audio description and audio subtitling in a dubbing country: Case studies
- Riitta Lahtinen, Russ Palmer*  
105 Environmental description
- 115 Bionotes Authors



# Introduction

ELISA PEREGO  
University of Trieste, Italy

The term ‘translation’ is becoming more and more flexible to include processes that do not necessarily involve the traditional conversion from one language to another. The possibility of converting one semiotic channel into another enables us to incorporate in the field of audiovisual translation practices such as the conversion of spoken language into visuals (i.e. subtitling for the deaf and hard of hearing or signed language films) or of visuals into spoken language (i.e. audio description – AD – for the blind). Both processes are meant for sensorially impaired audiences and they are associated with the need to cater for their social inclusion, increasing independence, and the chance and right to enjoy media products. This volume, which deals specifically with several aspects of audio description for the blind and sight impaired, brings together varied contributions with more or less traditional approaches. These include the linguistic analysis of AD in various languages, the illustration of various strategies to overcome possible obstacles while audio describing, aspects of audio description which have only recently been dealt with (e.g. the need to depart from strict and limiting guidelines, the examination of the acceptability and the reception of text-to-speech audio description – TTS AD –) and some methodological considerations regarding reception research in AD. A chapter on environmental description offers a thought-provoking view which encompasses whatever lies beyond the audio description of films. The topics dealt with are up to date, especially in Italy,

and the publication of such a volume is propitious because it coincides with the launch of the European ADLAB (Audio Description: Lifelong Access to the Blind) project ([www.adlabproject.eu](http://www.adlabproject.eu)).

**Pilar Orero's** contribution opens the volume with a focus on the ability to interpret films, which is a too often neglected pre-requisite for effective audio description. Watching a film is not as simple an activity as one might think. Films have a grammar, quite a complex one, just as languages do. Drafting comprehensive and effective audio descriptions is made possible by a deep understanding of the language(s) and the grammar of film. After an overview on the ways of seeing and reading films based on Berger (2008) and Metz (1997), Orero focuses on the manifold interpretations films are subject to, even by the same viewer at different moments in time. Such lack of uniqueness in interpretation has to be taken into account by the describer. His/her job is to deal with film language and to translate it choosing the most adequate interpretation to convey. Since there is no correct or unique interpretation, the describer should be able to understand the language of films, recognize the different reading levels, identify the denotative and the connotative meanings and the rhetoric of the visuals, and finally grasp the intentions of the director, before drafting an audio description. Only departing from the guidelines typically calling for total objectivity (superficiality, according to Orero) will he/she be able to create a usable product. Only explicating the semantic load of the visuals will he/she be able to cater for the needs of the blind audience. This perspective is quite recent and it is shared by many scholars in the field, as recently emerged, for instance, during the Advanced Research Seminar on Audio Description held in Barcelona in March 2011. European guidelines seem quite rigorous and rigid. Following them strictly may lead to the production of objective but cold descriptions, lacking precision and real comprehensibility. This is also made clear in the volume by Nathalie Mälzer-Semlinger. In a film there is often much more going on than what is objectively being described. A man touching a woman's hand, as reported in the AD analyzed by Orero (p. 23), or the recurrent colour of a blouse, which is insistently mentioned in the AD analyzed by Mälzer-Semlinger (pp. 31-32), carry such a meaningful, irreplaceable and rich semantic load that they cannot be left out, even though explaining, explicating and interpreting are not (yet) words found in present AD guidelines. This is one of the reasons why in her contribution **Nathalie Mälzer-Semlinger** focuses explicitly on the need for more flexibility in (German) AD guidelines, which should be adapted by audio describers on the basis of the narrative strategies used in the original film. This is particularly important in feature films where the narrative is thick and intricate, allusive and ambivalent, rich in symbolic elements at the visual level. In such situations, the required objectivity and syntactic simplicity called for in the (German) guidelines might not be the proper means to deliver an effective description. When needed, the describer should feel free to offer the blind audience explicit mention of details and interpretation of the cinematographic

codes. Even though the describer might run the risk of over-emphasizing seemingly incidental details which occur in the narrative, at times they have to make the motivations of the narrative strategies explicit and available to the audience. Balancing objectivity and explanations is a complex procedure – it is difficult to name covert elements of the visual narrative, such as, for instance, the meaningful and evocative direction of a character's gaze. It is however necessary to grant the blind audience actual participation and full enjoyment of the film experience.

Leaving guidelines aside for a while, the volume moves on to a more linguistic perspective. **Vera Arma** inspires readers with a reflection on the language of audio description. Audio description is a written to be read text type, which as such tends to stand closer to the written end of the spoken-written continuum. This is particularly so as far as Italian is concerned. Although her reflections are based on a corpus-based analysis of one film only, her preliminary results suggest quite a strong tendency in Italian audio described texts to reproduce formal and written register features in the form of lexical choices (which are unusual in spoken language), word order (cf. the large number of adjectives preceding nouns) and syntax (cf. the frequent embedded structures). A comparison of the English and the Italian AD of the same film also shows that while English is more linear and straightforward, regular and simple, Italian bears a resemblance to literary language and it is more varied and complex. Differences between the two languages and the two ADs have been detected also in the selection of information to be described, with Italian being more subjective than English. A quick hint is made in the paper on the possible effect of the cognitive load of AD receivers in determining the selection of material to be audio described. Though this topic is not analyzed explicitly by Arma, it certainly is one which deserves more in-depth analysis: reception studies, which are nowadays developing, can in fact contribute to the improvement of media usability and accessibility. They are primarily concerned with discovering how viewers interact with films and their various forms of audiovisual translation, thus to be able to guide translators to the creation of ever more ergonomic products.

A paper totally devoted to reception studies is that of **Agnieszka Chmiel** and **Ivona Mazur**, who focus specifically on methodology in AD research on reception. As they claim, reception studies in AD can directly contribute to the quality of audio description because they enable researchers to identify viewers' preferences and overall feedback, which is the first step toward drafting quality and user-oriented ADs. However, conducting research in the realm of reception studies is not easy, especially if working with visually impaired people (VIP). The authors illustrate the most common difficulties in accessing and working with such a particular sample of the population, and those in building a questionnaire that takes into account VIP's real needs. In the paper methodological difficulties are described and solutions to overcome them are offered. A sample questionnaire is included as an appendix, which can function as an invaluable tool for research

in AD. A further significant step toward the production of ergonomic and quality products has been taken by **Agnieszka Szarkowska** and **Anna Jankowska**, who encourage researchers to consider the cognitive aspect of audio description and the usability issues which should always accompany the study of audiovisual translation in all its forms. In particular, the authors examine the case of audio description trying to find its place in voiced-over products and they consider the use of synthetic text-to-speech (TTS) systems as a cheaper alternative to traditionally produced audio description, and as a means to increase both the number and the range of audio described films on the market. They point out how the European audiovisual landscape is changing and offer a fresh look at the specific, complex and challenging case of Poland, with dubbing, subtitling and voice-over existing side by side. They demonstrate how TTS AD can be combined with voice-over to produce AD to foreign films and they present the results of a survey conducted among a group of blind and partially sighted subjects after watching a TTS audio described film excerpt. Contrary to what many may think, the results of the survey demonstrate that the participants are quite open to the idea of TTS AD both as an interim solution – until there are more audio described films available – and as a permanent solution. The paper shows that misconceptions about the overload of information resulting from the combination of various sources of information are starting to be adjusted. What appears to be a hindering factor (i.e. the combination of audio description with voiced-over/audio subtitled foreign programs) is not considered as an obstacle to film enjoyment by real end-users. Although still preliminary, these results might help to eradicate prejudices and reluctance to the production of an increasing number of accessibility services throughout Europe.

Besides resorting to reception studies, taking the suggestions of practitioners into account could prove to be a useful shortcut. **Bernd Benecke's** contribution, for instance, generates from his hands-on experience as a professional audio describer in Germany, which is a dubbing country. Dubbed material can be easier to audio describe if compared to subtitled material. However, dubbed films can occasionally include subtitles which need to be made available to the blind. Written subtitles are usually made available by an extra voice talent who reads them out aloud (that is, resorting to audio subtitling). The integration of audio description and audio subtitling however is not easy. Benecke shows examples of real practice and gives tips to get over possible obstacles posed by specific situations. He considers four specific cases where subtitles are needed in the product's dubbed version (i.e. when one or more characters use sign language; when a foreign language is used in documentary films; when songs in dubbed films are kept in the original language; when a blend of various foreign languages are used in the same film) and demonstrates where, why and how describers can deal with such complex but common situations.

Although the volume focuses on audio description as a tool for making films available to the blind audience, it should be remembered that (audio)

description serves a wider range of situations and people: unfortunately, sensory impairment can be quite pervasive and cover more than one disability. Subjects may have complex and multiple special needs (blind/visually impaired, Deaf/hard of hearing, Deafblind) which are difficult to cater for. In their closing contribution, **Riitta Lahtinen** and **Russ Palmer** go beyond blindness and film or TV audio description and give an account of the more general process of environmental description meant for multiple sensory impaired people. Mainly focussing on deafblindness and dual sensory impairment, the authors present various methods and techniques to describe the environment which depart from language (spoken, written or sign language) and include the wider use of the senses. In particular, they make the reader aware of haptic perception, i.e. the process of recognizing objects through touch, and of one of its branches, i.e. drawing with the index finger on the back of the receiver to convey direction-related information. Though not dealing with the core topic of the volume the overview closes it by widening the horizons on the infinite ways reality can be made accessible to an impaired audience.

#### ACKNOWLEDGEMENTS

Special thanks go to the friends and colleagues who participated in the audio description session in the *Emerging topics in translation and interpreting* conference, held in the Department of Language, Translation and Interpreting Studies of the University of Trieste (16-18 June 2010), to Vera Marchesi for her initial editing work, and to Christopher Taylor for supervising the work all the way through the conference and the publication process.



# Film reading for writing audio descriptions: A word is worth a thousand images?

PILAR ORERO\*

Universitat Autònoma de Barcelona, Spain

## ABSTRACT

*While English, German, Catalan, Music, Mathematics are languages which have a vocabulary, a grammar and a syntax – which needs to be learnt before being able to read – film language is understood by babies before they can speak or read. Films can be enjoyed naturally without acquiring any fluency in its language, and this natural approach seems to be taken by many when drafting audio descriptions. Though much international attention has been paid recently to draft audio descriptions standards and guidelines (Benecke 2004; Ofcom 2006; Orero and Wharton 2007; Puigdomènech et al. 2007; Remael 2005; Snyder 2006; AENOR 2005; Vercauteren 2007), little attention has been devoted to the most basic elements of film: its vocabulary, how to read it and its meaning (an exception could be made with sound since we already have articles by Remael forthcoming and Igareda forthcoming). This article departs from basic concepts such as the artistic experience, its channels of reception and how films are presented and perceived. Film languages are then discussed to focus on the image and the many possibilities of its reading. It is through the integration of all the readings and meanings that a deep understanding of the film is achieved; hence a comprehensive audio description can be drafted. It is interesting to note the differences between reading – which is the focus of this article – and telling a story visually. This latter issue is key when drafting audio descriptions for films, since narration will play the leading role, but it is not the focus of this article.*

## INTRODUCTION

Many basic questions are needed to understand the reading complexity deployed in a film through its multiple channels. How to see? How to read? How to listen? How to understand? Ultimately, how to feel? Ironically, Metz (1974: 47) coined the description of a film as being “difficult to explain because it is easy to understand”. The simplicity of understanding a film is based on the accessibility to film narrative and visual imagery which does not require any intellectual training nor understanding. Most people can enjoy films – at least at the most basic level. This is the case for users who do not hold any further function than that of being the consumer at the end of the film production chain. Different functions and responsibilities are held by different constituencies: the film critic, the student or the person – or team – who drafts the audio description (AD), the viewer. John Berger (2008: 25) brings the topic of art ownership to the fore in *Ways of seeing* where he questions where art belongs “to those who can apply it to their own lives, or to a cultural hierarchy of relic specialists”. Veering the argument towards real life and consumers, if we agree that films are produced to be seen by viewers, Bordwell (1985: 30) comments that humans pose many restrictions when seeing a movie, and abandons the concept of the ‘ideal viewer’. Though there seems to be no way forward to define, or group, those who consume films, since the enjoyment of a film depends on each individual at a different time, it is important to analyse the relationship between the artist, the work produced and how a work can be consumed. Monaco (1977: 176) drafts a triangle of the artistic experience, depicted here in Figure 1.

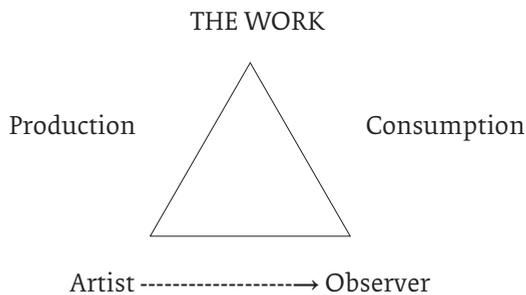


Figure 1 . Triangle of the artistic experience (Monaco 1977: 34).

The artist is at one end of the triangle and produces the work of art, which epigonically is consumed by the observer. While audio description could be located at the level of the ‘observer’, its function goes beyond mere ‘observation’ as an ultimate goal. Audio description takes a double production loop, since the script has to be produced, leading to the work being consumed for the second

time when it reaches its final destination – though it can be studied and analysed further for academic purposes.

The triangle of an audio description experience will replicate the same relationship that the source text, or film, originally had as can be seen in Figure 2.

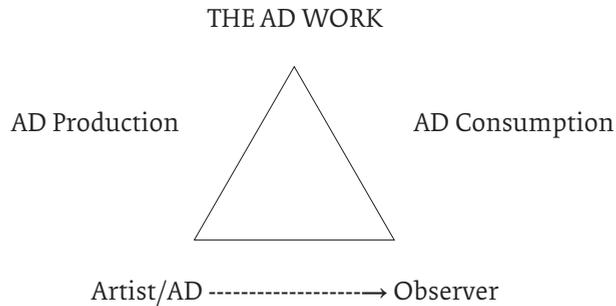


Figure 2 . Triangle of the AD artistic experience.

The audio describer has a similar function to that of the language translator, who has to render the work written in a source language into a work in a target language. Audio description too deals with a source language, in this case filmic, which has to translate into written language which will undergo further processes of production – locution and recording – before the product is ready for the user.

Interestingly we can compare the audio describer with the translator, but it seems there is a gap between the observer, or user, of a translation and that of media accessibility, i.e. audio description, subtitling for the deaf and hard of hearing and audio subtitling. The user of a translation is the anonymous audience – with a very heterogeneous background – who has no input in the style and the language in which the translation should be rendered. This does not mean translations are not analysed and there is a branch of Translation Studies which focuses on reception studies; also from time to time there is a popular outcry<sup>1</sup> regarding translation. In media accessibility we find users – and their associations – lobbying for an objective product, as if it were possible to read or mentally process verbatim subtitles, or to describe dynamic visual narrative as a still photograph. There is a mismatch between expectations and what is reasonable, or can be delivered. There is also a lack of understanding of the many processes involved in the multisemiotic transformation needed to create a subtitle by respeaking or an audio description, not to mention the many technological challenges. If an image is worth a thousand words and we have strict time limitation, how can we fulfil users requests and create an adequate audio description? We don't even have a definition for the term 'adequacy' in AD.

In the same way as reading in English means to gaze at the page from left to right and from top to bottom, a film can also be read, but how? Though image reading differs from written text reading, films can be read, according to Monaco (2009: 174), at three levels: physically, mentally and psychologically. From a physiological point of view, we should consider the way gaze patterns are arranged, with the best readers being those who have “the most efficient and extensive saccadic patterns” (Monaco 2009: 174). An example of this reading is when we follow camera framing, light and movement (Douglass and Harnden 1996; Treuting 2006). It has been shown how directorial techniques influence visual experience. Marchant *et al.* (2009: 159) have established the “commonality factor [which] affords a quantitative measure of how viewers view dynamic scenes over time”, hence it is possible to work out the number of people watching the same area of interest in a dynamic film sequence. From an ethnographic point of view, we should consider reading and understanding cultural and visual conventions present in the film. The best readers are those who have the “greatest experience and knowledge” (Monaco *ibidem*). The psychological level is the one where the two previous reading modes are integrated, and the best readers are those who are “best able to assimilate the various sets of meanings they perceive and then integrate the experience” (Monaco *ibidem*).

The three different reading experiences explain how after seeing a film people understand, and ultimately judge, the film differently – the absence of Bordwell’s “ideal viewer”. In addition, Lehman and Luhr (1999: 169) comment as to how the same film “can also have different meanings even for the same viewer at different times”. They go further stating that “no movie has one “right” meaning that every viewer can ‘get’ by approaching it ‘correctly’”. As we saw in Figure 1 in the previous section, the consumer may be a passive user but could also be an active agent – the audio describers in Figure 2 – since they can potentially participate in the process, as “the meanings of a film are produced by viewers in their interactions with it” (Lehman and Luhr *ibidem*).

From the physiological reading we can analyse and agree on the different areas of interest, where the director focused the viewer’s attention. It is possible to study and understand the ‘commonalities’, and hopefully we should be able to learn to draft areas of interest and commonalities to highlight in the audio description. Understanding cultural markers is a matter which has abundant literature, and working towards an encyclopaedic knowledge, or how to obtain it, should be one of the skills developed by the audio describer.

Now regarding the psychological reading, films are a system of communication – drawing this definition from semiotics. It is possible to understand them as a language. Metz (1974: 47) explains this as “it is not because the cinema is language that it can tell such fine stories, but rather it has become language because it has told such fine stories”. Understanding the language of films will be

a step forward in the creation of an AD language, though the different techniques used for its translation and transformation go beyond the scope of this article.

We'll follow the basic principles of Structuralism in Linguistics, where language is a system of arbitrary signs. Each sign has a signified and a signifier, and in films the same relationship can be established. While in literature the main locus of art resides in the relationship between the signifier and the signified, in films the signifier and the signified are usually identical. Looking at the word "rose" (Monaco 1977: 177), it can be modified (rosy, rosier, rosiest, risen) and therefore lead to confusion (rows, ruse, arose). In literature a 'rose' – as a motive – goes beyond the botanic flower, and it can represent concepts such as finite beauty or perfection in nature, it can even suggest a smell, or the ambiguous cause and effect: beautiful and yet thorny, and so painful and dangerous to achieve. This shift from the visual to the abstract is less common in films than in literature. Films don't suggest, they usually state. A person reading a novel or a poem can imagine whereas a person watching a film sees. While in written language, the surprise, audacity and effect lies in the difference, in the relationship and tension between the signifier and the signified, in film the same does not apply. In fact, it may be argued that the power – and popularity – of films lies in the lack of suggestion, the simplicity in understanding and its plain, or straight, reading possibilities. Following the example of the roses in Sam Mendes *American Beauty* (2000, USA) the film can be read at many levels. The title could make reference to one of the characters of the film, Angela Hayes, a beautiful and insinuating teenager who is an American beauty. It could also be stated that Angela, the American beauty, is compared to the rose American Beauty which is only grown in artificial conditions to be perfect. These are just two of the several possible interpretations (Anker 2004). Which one is the correct meaning? Which one should be narrated in the AD?

## FILM LANGUAGE

The phoneme is the basic unit of meaning in written language. This is because the change of a phoneme in a word can change its meaning, again from 'rose' to 'rise'. In films the smallest unit could be that of a frame, but since films are essentially dynamic, time also has to be taken into consideration, hence the scene could be a candidate for a basic unit. Both frame and scene may contain an infinite amount of visual information – not to mention the soundtrack. While in technical terms a single frame is the smallest physical unit, the time span – and the sound – forces us to take films as a continuum of meaning.

The language of films has been described by Monaco (1977: 178) as consisting of "short-circuit signs in which the signifier nearly equals the signified; and depends on a continuous, nondiscrete system in which we can't identify a basic unit and which therefore we can't describe quantitatively". Film analysis poses

endless difficulties given its complex nature: time, sound and images combine to create “an easy art, the cinema [which] is in constant danger of falling victim to this easiness [...] A film is difficult to explain because it is easy to understand” (Metz 1976: 47). The paradoxical nature of film language manages to communicate meaning through two different manners: denotatively and connotatively. Films denote meaning in the sense that an image or sound are what they are and there is no need for further understanding “Film is what you can’t imagine” (Monaco 1977:179). Films can reproduce with great accuracy physical realities and the physical world. This visual richness may defy description, since an image is worth a thousand words.

The connotative meaning in films is achieved by the possibility of showing any cultural representation such as dance, music, painting, titles, etc. which in itself have a wealth of symbolism attached. Films also have the connotative meaning of the spoken language through the soundtrack, and that of the written language through their titles. There are two further ways in which film have connotative meaning which are exclusive to this medium. A film has a paradigmatic connotation when the director has chosen specific cinematic aids to portray an effect, such as a camera angle or move, a colour filter, etc. Monaco (*ibidem*) suggests as examples of paradigmatic connotation a low-angle shot of a rose, which “conveys a sense that the flower is for some reason dominant, overpowering” against an overhead shot of a rose which “will diminish its importance”. The comparison of a shot with other possible shots – of the same object that we don’t see – is paradigmatic, while if the comparison is with preceding or following shots which we see is syntagmatic. Basically, we are dealing with an image and its context: how to shoot (paradigmatic) and how to present the shot (syntagmatic). The latter is where editing or montage are a basic building block in cinema language. Other basic elements of denotation and connotation in a film are drawn from Wollen (1972), who suggested three orders of cinematic signs: icon, index, symbol. The icon is when the signifier represents the signified, when it looks the same. The index is when there is an inherent relationship between the signified and signifier, and finally the symbol is the arbitrary sign which is represented by a convention. The icon is mostly visual while the symbol strives for written and spoken languages. The index is in between literary symbol and cinematic icon, and according to both Wollen (1972) and Monaco (1977), it is the way in which cinema can convey meaning.

Examples of indexical meaning are that of the turning of calendar leaves for passing time, or the sunset for the end of life or a relationship. Both examples are now too obvious and it is in the innovation of indexical meaning where a director’s creativity is at stake.

Following the route of film language and vocabulary we could also visit film rhetorics, where we could find the figures of metaphor, metonymy and synecdoche, as the three basic forms of indexical transfer of meaning. Understanding visual rhetoric figures is also required for a close reading of a film.

## AUDIO DESCRIBING FILMS

How can the above theoretical concepts be of any use when drafting audio descriptions? This is where the focus of this article lies. Following Wollen's classification of meaning in films, describing the icon should not pose too much of a problem, since the visual image coincides with its meaning. This may be where one of the many levels of understanding comes into play, and the most superficial objective reading. Apparently, according to users demands, this is the level of AD which should be on offer, a simple reading of the symbol, without any further deepening and interpretation. For example, in Guy Ritchie's *Rocknrolla* (2008, UK) three professions are presented: the lawyer, the judge and the councillor, as can be seen in the next three frames (Fig. 3 to 5):



Figure 3. Frame depicting the judge.



Figure 4. Frame depicting the lawyer.



Figure 5. Frame depicting the councillor.

In the case of the man on the phone in Figure 5, the conversation in the soundtrack (“Councillor, did you get that car?”) is used as a means of disambiguation. The man who is speaking on the phone is a councillor.

In the three previous examples the men depicted are icons for the three professions. But there are many other images which are used to help with the characterisation and aid in the understanding and enjoyment of the film. When presenting the character of Yuri, a Russian millionaire gangster, more details are provided from the very first time we encounter this character. This is the case of the office scene where Yuri meets Lenny, one of the leading characters of *Rockandrolla*. The audio description informs us that “Lenny attends a meeting”, which is true, because this is what he is doing, though there is much more visual information that is omitted, this time not due to time restrictions. The place where the meeting takes place is presented as the sequence of the following frames:



Figure 6. Frames depicting the view of the offices in the football ground Emirates Stadium, London.

Up to this moment, all we can see on screen is that Lenny and his chaperones are entering a very modern building; glass walls, rounded structures, all very modern looking. Surprisingly, though, such a large modern building is empty of employees. This fact is never mentioned, as is the case with the information regarding the cultural markers which are used to characterise the other main character of the scene (cf. Fig. 7).



Figure 7. Frame depicting a samovar in the center with Yuri, dressed in casual clothes.

The information regarding the location of the meeting, Yuri's attire – also in contrast with the rest of the men in the room – and the samovar are also spared. The meeting finally takes place in the box overlooking the impressive background of red seats (Fig. 8).



Figure 8. Frames depicting the Emirates Stadium in London.

The information offered in the audio description is “They are in a box of a Premiership stadium”. The connection, which had not been made before, is now clear. The Russian millionaire is the owner of the Premiership football team, and they are in his office. It is a direct reference to Roman Abramovich– who is a Russian millionaire and the owner of Chelsea Football Club. Should this information be mentioned in the audio description? This is what some regard as interpretation – which in Translation Studies will be considered as explicitation – and belongs to an ‘ethnographic’ reading, since cultural aspects are taken into consideration.

We have seen an iconic representation in Figures 3, 4 and 5, and an indexical representation in Figures 6 and 7. The last reading is psychological, whereas the two previous readings are integrated. The psychological reader, according to



Figure 9. Frames depicting a seduction scene.

Monaco (1977:179), is the “best able to assimilate the various sets of meanings they perceive and then integrate the experience”.

To follow with examples from Guy Ritchie’s *Rocknrolla*, Fig. 9 illustrates a seduction scene on a boat in the Thames river.

While the audio description offers the information “He touches her hand”, there is much more going on. They are meeting on an impressive yacht, moored in London, where Yuri Omovich utters the famous quote “They say there are only two days you enjoy a boat; the day you buy it and the day you sell it”. Yuri clearly intends to seduce his accountant, the exquisite Stella (Fig. 9). This is a clear example of two possible readings of this scene, one of which is the superficial reading, where the action is described as it actually occurs. The second reading implies the understanding that they are in an exclusive yacht, on the Thames, at night, that Uri is ‘chatting up’ Stella, who is equally flirtatious. In this scene we are shown the attraction Yuri has for Stella, which will be fully developed later on in the film. We also understand Stella’s character much better, since she is married to the solicitor, who is gay, and she cares only for power and money.

It is true that some scenes do not lend themselves to symbolic readings, but there are some which are rich with references and can easily be interpreted, such as the dancing scene where Stella and OneTwo – the narrator of the film – not only engage in business but also in sexual flirtation. The party where the dancing takes place is the house of a wealthy, decadent person who is the focus of the British tabloid press. This is portrayed in the film by the opening frames of the scene, where the door of the mansion is opened by a man in a smart jacket but with no trousers (Fig. 10).



Figure 10. Frame depicting a man who opens the door at the party.

If there remained any doubts about the place where the wild party takes place, we also are offered the following frames:



Figure 11. Frames depicting actions inside the party.

In the previous two frames we see the first scene in which OneTwo and his friends enter the party. A man dressed as a fox hunting jockey rides another man in the pursuit of a girl in bra and knickers wearing fox's ears and tail. The atmosphere of decadence is highlighted by the waiter, who is not wearing a shirt. Other clues are provided and they are even reinforced with the use of graphic markers, such as the following frame, where the actions of the the characters – dancing, flirting, passing information – add a semantic load to the situation where the action is taking place. This prominence of the environment as a marker in the development of the story is what Vercauteren (2010) has coined as a “spaceality AD”, and it should be considered beyond a mere description of the parts that contribute towards a meaningful interpretation which will lead to a coherent visual narrative and its enjoyment.

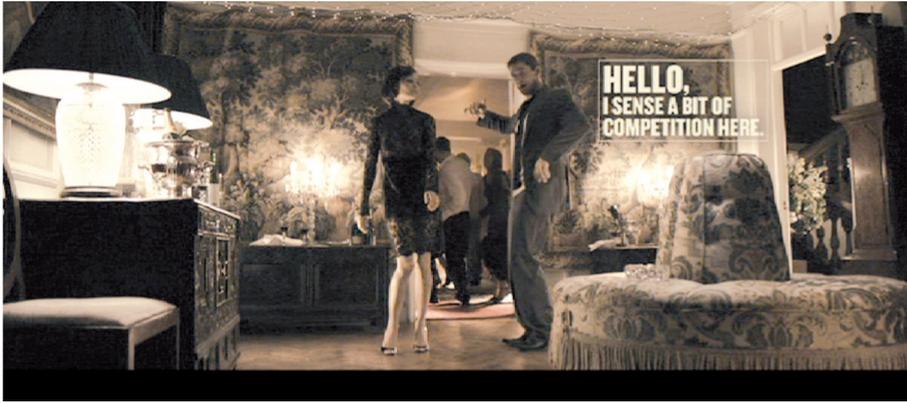


Figure 12. This frame recreates the front cover of the *Hello* magazine.

Against the possibility of having a unique and correct reading of a film and its audio description, it is important to understand that films have a complex language which require many levels of reading. This article has shown some theoretical posits for three different readings of films. Veering from the objective photographic description of static visual imagery towards an interpretation of the clues which are offered at both iconic and symbolic level.

There is a need to create ADs which are the result of a deeper analysis, understanding and interpretation of films, rather than follow existing guidelines which insist on a superficial reading such as in Figure 13. In this excerpt the concept of ‘rich’ is avoided and an ambiguous AD (“An open top car swings round the corner, its five occupants laughing and squealing”) is offered.



Figure 13. Frames depicting a convertible car with wealthy passengers, which contrasts with Fig. 14, where long rows of destitutes queing for a meal and a free bed in a San Francisco hostel, in the film *In Pursuit of Happiness*.



Figure 14. Frames depicting a long queue of homeless people.

In this scene the AD reads “the queue of people looking for places stretches up around the block. As the queue shuffles forward more people hurry into line carrying their possession in huge shapeless bags”. The message of the juxtaposition of the two images is clearly edited to show the contrast between rich and poor, though given the descriptive nature of the AD the intentionality is lost.

Audio describers should follow a course in film language, grammar, syntax and its readings, in order to avoid superficial – wrongly named objective – descriptions for richer and more meaningful readings which will do justice to the visual film’s narrative and its intentions.

## NOTES

\* This article is part of the ongoing research Project “SDH and AD: First scientific approaches and their application” (HUM2006-03653FILO), financed by the Spanish Ministerio de Educación and the Catalan Research Fund (2009SGR700).

<sup>1</sup> See for example the English Thomas Mann translation debate in Koch-Emmery 1953 and Mandel 1982.

## REFERENCES

- AENOR. 2005. *Norma UNE: 153020. Audiodescripción para personas con discapacidad visual. Requisitos para la audiodescripción y elaboración de audioguías*. Madrid: AENOR.
- Anker, R. M. 2004. “The War of the Roses: Meaning and Epiphany in *American Beauty*.” In *Catching Light: Looking for God in the Movies*, 345–363. Grand Rapids, Michigan: William B. Eerdmans Publishing Company.
- Benecke, B. and Elmar, D. 2004. *Wenn aus Bildern Worte werden*. Munich: Bayerischer Rundfunk.
- Berger, J. 1976. *Ways of Seeing*. (Reissued in Penguin 2008). London: Penguin.
- Bordwell, D. 1985. *Narration in the Fiction Film*. London: Methuen.
- Braun, S. 2007. “Audio Description from a Discourse Perspective: A Socially Relevant Framework for Research and Training” *LANS* 6: 357-372.
- Braun, S. 2008. “Audiodescription Research: State of the Art and beyond.” *Translation Studies in the New Millennium* 6: 14-30.
- Douglass, J. S. and Harden, Glenn P. 1996. *The Art of Technique: An Aesthetic Approach to Film and Video Production*. Needham Heights MA: Allyn & Bacon.
- Igareda, P. (forthcoming) “Lyrics against Imaged. Music in Audio Description”. *MonTi*.
- Koch-Emmery, E. 1953. “Thomas Mann in English Translation”. In *German Life and Letters* n.s. 6, 275-284.
- Lehman, P. and Luhr, W. 1999. *Thinking about Movies: Watching, Questioning, Enjoying*. London: Harcourt and Brace. [quotes from 2<sup>nd</sup> edition, 2003 Oxford: Blackwell Publishing].
- Mandel, S. 1982. “Helen Tracy Lowe-Porter: Once a Translator, always a Translator.” *Denver Quarterly* 17: 2.
- Marchant, P., Raybould, D., Renshaw, T. and Stevens, R. 2009. “Are you Seeing what I’m Seeing? An Eye-Tracking Evaluation of Dynamic Scenes”. *Digital Creativity* 20 (3): 153-163.
- Metz, C. 1974. *Film Language: A Semiotics of the Cinema*. New York: OUP.
- Monaco, J. 1977. *How to Read a Film* [4<sup>th</sup> edition 2009]. Oxford: OUP.
- Ofcom. 2006. [Revised version]. *Guidelines on the Provision of Television Access Services*. On line: [www.ofcom.org.uk/tv/ifi/guidance/tv\\_access\\_serv/guidelines/guidelines.pdf](http://www.ofcom.org.uk/tv/ifi/guidance/tv_access_serv/guidelines/guidelines.pdf)
- Orero, P. and Wharton, S. 2007. “The Audio Description of a Spanish Phenomenon: Torrente.” *JosTrans* 7: 164-178.
- Puigdomènech, L., Matamala, A. and Orero, P. 2007. “Audio

Description of Films: State of the Art and a Protocol Proposal.” Proceedings Intermedia 2007 Lodz Conference on Audiovisual Translation and Interpreting, 13-14 April 2007. Łódź, Polonia.

Remael, A. 2005. *Audio Description for Recorded TV, Cinema and DVD. An Experimental Stylesheet for Teaching Purposes*. www.hivt.be

Remael, A. (forthcoming). “For the Use of Sound. Audio Description and Film Sound – a Few Key Issues and a Brief Case Study.” *Proceedings of the Surrey Conference*.

Snyder, J. 2006. “Audio Description - An Aid to Literacy.” In *Languages and the Media 2006*. [http://www.languages-media.com/lang\\_media\\_2006/protected/snyder\\_joel.pdf](http://www.languages-media.com/lang_media_2006/protected/snyder_joel.pdf)

Treuting, J. 2006. “Eye Tracking and the Cinema: A Study of Film Theory and Visual Perception.” *SMPTE Motion Imaging* January, 2006: 31-40.

Vercauteren, G. 2007. “Towards a European Guideline for Audio Description.” In *Media for All. Accessibility in Audiovisual Translation*, Díaz Cintas J., Orero P. and Remael A. (eds.), 139-150. Amsterdam: Rodopi.

Vercauteren, G. 2010. “The Audio Description of Space in Film.” Paper presented at the Trieste Conference, June 2010.

Wollen, P. 1972. *Signs and Meaning in the Cinema*. New York: Viking.

# Narration or description: What should audio description “look” like?

NATHALIE MÄLZER-SEMLINGER  
University of Hildesheim, Germany

## ABSTRACT

*Like the various forms of audiovisual translation (AVT), ‘audio films’ are a hybrid, consisting of both the original film and an audio description (AD) inserted into it as a voice-over that provides the listener with a substitute for the visual content. According to the German guidelines, AD should be objective in order not to compromise the original work. This raises two questions: firstly, is it appropriate for the AD to be descriptive if one assumes that a feature film has a narrative structure, which is primarily represented on the level of images? And secondly, if the aesthetics of the film are essentially a function of its visual content, how can objectivity be reconciled with the stylistic and aesthetic objectives of the movie? This analysis is a contribution to the question of how sound and visual information of a feature film interact to tell a story and, as a consequence, what audio description should ‘look’ like in order to respect both the function of the original and the needs of the target audience.*

Describing and telling are two basic modes of representation in narrative texts, including orally presented texts in audio-visual media. But only telling is constitutive for narratives. Nevertheless, describing and telling are interdependent, as Genette argues (1981: 162f), so it may be surprising that the describing mode is rejected by some theorists. One such critic is Lukács, who discussed these two modes in his 1936 essay *Describing or telling*. Lukács, however, distinguishes between descriptions connected to the characters and the action, and descriptions whose function is reduced to a simple *effet de réel*, as Roland Barthes called it. To keep in mind this distinction seems very useful to me, also as regards audio description.

Analysing the function of description in a narrative is an important first step in the creation of audio descriptions – at least if the film being described is a feature film. As the term suggests, the goal of audio description is to furnish orally presented verbal descriptions as a substitute for the visual level of the film. Together with the soundtrack of the original, audio description provides an audio version of the original movie. It should aim at achieving dynamic equivalence with the original, as other kinds of audiovisual translations do. One of the crucial problems of audio description is the lack of time, as it has to fit in the pauses between the pieces of dialogue in the original film, like marquetry. Another problem is the series of stylistic restrictions set by the guidelines of Dosch and Benecke (1997), according to which the describer should choose brief and simple sentences. This shows once more the importance of analysing which elements of the picture are relevant and need to be described. Important criteria for deciding this are – in my view – on the one hand, the question of the interaction between the elements of the picture and the soundtrack, and, on the other, that of their function for the whole narrative. It is thus important to remember Umberto Eco's distinction between the cinematographic and the filmic code. While the cinematographic code simply codifies effects of reality produced by technical devices – such as the camera – the filmic code deals with the communication of narrative messages (Eco 1985: 250). In other words, in a feature film we are not confronted with effects of reality, but with narrative functions.

In his analysis of film dialogue, Francis Vanoye (1985: 99-118) distinguishes between two functions: a horizontal and a vertical one – or let's say a function addressing the audience, piercing the fourth wall – a distinction also adopted by Aline Remael (2008: 60). This distinction corresponds to the difference between the mimetic and narrative functions of the filmic elements, and of course, both can be fulfilled by the same element. I would like to adopt this distinction and apply it also to the visual elements of the film, adding an aesthetic and an entertaining function to the vertical ones.

In the following paragraphs I will consider the question of how audio description deals with elements of the visual (or iconic) code that have a double function – a horizontal and a vertical one – and how it copes with such ambivalence. I will use an example to illustrate the limits reached by audio

description when it has to “translate” an iconic code into a linguistic code, and what kind of stylistic means could be used in order to reach the greatest possible equivalence with the original.

The example I have chosen for this purpose is *Yella*, directed by Christian Petzold in 2007 (Germany). The narratological concept of unreliable narration can be applied to this work: Yella dies in a car accident (the car falls into the Elbe River), but the audience thinks the woman had survived. This is possible because of an almost imperceptible change of narrative perspective: we switch from the external to the internal or subjective perspective of the protagonist, who is dreaming her further life while she is actually dying. This means that we pass from an auctorial to an actorial narration – from a zero-focalisation to an internal focalisation (Martinez and Scheffel 1999: 64). The crucial information that Yella has not survived the accident, but has perished by drowning in the waters of the Elbe is withheld. We switch back to the external perspective only at the end of the film, when we finally find out about the real consequences of the accident. There is, however, a whole series of signs in this film indicating that we have left the first level of the diegesis and that we are in an intradiegetic dream of the protagonist. But we do not have enough information to interpret these signs correctly. We can try to explain them by assuming that maybe we haven’t understood everything or that the protagonist has a distorted vision of the world because of her traumatic experience. On the visual level, these signs could be interpreted as an ‘aesthetic surplus’. An example of these signs is the recurrence of the colour red – the fact that the characters are always wearing the same clothes, or the omnipresent theme of water. All these signs share a certain ambivalence and do not indicate clearly that we are in the protagonist’s dream world, composed of a day’s residues.

Furthermore, these details are characterised by a certain casualness, by means of which images, especially moving images, can show things and persons without the viewer ascribing a meaning to each detail. As abstraction is difficult to achieve on the visual level of the film, the viewer assumes that some elements of the picture are incidental, or, as claimed by Eco (1985), that they belong to the cinematographic code, but have no function within the filmic code. This concreteness of the picture gives the filmmaker the opportunity to play with the ambivalence of the elements in it and to leave the viewer in doubt as to whether these elements fulfil a narrative function, a mimetic function or both.

Now, where do we find this ambivalence in *Yella*? And is it possible for the audio description, in its current form, to preserve this ambivalence?

On the visual level, the film seems to preserve the illusion of reality, but, as we saw, the intradiegetic level of the story is made up of fragments of the diegetic level and is filled with signs that indicate to us and to the protagonist that she is drowning.

There is the recurrence of the colour red in Yella’s surroundings. The opening credits already link her name with this colour, and then there is the red of her

blouse, the red dustbin liners on the train she takes after the accident, the red of the traffic lights wherever she goes, her new friend's red car, the red bag over the bus stop sign she passed – all of these details seem to be irrelevant for the assumed diegetic level, but they become important symbols on the level of the dream. How does the current audio description deal with these ambivalent symbols?

One of these signs is picked up by the audio description – the red blouse that Yella wears for several days. In each new sequence the fact that Yella is wearing this red blouse is mentioned. However, the other signs are not picked up by the audio description – which is understandable, since even the repeated mention of the red blouse already comes across as very insistent and could seem pedantic or mysterious to the audience.

As these kinds of details of the image are shown casually and can be interpreted as 'reality filmed by chance' – as a cinematographic code without meaning for the action – mentioning them in the audio description places much too strong of a focus on them, taking away all the ambivalence. The strategy of unreliable narration is thus insufficiently taken into account by the audio description.

Here we have a problem that is due to the ambivalence of the filmic code and to the different form of presentation of iconic and linguistic codes. Of course this does not mean that linguistic codes cannot be ambivalent. But ambivalence is difficult to achieve on the basis of the guidelines for audio description. The demand of formulating simple sentences that carry just one piece of information per sentence (Vercauteren 2007: 144) may be comprehensible in consideration of the target audience of audio films, but in our – admittedly tricky – case, this demand contradicts the narrative function of the film. Casualness, ambivalence and the multiple connotations of signs cannot be achieved by a text with a paratactic syntax with one piece of information per sentence. In order to produce similar effects as the visual level of the film, the use of more complex sentences and other stylistic means should be permitted. Kluckhohn (2005) has already mentioned the importance of word order in audio description, but she also clings to the necessity of paratactic syntax, although it would be possible – in my view – to move ambivalent details to more unobtrusive places in the sentence by using hypotactic syntax to give varying degrees of emphasis to the information (Weinrich 1971: 211-237). The above-mentioned symbolic elements of the visual level, like the red lights or other 'warning signs', could be embedded in a larger context. The argument that the audience would find it difficult to understand the text is not valid here, because the goal is to move the focus away from the ambivalent details, without withholding them completely from the audience.

A similar problem arises in another scene in the film that takes place before the accident. We are confronted here with a symbolic anticipation of the further action. Yella is sitting in her father's living room, there is a pennant on the wall above her head with the inscription *Veritas – Qualitätsarbeiter* ('quality worker') – between these two words there is a labyrinthine line leading to a dead end. As far as this strange detail is perceived by the spectator, it will probably be attributed

a mimetic function and understood as a pennant of the GDR, a likely thing to be found hanging on the living room wall of an elderly man in Wittenberge, a city in the former GDR. We even get some information on the narrative level about Yella's father, about his past and the current problem of unemployment brought about by the new social order. But in view of the further story, this pennant pointing like an arrow at Yella can be interpreted as an inauspicious portent announcing her early death. This impression is reinforced by the calendar hanging on the other wall that seems to add "Your days are numbered". But Yella turns her back to it, so she can't see it, just as she won't see the truth of the accident.

Similar signs are the pictures hanging in Yella's hotel room, above her bed and all showing motifs of water that indicate her actual whereabouts: under water in the river. These signs, like the red traffic lights, are hidden in the scenery and can't be therefore easily taken into account by the audio description without surprising the audience or being interpreted as a strong hint. This kind of detail could only be embedded in a larger description of the room.

Once again, when it comes to longer descriptions, the describer is confronted with the potential lack of time. But even if he had enough time, there would still be the problem of the motivation for such a description. Is the *effet de réel* a sufficient motivation for a description of seemingly incidental details? Can the describer invoke the fact that every description provides additional information for the target audience, allowing them to be as well informed as the seeing audience?

Given the fact that the feature film is a narrative and that the describer hasn't got much time for descriptions, the elements of the picture should not be described for their own sake. The *function* of the elements to be described should play an important part in choosing them, as should the question of *how* they are to be described. For example, I think it is unnecessary to always describe in the same mechanical way what the characters look like: filmic pictures may show or not show us things or characters, but they are not able – like language – to confine their depiction to certain characteristics. It is thus definitely worth asking what motivates a description – even if the reasons are not the same as those mentioned by Lukács. This motivation should lie in connecting the descriptions to the actions of the characters.

There is such an opportunity in the scene where Yella walks through the empty train, striding past red bin-liners. The German audio description says "*Sie läuft an Müllsäcken vorbei durch den Gang*" – 'She walks down the corridor past bin-liners' (14.47). Because the colour red is not mentioned here, this element cannot be retrospectively interpreted as a sign, and cannot, therefore, adopt a narrative function. On the other hand, the insistent repetition of the fact that Yella always wears a red blouse could be toned down by avoiding the mention of the colour every time. It would be sufficient to mention it from time to time just to refer to the fact that it is the same blouse.

A good possibility for linking descriptions to the characters is the direction of their gazes. For example, in one shot Yella's glance falls on the monitor of

Philipp's laptop at the very moment the screensaver mode is activated: a huge wave breaks across the screen. It is a pity that the audio description ignores this detail, as it would have been sufficiently motivated by Yella's glance.

But audio description should not only describe what the character looks at, but also the gaze itself. Two of Yella's significant gazes are not mentioned or are not sufficiently described. The first is when Yella opens her eyes for the first time after her accident: this gaze marks the transition between the external and the internal focalisation of the narration. The eerie effect of it is produced by Yella lying there for about thirty seconds without moving. As she opens her eyes, the fact that she is staring is not mentioned; instead, the scene is described without any ambivalence "*Yella öffnet die Augen. Sie blickt in eine Baumkrone.*" (11.52) ('Yella opens her eyes. She looks up into a treetop'). A more ambivalent formulation would be "*Yellas Augen öffnen sich, starr ist ihr Blick nach oben gerichtet, in eine Baumkrone*" ('Yella's eyes open, she stares upward, into a treetop'). This formulation leaves enough room for the audience to interpret this gaze either as the look of a survivor or of an undead.

Yella's second important gaze occurs in her hotel room, where she is sitting by herself. We see her performing meaningless movements, which are nevertheless described by the audio description, e.g. "*Mit dem Fuß angelt sie nach ihren Pumps. Dann legt sie die Hände auf ihre Knie*" ('She reaches with her feet for her shoes. Then she puts her hands on her knees') (49.00). What is important is that the next moment she turns her head and looks straight into the camera. This gaze – a stylistic device that is never unmotivated in feature films (Metz 1997: 30-42) – can be explained in retrospect by the fact that there are two narrative levels here, between which a metalepsis occurs. This glance is not meant to unmask the technical device, but Yella undertakes the dangerous attempt to switch back to the objective level of narration and to face the truth – even if she does it without success. Only at the end of the film does her gaze pierce through and reveal to us that Yella is dead. The fact that this look is not mentioned in the audio description can be explained by the difficulty in naming the direction of her gaze: saying "she looks into the camera" or "at the audience" may be misleading for the audience. But it would have been easy to include in the description that her glances often look distressed – indicating that she has doubts about the reality of what she sees and hears.

I hope that this brief analysis of the film *Yella* and the few examples given from the audio film have made it clear that there is a tendency in audio description to describe visual elements of a film only because of their mimetic function and to neglect their narrative function. This phenomenon can be explained partly by the lack of time for more extensive descriptions, but it is also due to the stylistic guidelines for audio description. If the goal really is the participation of blind or visually impaired people, allowing them to share in the filmic experience, then the narrative, aesthetic and entertaining elements of the film should be better taken into account. I would, therefore, like to argue for a weakening of the strict

rules for audio description and especially to allow the use of a more complex syntax that would be able to give varying emphases to the information and create room for connotations and ambivalence.

This presupposes of course that there is enough time and that the description can be connected to the characters and their actions. If this is not possible, the question is whether the describer should be allowed to motivate the description of certain details in other ways. Concerning the pictures with water motifs, for example, could the audio describer say “Yella’s gaze falls on these pictures” in order to motivate their description? Or is it sufficient to mention that she is sitting right under them? Or would it be enough motivation if the pictures are mentioned in the context of a wider description of the whole room? For example, “*Sie betritt ein geschmackvoll eingerichtetes Hotelzimmer, wo die Farben des Dekors genau aufeinander abgestimmt sind. Alles ist in einem wässrigen Blau gehalten*” (‘She steps into a tastefully decorated hotel room with a harmonious colour scheme: watery blue is the dominant colour’). This kind of description would seemingly put the focus on the atmosphere of the room and fulfil an *effet de réel*, while casually letting drop the word ‘water’ or ‘watery’. The choice of means certainly depends on the time at the describer’s disposal. But these options are surely preferable to an unmotivated short description of the pictures, as it is the only way to preserve the ambivalence of the pictures’ function.

In view of Kautz-Vella’s study (1998), which underlines the proximity between audio film and radio drama without obliterating the differences between these two types of text, it would be conceivable to supplement the soundtrack of the original film. It could be interesting to think about adding sound or music that would provide more interpretative options for the audience – in *Yella* that could be more sounds of water or instrumental music with similar associations. As long as these adaptations are subordinate to the style of the original film, I would not consider this to be an unjustified interference with the original, as Fix and Morgner (2005: 150) fear it would be. Disregarding the narrative strategies of the original film in a schematic audio description seems much more problematic to me than such an intervention. It is for that reason that I think a weakening of the guidelines and an examination of the narrative possibilities of radio drama would greatly enrich the expressiveness of audio film. After all, the target audience of audio films is heterogeneous and there are certainly a lot of people able to understand hypotactic and more complex syntax. For them a strict implementation of the guidelines could be considered as a kind of spoon-feeding, only lowering their pleasure in the film.

## REFERENCES

- Barthes, R. 1968. "L'Effet de réel." *Communications* 11: 84-89.
- Dosch, E. and Benecke, B. 1997. *Wenn aus Bildern Worte werden. Ein Handbuch für Filmbeschreiber*. München: Bayrischer Blindenbund E. V.
- Eco, U. 1985. *Einführung in die Semiotik* (5<sup>th</sup> ed.). München: Wilhelm Fink.
- Fix, U. and Morgner, H. 2005. "Narration im Hörfilm – Theorie und Analyse." In *Hörfilm. Bildkompensation durch Sprache*, Fix U. (ed.), 117-152. Berlin: Erich Schmidt.
- Genette, G. 1981. "Frontières du récit." In *L'Analyse structurale du récit*, Barthes R. (ed.), 158-169. Paris: Seuil.
- Kautz-Vella, H. 1998. *Filme beschreiben für Blinde*. Berlin: Rubikon.
- Kluckhohn, K. 2005. "Informationsstrukturierung als Kompensationsstrategie – Audiodeskription und Syntax." In *Hörfilm. Bildkompensation durch Sprache*, Fix U. (ed.), 49-66. Berlin: Erich Schmidt.
- Lukács, G. 1936. "Erzählen oder Beschreiben." In *Kunst und objektive Wahrheit*, Mittenzwei W. (ed.), 113-165. Leipzig: Reclam.
- Martinez, M. and Scheffel, M. 1999. *Einführung in die Erzähltheorie*. München: C. H. Beck.
- Metz, C. 1997. *Die unpersönliche Enunziation oder der Ort des Films*. Paris: Klincksieck.
- Remael, A. 2008. "Screenwriting, Scripted and Unscripted Language: What do Subtitlers Need to Know?" In *The Didactics of Audiovisual Translation*, Díaz Cintas J. (ed.), 57-67. Philadelphia: John Benjamins.
- Vanoye, F. 1985. *La parole au cinéma*. Paris: Clancier-Guénaud.
- Vercauteren, G. 2007. "Towards a European Guideline for Audiodescription." In *Media for All*, Díaz Cintas J., Orero P., Remael A. (eds.), 139-149. Amsterdam: Rodopi.
- Weinrich, H. 1971. *Tempus. Besprochene und erzählte Welt* (2<sup>nd</sup> ed.). Stuttgart: Kohlhammer.

# “Why can’t you wear black shoes like the other mothers?” Preliminary investigation on the Italian language of audio description

SAVERIA ARMA  
Università Federico II, Napoli, Italy

## ABSTRACT

*According to ITC guidelines (2000), audio description should be as objective as possible, provide only relevant information, avoid any personal judgement and patronising attitude. However, since audio description is relatively young and develops at different paces worldwide, theoretical approaches and practical applications are not homogenous. In Italy, for instance, audio description is not an academic discipline yet and initiatives aiming at increasing its use are very loosely connected. However, the amount of TV audio description is not irrelevant, nor is the yearly production by no-profit associations. The language used seems to be influenced by the relatively isolated, slow and new development of audio description in Italy, but also by its literary tradition and the ‘cultural’ attitude of professionals towards spoken and written language. Through corpus-driven analysis of an Italian and an English audio description script of the film *Chocolat* (L. Hallström, 2000, USA-UK), this contribution aims at drawing attention to some features of the Italian language used, i.e. on the use of a written register and of formal and typically written syntactical structures. In addition, the article challenges relevance and objectivity in the Italian AD script which, although very far from ITC principles, generally seems to meet the expectations of the Italian blind audience.*

“If you cannot say what you mean,  
(...), you will never mean what you say”.  
*The last emperor* (1987)

## INTRODUCTION

Audio description is a relatively new practice aimed at enhancing the accessibility of different types of audiovisual products primarily to the blind and the visually impaired, using a pre-recorded or live audio track which ‘translates’ into words visual elements otherwise only accessible to sighted users. In recent times audio description has attracted scholars from various disciplines (Braun 2007; Orero 2005; Vercauteren 2007) and it has started to be taught as an academic discipline in language and translation-related faculties in a few countries (such as the UK, Spain and Belgium). However, the discussion on the position of audio description vis-à-vis Translation Studies and Audiovisual Translation is still ongoing (Braun 2007; Gambier 2004; Hernandez and Mendiluce 2004; Hyks 2005; Orero 2005). On a more practical side, ITC guidelines (2000) are rated among the main references for all audio description professionals, both in the UK and abroad, while the applicability of some recommendations – especially with regard to objectivity, linguistic choices and selection of relevant information – is still discussed (Benecke 2007; Braun 2007; Snyder 2005). Indeed, if a number of articles and presentations have dealt so far with controversial issues of audio description (Braun 2007; Fix 2005; Bourne and Himenez Hurtado 2007; Matamala and Rami 2009), only few research works have been conducted on corpora to find more extensive evidence of the characteristics of the language of audio description (Piety 2004; Salway 2007). In particular, the TIWO (Television in Words) project has analysed 91 film scripts paving the way for a definition of audio description language as a Language for Special Purposes, given the regular presence of highly statistically evident idiosyncrasies, and of grammatical and semantic patterns which are rather unusual in general language. Nevertheless, if regularity and specificity of “actions” (Salway 2007; Vassiliou 2006) are typical of the English audio description language, it may not be the same for other languages. Drawing on this first research hypothesis, this paper opts for a comparative approach for a very preliminary investigation on linguistic aspects of Italian audio description and puts forward hypotheses on the reasons for the differences found. It provides some information on blindness and audio description in Italy and it draws on a comparative analysis of AD scripts, preceded by a short description of both the film and the methodology used. The text analysis, conducted by means of the software AntConc 3.2.1, is divided into two main parts. The first one is meant to position the language(s) used in the scripts against the background of spoken and written language by focusing on the most frequent verb entries and on the sentence structure. The second part focuses on the information selection process and analyzes relevance and objectivity in the description of colours and characters.

Before analysing the scripts, it seems useful to briefly provide some information on the situation of blind and visually impaired people and of audio description in Italy, since the existing academic literature on audio description does not cover the Italian situation. In Italy, degenerative diseases are the most prominent causes of blindness and low vision. Today there are over 352,000 blind people in Italy (about 1.7% of the whole population), with the highest incidence among those aged between 65 and 90<sup>1</sup>. Italian law provides for rehabilitation services and subsidies to people who are totally blind or affected by severe low vision only; no protection is guaranteed to people with low or moderate low vision. When it comes to media access, very little attention is paid to audio description. The national 2007-2009 agreement between the Government and the Italian Public TV and Radio Broadcaster (RAI) states that the number of programs audio described should increase to up to 60% of all programs, and that RAI should make efforts to enhance the quality of medium wave radio signals on which audio description today is still received by most households (though streaming mode is also available on the RAI website)<sup>2</sup>. Apart from RAI, blind people can enjoy audio description only on a few occasions<sup>3</sup>, unless they subscribe to the service offered by the Trento-based social cooperative *Senza Barriere Onlus*<sup>4</sup>. A number of associations are striving for audio description to be considered not only for DVD and TV, but also for cinema and other forms of art (theatre, sport, museums, etc.). Among them, it is worth mentioning that associations like *Conseguenze* and *Blindsight Project* have succeeded in having some films audio described for important festivals (Rome Fiction Fest, International Film Festival in Rome, Venice Film Festival); the newly born association *CulturAbile* brought audio description for the first time ever in Italy to a dance performance by Susanna Beltrami in Rome on the 5<sup>th</sup> December 2010 during the *Prima giornata nazionale dell'arte senza barriere*.

On the research side, in Italy audio description is not yet an academic discipline: so far, no academic articles have covered audio description and only a few dissertations have dealt with it (Antifona 2002; Pini 2005)<sup>5</sup>.

No research has been conducted on the linguistic features of Italian audio description, but the idea of RAI is that audio description should use “essential, simple and clear terminology, so as to leave suitable room for personal interpretation and emotional involvement” and that it should help “understand why silent pauses occur, discover hidden *nuances* of the plot and possible hidden meanings of the characters’ behaviours”<sup>6</sup>. The blind associations generally agree on the fact that the percentage of programs should be increased and that more types of programs should be audio described<sup>7</sup>; however, from a linguistic point of view, no evidence of complaint can be found, so we believe that the blind are generally happy with the service provided and that all other audio description providers are looking at RAI’s example.

## THE FILM

The analysis is based on the comparison of two audio described versions of the film *Chocolat*, one in Italian, the other in English. The choice of this film is due to the availability of both audio described versions at the time when the research work started (mid 2009). *Chocolat* is a romantic feature film based on the 1999 novel by Joanna Harris and starred, among others, Juliette Binoche (in the role of Vianne) and Johnny Depp (in the role of Roux) as protagonists. The film is the story of a young single mother, Vianne and her little daughter Anouk who arrive in the stubbornly religious and conservative French village of Lansquenet-sous-Tannes. Here, Vianne opens the *Chocolaterie Maya* in a disused bakery facing the church just before Easter. Conte de Reynaud, the opinionated mayor of the village, makes every effort to have the *chocolaterie* go bankrupt, hoping to preserve the morality of his citizens, tempted to over-indulge by Vianne's chocolate specialities. Vianne is willing to organise a chocolate party on Easter Sunday and soon the village is split into two factions. Through the ancient art of chocolate – Vianne's origins are Maya – Vianne feeds the art of getting to know the people around her and women in particular. Josephine is one of them. She is the battered wife of Serge, a rude bar tenant; thanks to Vianne, Josephine leaves Serge and learns the art of chocolate. The other woman is Armande, Vianne's landlord; depicted as an old, distrustful and suspicious woman, soon she turns out to be open-minded, free and tolerant. Thanks to Vianne, Armande meets her nephew, whom her daughter had never allowed her to see fearing that her zest for life and refusal to conform could have influenced the child negatively. While the inhabitants of Lansquenet refuse to welcome a group of gypsy people arrived in Lansquenet, Vianne shows true curiosity and a sense of friendship towards them. Indeed, she falls in love with Roux, a gypsy who lives on a boat. After Armande's birthday is celebrated on the gypsy's boat and the fire caused by jealous Serge, Vianne and Roux seem to be forced to separate forever. Meanwhile, Conte de Reynaud desperately tries to resist chocolate temptations until Easter but he is victim of his own insecurity and preconceptions: now fooled, he eats chocolate in the window of Vianne's shop and falls asleep. When Vianne finally decides to leave Lansquenet, the soul of her dead mother suggests her to stay. Indeed, Roux comes back to her, and they finally settle in Lansquenet with Anouk. The film closes on the successful party organised by Vianne at Easter. *Chocolat* is a film about indulgence and guilt, pleasure and fears; chocolate is a mood-changing substance, and plays as a *leitmotiv* in the whole film. The smell of chocolate is the key to the heart of people and has magic properties, though it always remains something quite ordinary.

Though the story is simple and linear, the plot is more articulated: indeed, events are narrated by adult Anouk, whose voice represents an omniscient narrator who often steps into the film narration to comment, make explicit,

anticipate or postpone some clues. In addition, her voice is clearly that of an old woman, which adds something magic to the narration. The narration opens with the arrival of Vianne and Anouk in Lansquenet, the external voice explaining that Comte De Reynaud is inviting all citizens to come. Events are then narrated with the same time progression as the story, except for a few feedbacks and digressions. From an audio description perspective, some form of description is already provided by the external narrating voice and audio description should harmonise the information provided in a consistent progression. Indeed, the voice of the external narrator and the audio description are kept well separated in both versions, due to the fact that audio description is performed by a male voice in both cases.

## METHODOLOGY

The analysis has been conducted on the scripts of the audio described versions of the film *Chocolat* in Italian and English. The English script analysed was made available by RNIB, while the Italian one was produced by the social cooperative *Senza Barriere* and it is not a translation of the English version. The English version was accessed through the TIWO project after special authorisation for research purposes. The Italian script was transcribed from the film DVD itself by means of a speech recognition software (Dragon Naturally Speaking 10). The dictation was followed by a revision and editing phase; all time-in and time-out codes, information about the casting and credits were removed from the English script. Finally, the two texts were analysed using the software AntConc 3.2.1 which provides fast and reliable results for corpus driven text analysis<sup>8</sup>. However, a pure corpus-based text analysis was not the main goal of our analysis; the use of a corpus-analysis software has to be considered as purely 'instrumental' to our research purposes. Due to limited availability of data (only 2 texts were processed), all hypotheses are a preliminary research stage to be double-checked and verified against more extensive comparative research. Quantitative and qualitative observations have been combined, as in paragraph 3.3.1. Examples about sentence structure, relevance and objectivity (respectively § 3.3.1.2 and § 3.3.2) have been retrieved manually, because the software does not allow for Part-of-Speech tagging.

## TEXT ANALYSIS

### AUDIO DESCRIPTION BETWEEN WRITTEN AND SPOKEN LANGUAGE

Audio description is written to be read: any audio description production process has a written phase but the final product is an audio track produced orally. However, the language of audio description does not entirely match with spoken or with

written language (Piety 2004). These are not only linguistic phenomena but also situational, proxemic, sociolinguistic and, more generally, cultural events which require a very complex representation. If, as stated by Chafe (1982: 45-49) “while speakers interact with their audience, writers do not”, even though AD is actually meant to have a strong interactional function and a feedback effect on the target audience, it cannot be defined as ‘spoken’. Indeed, according to Chafe (1982: 45), while the written language is more detached, spoken language is more involved and “more concerned with experiential richness”. Evidence for the detachment of written language could be found in a statistically higher use of passive forms compared to spoken language (Chafe 1982). In addition, spoken language shows a higher frequency of the first person pronoun (‘I’), which is related to the control of the information flow during the communication process (Chafe 1982: 47 and *passim*). Audio description makes consistent use of the third person but usually opts for active instead of passive forms. Rhythm, intonation, volumes, voice tone variations, pauses and reformulations are prototypical of spoken language (Halliday 1985) but cannot find relevant equivalence in the written language. In fact, differences between spoken and written language cannot be reduced to the presence/absence of given (para)linguistic features but need to be ‘measured’ on the degree of representativeness of those features in the phenomena analysed<sup>2</sup>. Indeed, while Bazzanella (1994) states that differences between spoken and written language have a formal nature and Marcato (1985) says they are more situational, Lehmann (1988) stresses that linguistic phenomena should be described as part of a *continuum*: texts are not spoken OR written, they stand in between spoken-spoken texts (spontaneous, non-planned conversation) and written-written texts (formal, planned written texts). In such a perspective, also the linguistic features of audio description would be considered more from a qualitative than from a quantitative point of view. Benecke and Dosch (2004: 24) state that in audio description “formal, written language must be avoided as this hinders a lively description following the motion and life of the movie”. Similarly, Orero (2007) stresses that “AD should be a neutral discourse written to be read aloud and narrated, rather than a description of the film which by its very similitude to the filmic discourse appears part of that film”.

The following paragraphs provide some examples to show how the Italian language used in the script of *Chocolat* strongly tends to be close to written language. Therefore, it seems useful to provide some insights into spoken and written language in Italy.

For many years written language has been considered as prevailing on the spoken language, described in some sort of negative ontology (Biber 1988; Cortelazzo 1985; Dardano 1994; Sabatini 1990; Sobrero 1993; Voghera 1992). Over recent years, however, the phenomenology of spoken and written phenomena has been investigated also thanks to the analysis of various spoken language corpora. Indeed, drawing on an elaboration of the distinction made by Bazzanella (1994), the most outstanding differences between spoken and written language in

Italian have been recently listed (Arma 2007). From a phonological/graphemic point of view, written language is characterized, for instance, by a pre-organized structure, by the graphemic/visual channel, by low incidence of supra-segmental, deictic and paralinguistic features, as well as by a reduced phatic function and no immediate feedback. Italian spoken language, on the contrary, appears to be characterized by low pre-organization, by the use of the phonic/acoustic channel, by a high level of suprasegmental features and immediate feedback. These macro- and micro-features trigger syntactical, morphological, grammatical and lexical aspects, typical of spoken and written language. In particular, word order in written language is subject to organization, is characterised by a Subject+(passive)Verb+(Agent)Object structure, a large number of verbal units, explicitness, reduced variation in discourse organisation, high lexical density, higher use of specific terminology, strategic repetitions and pronominal anaphoric expressions. From the same perspective, Italian spoken language is characterized by marked word order within a sentence, dislocations and topicalisations, a high incidence of noun strings, higher variation in discourse organisation, pauses, hesitations, low lexical density and lexical choices belonging to general language, as well as repetitions, (self)corrections, modal particles, and an extensive use of the phatic function. These features refer to standard forms of spoken and written language, however some forms could be unusual in standard written language (Arma 2007), and some others are much more frequent in the standard written language than in the spoken one. Of course the range of Italian variants should also take into account sub-standard, regional, popular, dialectic, social and diaphasic language varieties (Sabatini 1985; Arma 2007).

#### TO LOOK OR TO BE?

The analysis of the scripts started with frequency lists. We decided to look at the most frequent verbal entry in both scripts, as there appears to be a correlation between the use of these words and the text semantics.

The number of tokens is 4,840 in the English version (with 1,258 different types), and 6,108 (with 1,925 types) in the Italian one. This makes 3.17 for the type/token ratio field in Italian and 3.85 in English. Although such data should not be considered as reliable *per se*, they are useful to make some hypotheses on language variation. Indeed, the first most frequent verb in the English script is the third person of the verb 'to look'<sup>10</sup>. For this reason, we have chosen to analyse the semantic environment of this KWIC both in English and in Italian. 'Looks' ranks 23<sup>rd</sup> among the most frequent words, and recurs 34 times. It is always used as a verb throughout the script; indeed, it always indicates the directionality of a character's eye movement on the scene. Table 1 below shows some collocations of the word 'looks' in the text.

AntConc 3.2.1w (Windows) 2007  
File Global Settings Tool Preferences About

Corpus Files  
chocolat inglese

Hit	KWIC	File
6	After she's gone Reynaud looks at the food - a silver c	chocolat inglese.txt
7	to the chocolaterie. She looks around the empty shop. St	chocolat inglese.txt
8	te into her hand. Vianne looks at her thoughtfully then	chocolat inglese.txt
9	nd her shoulders, Vianne looks down at the half-finished	chocolat inglese.txt
10	bed beside Anouk. Vianne looks towards an ornate ceramic	chocolat inglese.txt
11	ren gather round. Vianne looks across at the barge. A !	chocolat inglese.txt
12	Armande's house. Vianne looks at her tearfully. On his	chocolat inglese.txt
13	.. In her kitchen Vianne looks round as a sharp gust of	chocolat inglese.txt
14	spers to Yvette. Armande looks at the two empty chairs.	chocolat inglese.txt
15	Armande's garden, Vianne looks around the table with ple	chocolat inglese.txt
16	using for breath Armande looks up and catches sight of l	chocolat inglese.txt
17	nd on the wooden seat and looks at Roux sitting opposite.	chocolat inglese.txt
18	nd sits beside her, she looks up into his face and giv	chocolat inglese.txt
19	e wakes. She sits up and looks towards the trees on the	chocolat inglese.txt

Search Term: looks Word: Case: Regex: Concordance Hits: 34 Search Window Size: 50  
Total No. 1 Files Processed: 1  
Kwic Sort: Level: 0  
Save Window, Exit

Table 1. Collocations of 'looks' in the AD script of *Chocolat*.

Indeed, 'looks' mainly collocates with prepositions such as 'at', 'around', 'up', 'down', 'across', 'towards'. These results are in line with those obtained on the bigger TIWO corpus (Salway 2005, 2007). In that case, too, 'looks' was seen to be the most frequent verbal occurrence (Vassiliou 2006). Looking at the Italian script, we observe that the most frequent item is the third person singular of the verb 'to be' (i.e. 'è', from the infinitive 'essere'), ranking 25<sup>th</sup> with 33 occurrences in the script. From a contrastive perspective, this could mean that the directionality of eye movement is not primarily outlined in the Italian AD script. In order to check this hypothesis, we looked at the occurrences of the Italian verb *guardare*, which primarily translates the English verb 'to look'<sup>11</sup> and occurs only 5 times in the script:

AntConc 3.2.1w (Windows) 2007  
File Global Settings Tool Preferences About

Corpus Files  
chocolat italie

Hit	KWIC	File
1	o strano, protettivo, si guard intorno si avvicina ad t	chocolat italiano.txt
2	rizzata a lui. Subito la guard stupito È l'invito all	chocolat italiano.txt
3	uta signora la Bierau e guard gli altri amici che si	chocolat italiano.txt
4	rmai dovunque. Vianne si guard intorno sgomenta in cert	chocolat italiano.txt
5	uce Serge. Il sindaco lo guard con aria interrogativa.	chocolat italiano.txt

Search Term: guarda Word: Case: Regex: Concordance Hits: 5 Search Window Size: 50  
Total No. 1 Files Processed: 1  
Kwic Sort: Level: 0  
Save Window, Exit

Table 2. Occurrences of *guardare* in the AD script of *Chocolat*.

However, these results are not sufficient to prove our hypothesis, since the semantic spectrum of the English ‘to look’ and the Italian ‘guardare’ is quite extended. Indeed, 34 occurrences of ‘looks’ should be added to 6 hits of ‘stare\*’, 5 hits of ‘gaze\*’, 6 hits of ‘watch\*’, 1 hit of ‘glance\*’ and ‘peep\*’; 4 expressions contain ‘catching sight’; 14 hits are registered for ‘see\*’ but do not express eye movements<sup>12</sup>. This makes a total of 56 verbal hits expressing eye movements on the scene. The Italian version shows 5 hits of *guarda* and 10 hits of *osserva*. However with a closer analysis, it becomes clear that the semantic field is split into a variety of verbs and expression, such as *fissa\** (‘stares’, 11 hits), *rivolge\** (1 hit), *squadra\** (3 hits), *scruta\** (4 hits), *sbircia\** (1 hit), *volge lo sguardo* (1 hit\*), *lanciare uno sguardo* (3 hits), *posare lo sguardo* (1 hit), *abbassare lo sguardo* (1 hit), *far scorrere lo sguardo* (1 hit), *scambiarsi uno sguardo* (1 hit). If we look at the KWIC *eye\** (in Italian *occh\**), we notice that it produces 14 hits (among them ‘to catch the eye on’, ‘to keep one’s eye on’, ‘to close eyes’, ‘to open eyes’). The Italian *occh\** produces 19 hits; among them, we find *lanciare un’occhiata*, *posare gli occhi su*, *cercare con gli occhi* and *incrociare con gli occhi*. As a preliminary observation, evidence shows that where the English audio description makes consistent use of ‘look\*’, the Italian often chooses different high register verbal expressions – unusual in spoken language and rather typical of written literary language. To this end, various written and spoken language corpora were interrogated (CoLFIS, CORIS/CODIS, CLIPS)<sup>13</sup>.

#### SENTENCE STRUCTURE

From the syntactical point of view, both English and Italian sentences tend to respect the Subject+(passive)Verb+(Agent)Object; however while there are few secondary clauses in the English script<sup>14</sup>, the Italian one is characterised by many embedded secondary clauses, mostly noun phrases or temporal/causal clauses, typical of the Italian written language. Indeed, in the Italian script there are 339 subordinate clauses out of a total number of 853 clauses. Table 3 shows a few parallel examples:

English AD	Italian AD	Back translation
1a) As they pass the chocolaterie, Vianne is arranging the window display. She waves to them.	1b) <i>Vianne intanto sta sistemando la vetrina e vedendoli passare sulla piazza li saluta con un gesto della mano affacciandosi alla porta.</i>	In the meantime, Vianne is arranging the window display and as she sees them passing by in the square, she waves at them leaning by the door
2a) Reluctant to speak, they silently nod their appreciation of the succulent food.	2b) <i>Tutti sorridono estasiati, ogni tensione si scioglie finché esplode un riso liberatorio pieno di gioia semplice e autentica.</i>	Everybody is smiling, happily, the tension is gone bursting into a liberating, joyful, simple and genuine laughter.
3a) At home, the Comte de Reynaud takes a summer dress from the wardrobe and starts to cut it to ribbons with a huge pair of scissors.	3b) <i>In quel mentre il conte colto da un'irrefrenabile attacco di rabbia fruga nell'armadio della moglie per tagliuzzare i suoi vestiti con un grosso paio di forbici sperando così di vendicarsi per il tradimento subito.</i>	In that moment, the Comte de Reynaud is prey of an unstoppable rage and is searching his wife's closet to cut through all her clothes with a big pair of scissors. Doing so, he hopes to get revenge for the betrayal.
4a) As he draws the sweet morsel into his mouth his eyes close and he sighs with ecstasy.	4b) <i>La lingua allora come mossa da volontà propria fuoriesce a leccare quel dolce con una sensazione meravigliosa, paradisiaca.</i>	As if moved by its own will, the tongue reaches for the sweet and licks it, with a heavenly, blissful sensation

Table 3. Comparison of the English and the Italian sentence structure in the two versions of the audio description.

Interestingly, many examples (41 out of 126 adjectives in noun groups) in the Italian script show that the adjective in the noun group is put before the noun it refers to, such as in *graziosa signora* ('nice woman'), *strani disegni* ('strange

drawings'), *candidi fiocchi di neve* ('white snow flakes'), *aperta campagna* ('vast countryside'), *pessime condizioni* ('bad conditions'), *antiche terre* ('ancient places'), *vigorosa stretta di mano* ('strong handshake'), *vecchi fogli di giornale* ('old newspapers' pages), *delicati petali di rosa* ('delicate rose petals'), *inquietanti scene di morte* ('worrying death scenario'). Indeed the standard position of adjectives in Italian in the noun group is after the noun they refer to (Scarano 2000). Though both in spoken and written language the percentage of adjectives put before the noun is around 13% (Scarano 2000), differences in their use do not emerge from purely quantitative analysis but from considerations on their 'qualitative' use. In spoken language, adjectives put before the noun are frequent in formulaic expressions and generally with adjectives which explicitly require that syntactical position, where they lose their original meaning and therefore cannot be considered as lexical choices. In written language, the adjectival position before the noun stands for a more creative and free use of the language. Ultimately traditional grammar assigns to the choice of this position a non-restrictive function and a subjective value. Postponed adjectives in the noun group are assigned a more restrictive function in the identification of the properties described (Scarano 2000).

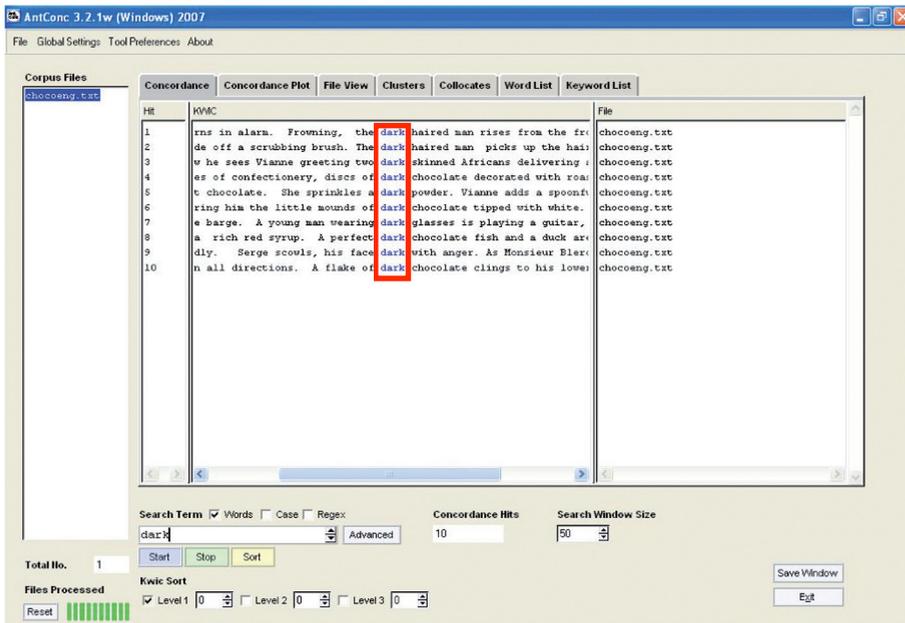
#### INFORMATION SELECTION: RELEVANCE AND OBJECTIVITY

Relevance in audio description is a central but controversial issue. On the one hand, only relevant information should be selected and provided in a way that fits into the time spans allowed by the dialogues and the sound track; on the other hand, the same film could be audio described in many different ways, since audio description is an "inexact science" and "there are many ways of getting to an end result" (Hyks 2006). Indeed, practices change from country to country and even from company to company. Audio describing indeed appears to be a complex activity which requires "intense assessment and decision processes" (Braun 2007: 6). Taking into account the cognitive load triggered by audio description could help to improve the information selection process. Often, *how* information is selected depends greatly on the skills of the audio describer, on his/her personal taste, or on the requirements of the target audience. In addition, as stressed by Pujol and Orero (2007) the output is influenced by the individual interpretation of reality. Once the primary information has been selected and prioritised, another problem is how objective or subjective is the description provided, to what extent the audio describer can add information and what type of information he/she can add. In the following paragraph, we will see how the Italian audio description of *Chocolat* deals with the selection of relevant information and to which extent the description can be considered subjective; to this purpose, we will deal with the description of colours and characters.

## COLOURS

According to ITC guidelines (2000), colours need to be audio described for a fully enjoyable description of objects and characters on the screen. For this reason, as an example, we have compared the occurrences of 'black/dark' in the English script to *ner\*/scur\** in the Italian one. Dark is the colour of chocolate, dark are the clothes of many inhabitants in Lansquenet, dark are the shoes of all women in the village, dark is the storm announcing the arrival of Vianne and Anouk. 'Dark' is a recurrent colour and there are reasons to believe that as it is a semantically relevant element it should not be neglected.

In the English script, 'dark' occurs 10 times, against the three occurrences of 'black'. It is used, in nearly all cases, to characterise an object or a character (e.g. dark-haired, dark skinned). In a single case only is it used in some kind of dead metaphor to tell that a person is very angry:



Hit	KWIC	File
1	rns in alarm. Frowning, the dark haired man rises from the fr	chocoeng.txt
2	de off a scrubbing brush. The dark haired man picks up the hair	chocoeng.txt
3	w he sees Vianne greeting two dark skinned Africans delivering	chocoeng.txt
4	es of confectionery, discs of dark chocolate decorated with real	chocoeng.txt
5	chocolate. She sprinkles a dark powder. Vianne adds a spoonf	chocoeng.txt
6	ring him the little mounds of dark chocolate tipped with white.	chocoeng.txt
7	a barge. A young man wearing dark glasses is playing a guitar,	chocoeng.txt
8	a rich red syrup. A perfect dark chocolate fish and a duck ar	chocoeng.txt
9	dly. Serge scowls, his face dark with anger. As Monsieur Bler	chocoeng.txt
10	n all directions. A flake of dark chocolate clings to his lips	chocoeng.txt

Table 4. Occurrences of 'dark' in the AD script of *Chocolat*.



Hit	KWIC	File
1	ith a wooden spoon. A little black and white dog sniffs eagerly	chocoeng.txt
2	Outside, three old ladies in black peer at the window display.	chocoeng.txt
3	er stick Armande crosses the black and white tiled floor. (doc	chocoeng.txt

Table 5. Occurrences of 'black' in the AD script of *Chocolat*.

The same colour, in the Italian script, occurs fewer times than expected:

Hit	KWIC	File
1	ette vestite completamente di nero si fanno osservare la vestitu	chocolat italiano.txt
2	lamme che si alzano nel cielo nero, Roux la tiene stretta nelle	chocolat italiano.txt
3	anno addensando grosse nuvole nere. Intanto il vento insidioso	chocolat italiano.txt

Table 6. Occurrences of *ner\** in the AD script of *Chocolat*.

Out of 7 occurrences, only 3 refer to the colour ‘black’ (the query ‘*ner\**’ was made without specifying masculine or feminine preference). *Scur\** only occurs once.

Another crucial colour in the story is red. Red are the clothes of Vianne and Anouk when they first arrive in Lansquenet, red is the colour of Vianne’s shoes (against the black shoes of all other women in the village), red is the colour of the flames in the fire caused by Serge to kill Vianne and Josephine, *Roux* in French means ‘red’ and is the name of the gypsy man whom Vianne falls in love with. The next table shows the hits of ‘red’ in the text:

Hit	KWIC	File
2	ning against the wind the red-cloaked figures, a woman	chocolat inglese.txt
3	es as Vianne, in a bright red apron, stands on a ladder	chocolat inglese.txt
4	out a new sign, painted red with letters of bright go	chocolat inglese.txt
5	awled asleep on the sofa, red-faced and unshaven. Yvette	chocolat inglese.txt
6	picks up a jar of bright red jam and lifts the silver	chocolat inglese.txt
7	wn the stone steps in her red high heels. Luc smiles at	chocolat inglese.txt
8	cks the pediment with her red shoes. She lashes the sta	chocolat inglese.txt
9	rying a bunch of pink and red roses, Serge ignores a c	chocolat inglese.txt
10	g cherries bob in a rich red syrup. A perfect dark chr	chocolat inglese.txt
11	o Armande. She unties the red ribbon. Armande unwraps he	chocolat inglese.txt
12	trees the sky is glowing red. She looks along the river	chocolat inglese.txt
13	wakes Anouk. She sees the red cloak. She tries to put ti	chocolat inglese.txt
14	s the string of a bobbing red balloon.	chocolat inglese.txt

Table 7. Occurrences of ‘red’ in the AD script of *Chocolat*.

Occurrences of Italian *ross\** (‘red’) are shown in next table:

Hit	KWIC	File
1	enco, avvolte in mantello rosso, arrivano al villaggio due forest	chocoita.txt
2	l cioccolato la polverina rossa. L'anziana afferra la tazza poc	chocoita.txt

Table 8. Occurrences of *ross\** in the AD script of *Chocolat*.

The Italian script contains only 2 occurrences of the hit *ross\**, one referring to the coat of Vianne and Anouk, the other to the pepper which Vianne puts into the chocolate. The Italian script seems less concerned with colour details and their semantic frequency and importance throughout the text.

#### DESCRIPTION OF CHARACTERS

The most outstanding differences in the approach to AD are in the characters description and in the selection of relevant physical or mental states to audio describe. As an example, we will show how Roux is described the first time he appears in the story. The second box contains the description of Serge, Josephine’s husband.

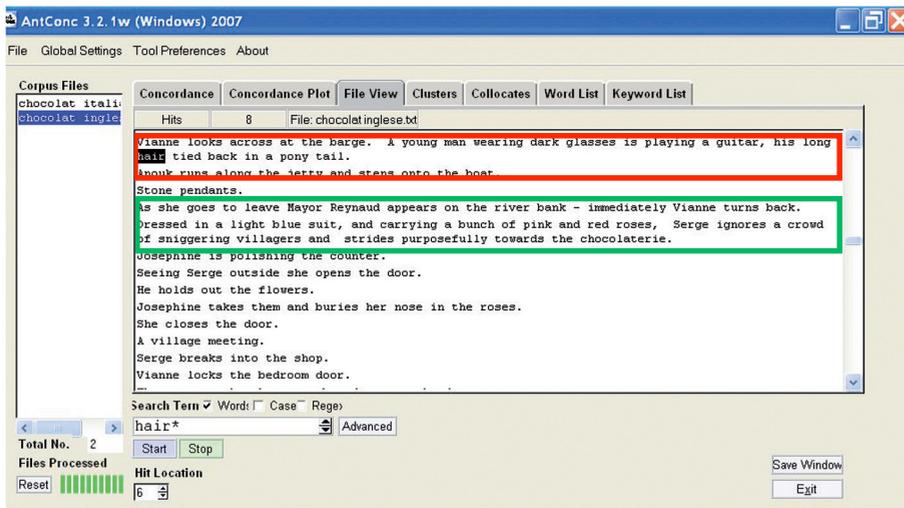


Table 9. Description of Roux and Serge in the English audio description.

Serge is described as “dressed in a light blue suit”, bringing “pink and red roses”; the villagers around him are “sniggering”. In addition, the English audio description describes the most relevant features of Roux’s appearance: “dark glasses”, “long tied hair”, “young man”.

The Italian version describes the scene as follows:<sup>15</sup>

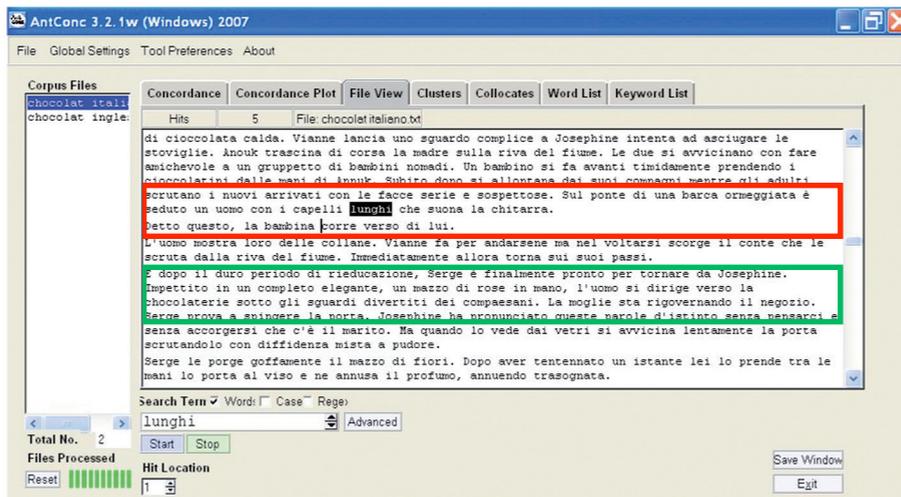


Table 10. Description of Roux and Serge in the Italian AD script of *Chocolat*.

The Italian version only describes Roux as a man with long hair, playing a guitar on a small boat. Serge is described as “*impettito in un completo elegante*” (‘proudly wearing an elegant suit’); the colour of his roses is not mentioned at all (“*un mazzo di rose in mano*”, i.e. ‘carrying a bunch of roses’). Moreover, the Italian version adds a narrative explication for Serge’s behaviour, and stresses that after very hard time for re-education to good manners, Serge is ready to go back to his wife.

## CLOSING REMARKS

This article drew on preliminary observations on the comparison of two audio described versions of the same film, i.e. *Chocolat*. We have stressed that the development of audio description in Italy does not only depend on the relatively isolated position of the country with regard to audio description, but it also seems to rely on the written literary tradition of the country. Both from the lexical and syntactical point of view, the Italian language used in this script tends more to variety and complexity than it does to regularity and simplicity. On the lexical side, we have analysed the collocations of the verb ‘to look’ in English and *guardare* (‘to look’) in Italian: the analysis has stressed that the Italian semantic spectrum is covered by a number of linguistic choices (mostly belonging to higher written register). On the syntactical side, we have shown that the Italian script contains a higher number of secondary clauses if compared to the English one. As to objectivity and information selection, we have shown how the Italian audio description is filled in with subjectively orientated interpretations rather than with objective observations and maintains a more ‘narrative’ mood if compared to the English one. More corpus-based comparative analysis and

reception studies would be needed, with specific regard to the relation between written language and narration, the 'hearing habits' of blind people (the role of radio could contribute to the analysis) and the satisfaction of the audience towards different approaches to audio description. This article is intended to be a first step in this direction.

- 1 [http://www.sedesoi.com/index\\_eng.asp](http://www.sedesoi.com/index_eng.asp) (accessed 1st December 2009).
- 2 In the first semester of 2009 RAI audiodescribed 6 films/ fiction for TV and 17 sit-comedies (accounting for 120 episodes).
- 3 Apart from TV, AD in Italy is available at the Teatro di Messina (Sicily). In summer 2009 an AD team coordinated by Elena di Giovanni provided two ADs for opera in Macerata. The association Cinema senza Barriere ('Cinema without barriers') yearly organizes barrier-free performances and has audiodescribed many films.
- 4 The blind associates are asked to pay an annual subscription fee; they are entitled to receive audio described films on a DVD containing the AD track.
- 5 As far as we know, no PhD thesis has dealt so far with AD. The first one should be discussed by the author of this article in 2011-2012.
- 6 Own translation of "[...] l'uso di una terminologia chiara e precisa, ma essenziale, che lasci il giusto spazio all'interpretazione personale e alla partecipazione emotiva di chi ascolta." And: "[...] capire il 'perché' di momenti di silenzio, scoprire "sfumature" della storia e delle relazioni che legano i personaggi".
- 7 [http://www.superabile.it/web/it/CANALI\\_TEMATICI/Viaggi\\_e\\_Tempo\\_libero/Inchieste\\_e\\_Dossier/info561554978.html](http://www.superabile.it/web/it/CANALI_TEMATICI/Viaggi_e_Tempo_libero/Inchieste_e_Dossier/info561554978.html) (accessed 1st December 2009).
- 8 This is a freeware for Windows, Macintosh OS X and Linux. It can be downloaded from Laurence Anthony Laboratory website (<http://www.antlab.sci.waseda.ac.jp/software.html>).
- 9 Source: <http://www.parlaritaliano.it/parlare/> (accessed 1st December 2009).
- 10 Entries from 1 to 24 include grammatical words like (from the most to the less frequent) 'the', 'a', 'and', 'her', 'of', 'to', 'at', 'in', 'on', 'with', 'into', 'The'; characters, personal and possessive adjectives/ pronouns rank before the entry 'looks' and respectively: 'Vianne' (97 hits), 'his' (81 hits), 'He' (46 hits), 'She' (46 hits), 'she' (46 hits), 'Anouk' (41 hits), 'chocolate' (41 hits), 'Reynaud' (36 hits), 'Josephine' (35 hits), 'he' (35 hits).
- 11 To this end, we used the Oxford-Paravia (2006) bilingual dictionary.
- 12 To find out the English verbs covering the semantic field of the verb 'to look', we used the Oxford Dictionary of English (1971). To find out the Italian verbs covering the semantic field of guardare, we used the Italian Dictionary by Devoto and Oli (2003).
- 13 The corpora that we mentioned are respectively accessible at: <http://www.istc.cnr.it/material/database/colfis/>, [http://corpus.cilta.unibo.it:8080/coris\\_ita.html](http://corpus.cilta.unibo.it:8080/coris_ita.html) and <http://www.cirass.unina.it/> (accessed 29 December 2009).
- 14 AntConc does not provide for Part-of-Speech tagging nor for counting primary and secondary clauses. The predominance of secondary clauses in the scripts has been verified against manual check and could therefore be subject to errors. In the English script, only 147 out of 707 are subordinate clauses.
- 15 Back translation of the description of Roux: "The villagers peer at the newcomers with serious and suspicious glances. On the deck of a moored boat, a man with long hair, playing guitar, is sitting. Once said that, the girl runs towards him"; back translation of the description of Serge: "And after the long re-education period, Serge is now ready to go back to Josephine. Proudly wearing an elegant suit, carrying a bunch of roses, the man heads towards the chocolate shop among the fellow citizens' amused glances. His wife is tidying up the shop".

- Antifona, C. 2002. *L'audiodescrizione. Uno strumento per l'integrazione sociale del non vedente*, unpublished MA dissertation. University of Rome 'La Sapienza'.
- Arma, S. 2007. *Dal parlato al (tra)scritto. La resocontazione parlamentare tra stenotipia e riconoscimento del parlato*, unpublished MA dissertation, University of Bologna, SSLMIT.
- Bazzanella, C. 1994. *Le facce del parlare. Un approccio pragmatico all'italiano parlato*. Scandicci (Firenze): La Nuova Italia.
- Benecke, B. 2007. *Audio Description: Phenomena of Information Sequencing*. EU-High-Level Scientific Conference Series MuTra 2007 LSP Translation Scenarios: Conference Proceedings. [http://www.euroconferences.info/proceedings/2007\\_Proceedings/2007\\_Benecke\\_Bernd.pdf](http://www.euroconferences.info/proceedings/2007_Proceedings/2007_Benecke_Bernd.pdf) (last accessed 10/11/2009).
- Benecke, B. and Dosch, E. 2004. *Wenn aus Bildern Worte werden*. Munich: Bayerischer Rundfunk.
- Biber, D. 1988. *Variation across speech and writing*. Cambridge: CUP.
- Bourne, J. and Jiménez Hurtado, C. 2007. "From the Visual to the Verbal in two Languages: A Contrastive Analysis of the Audio Description of The Hours in English and Spanish". In Díaz Cintas, J., Orero, P., and Remael, A. (eds.). *Media for All. Accessibility in Audiovisual Translation*, 175-188. Amsterdam: Rodopi.
- Braun, S. 2007. "Audio Description from a Discourse Perspective: A Socially Relevant Framework for Research and Training." In *Papers from the Centre for Translation Studies*. Paper 1. <http://epubs.surrey.ac.uk/translation/1> (last accessed 01/12/2009).
- Chafe, W. 1982. "Integration and Involvement in Speaking, Writing and Oral Literature." In *Spoken and Written Language. Exploring Orality and Literacy*, Tannen D. (ed.), 36-53. Norwood, NJ: Ablex.
- Cortelazzo, M. 1985. "Dal parlato al (tra)scritto: i resoconti stenografici dei discorsi parlamentari." In *Gesprochenes Italienisch in Geschichte und Gegenwart*, Holtus, G. and E. Radtke (eds.), 87-118. Tübingen: Narr.
- Dardano, M. 1994. "Profilo dell'italiano contemporaneo." In *Storia della lingua italiana, II, Scritto e parlato*, Serianni L. and Trifone P. (eds.), 343-430. Torino: Einaudi.
- Devoto, G. and Oli, G. 2002-2003. *Il dizionario della lingua italiana*. Firenze: Le Monnier.
- Il dizionario inglese-italiano, italiano-inglese*. 2006. Torino: Paravia, Oxford: Oxford University Press.
- Fix, U. 2005. *Hörfilm. Bildkompensation durch Sprache*. Berlin: Erich Schmidt.
- Gambier, Y. 2004. "La traduction audiovisuelle: un genre en expansion." In *Meta* 49 (1): 1-11.
- Halliday, M. A. K. 1985. *Spoken and Written Language*. Victoria, Australia Deakin University, translation by Dionisi, A., *Lingua parlata e scritta*, Roma: La Nuova Italia.
- Hernandez, Ana / Mendiluce, Gustavo. 2004. "Audes: Translating Images into Words for Spanish Visually Impaired People." *Meta* 49 (2): 264-277.
- Hyks, Veronika. 2006. "Translation versus Origination. Creativity within Limits". *Speech Delivered at Media for All 2006*. Berlin, 2006. *ITC Guidelines on Standards for Audio Description* (2000): <[http://www.ofcom.org.uk/tv/ifi/guidance/tv\\_access\\_serv/archive/audio\\_description\\_stnds/](http://www.ofcom.org.uk/tv/ifi/guidance/tv_access_serv/archive/audio_description_stnds/)>.
- Lehmann, C. 1988. "Towards a Typology of Clause Linkage." In

- Clause Combining in Grammar and Discourse, Haiman, J. and Thompson S. (eds.). Amsterdam; Philadelphia: Benjamins.
- Marcato, G. 1985. "Italiano parlato, comunicazione di base e oralità." In *Gesprochenes Italienisch in Geschichte und Gegenwart*, Holtus, G. and Radtke E. (eds.), 24-41. Tübingen: Narr.
- Matamala, A. and Rami, N. 2009. "Análisis comparativo de la audiodescripción española y alemana de *Good-bye Lenin*." *Hermeneus* 11: 249-266.
- Nencioni, G. 1976. "Parlato-parlato, parlato-scritto, parlato-recitato." In *Strumenti critici* 29, reprint. in id. (1983), *Di scritto e di parlato*, Bologna: Zanichelli.
- Orero, P. 2005. "Audio Description: Professional Recognition, Practice and Standards in Spain." *Translation Watch Quarterly* 1: 7-18.
- Orero, P. and Wharton, S. 2007. "The Audio Description of a Spanish Phenomenon: *Torrente 3*." *The Journal of Specialised Translation* 7: 164-178.
- Orero, P., and Pujol, J. 2007. "Audiodescription Precursors: Ekphrasis and Narrators." *Translation Watch Quarterly*: "Media Accessibility Issues and Standards of Audiovisual Translation" 3(2): 49-60.
- The Oxford English Dictionary*. 1971. Oxford: Oxford.
- Piety, P. 2004. "The Language System of Audio Description: An Investigation as a Discursive Process." *JVIB* 98: 8, 453-469.
- Pini, L. 2005. *L'audio-descrizione per disabili visivi*, unpublished MA Dissertation, University of Bologna, SSMIT Forlì.
- Sabatini, F. 1990. "Una lingua ritrovata: l'italiano parlato." *Studi latini e italiani*, IV: 215-234.
- Sobrero, A. 1993. *Introduzione all'italiano contemporaneo*. La
- variazione e gli usi*. Roma-Bari: Laterza.
- Salway, A. 2007. "A Corpus-Based Analysis of Audio Description." In *Media for All: Subtitling for the Deaf, Audio Description and Sign Language*, Díaz Cintas, J., Orero, P. and Remael, A. (eds.), 151-174. Amsterdam: Rodopi.
- Scarano, A. 2000. *Annotazioni sulle relazioni sintattiche, semantiche e informative degli aggettivi qualificativi. Uno studio su corpora di italiano scritto e parlato*. University of Florence. [www.lablita.dit.unifi.it/preprint/preprint-00bcollo5.pdf/.../at\\_download](http://www.lablita.dit.unifi.it/preprint/preprint-00bcollo5.pdf/.../at_download) (last accessed 01/12/2009).
- Snyder, J. 2005. "Audio Description. The Visual Made Verbal across Arts Disciplines - across the Globe." *Translating Today* 4: 15-17.
- Vassiliou, A. 2006. *Analysing Film Content: A Corpus-Based Approach*, dissertation submitted for the Degree of Doctor of Philosophy, Department of Computing School of Electronics and Physical Sciences, University of Surrey.
- Vercauteren, G. 2007. "Towards a European Guideline for Audio Description." In *Media for All: Subtitling for the Deaf, Audio Description and Sign Language*, Díaz Cintas, J., Orero, P. and Remael, A. (eds.), 215-229. Amsterdam: Rodopi.
- Voghera, M. 1992. *Sintassi e intonazione dell'italiano parlato*. Bologna: Il Mulino.



# AD reception research: Some methodological considerations\*

AGNIESZKA CHMIEL AND IWONA MAZUR  
Adam Mickiewicz University, Poznań, Poland

## ABSTRACT

*AD reception research, or collection of feedback from the blind and partially sighted as the target audience of audio described films, seems to be one of the best sources of information to be applied when creating both AD standards and audio descriptions proper. This paper presents experiences gained by the authors when conducting two reception studies. The first one involved a questionnaire distributed to 18 viewers with vision dysfunctions immediately after two screenings of audio described films. The other one is a larger-scale work-in-progress, whose results will be applied in the development of Polish AD standards reflecting the preferences of the blind and visually impaired viewers in Poland, where the participants are being interviewed and presented with AD samples. The authors discuss various methodological issues, including problems with obtaining a sufficient number of participants, reflecting feedback from visually-impaired AD consultants in the surveys and discovering user preferences. It is suggested that responses concerning objectivity or subjectivity of descriptions should be elicited indirectly (implicitly) rather than directly (explicitly) and that research results are more meaningful if interviews involve comprehension questions and AD samples.*

## INTRODUCTION

Audio description (AD) as an accessibility technique that offers access to video materials to the blind and partially sighted audiences has recently seen a boom not only in the number of audio described films and other works of art but also in the research area. The first ever periodical conference devoted fully to AD (Advanced Research Seminar on Audio Description organised by UAB in Barcelona) grew from approximately 25 participants during the first event in 2007 to over 80 four years and two seminars later. Many conferences focusing on audiovisual translation now host fully-fledged sessions on AD (to name but a few: Media For All Conference, Points of View in Kraków, Languages and the Media in Berlin). The blind audiences welcome the growing number of available audio descriptions, but is research on AD also advantageous to the ultimate audio description target group? How can the visually impaired people benefit from the growing interest of translation scholars and the rising number of scientific publications on audio description? In fact, many aspects of AD research have or may have a direct bearing on the practicalities of AD. For instance, researching the use of sound in audiovisual productions and its narrative function may bring about better and more sound-considerate audio descriptions (Van der Heijden 2007). Cognitive research focusing on memory load can make audio descriptions more acceptable and even more accessible to the target audiences by avoiding perceptual overloads identified in experimental research (cf. research project by Fresno 2011). However, in the current stage of AD practice development the greatest applicability is generated by research focusing on reception studies.

Reception studies in AD can directly contribute to the quality of audio description. They focus on the target audience of audio description and help collect feedback on various AD strategies and solutions. By learning the blind and partially sighted viewers' preferences, AD authors can more satisfactorily meet their expectations. The methodology of reception studies involves techniques that collect views and opinions of the visually impaired respondents. These might include interviews, questionnaires, online surveys, focus groups. They might be based on questions only or may include examples of audio descriptions to be evaluated by the respondents. As a result, preferences of the target groups regarding various aspects of AD are identified, which translates into recommendations, guidelines or standards for creating practical AD.

According to the ISO 9000 certificate broadly used for quality assurance in all areas of life, quality is whatever the customer requires and perceives as good quality (International Organization for Standardization). The rationale of reception studies in AD research is in line with this approach. In other words, since the blind and the partially sighted are the target audiences, or clients of audio description, let them tell audio describers what they want, what they can accept and what works best for them instead of simply presenting them with a product that is perceived as best by the AD author.

This incorporation of the visually impaired voice and opinion in the creation of AD may take various forms. As underlined above, reception studies lead to quantifiable preferences that may be later reflected in the AD preparation (this is the case with the UK's Audetel project, see below). Additionally, VIPs (visually impaired persons) may be involved in the process of creating AD from the very beginning – this is the case of Bayerischer Rundfunk in Germany where each AD team includes two sighted audio describers and one visually impaired expert (Dosch and Benecke 2004). Some AD scripts in Poland are also created by a mixed team (one sighted and one blind audio describer – Szymańska: personal communication). Both techniques have their strengths – reception studies let audio describers learn about general preferences of large target groups (including people with much and little exposure to and experience with AD) while ADs created with visually impaired team members reflect the blind expert's long-term experience with such an accessibility technique.

This paper will present some methodological considerations regarding reception research in AD gained by the authors when conducting two reception studies. It might be obvious that we need to ask the blind about their preferences in order to deliver good AD but it is much less clear how to elicit such feedback. Thus, the main question is not “if” but “how” to do it. And the latter question is the one this paper hopes to address successfully.

The section below will focus on discussing previous reception studies in the area of audio description and the methodologies employed. Subsequent sections will describe the AD-Verba project involving reception studies on AD in Poland and the difficulties experienced in the area of methodology.

## AD RECEPTION STUDIES – STATE OF THE ART

The unquestionable leader in the development of AD and in reception projects is the UK that can boast two wide-ranging studies of the blind audiences' preferences: the Audetel project and “Bollywood for all”. These were undertaken by the regulators and consumer associations rather than researchers themselves. We will describe them below focusing on the methodology rather than on the results. Descriptions of smaller reception studies undertaken by scholars (Cabeza-Cáceres, Maćzyńska and Szarkowska, Iglesias Fernández) will follow.

The Audetel project in the UK led to the creation of the most comprehensive audio description standards to date, i.e. the ITC (Independent Television Commission) Guidance On Standards for Audio Description (2000). It was implemented in 1992-1995 by an international EU-funded consortium of regulators, consumer associations and broadcasters set up in 1991. The project aimed at collecting feedback from the visually impaired respondents regarding AD of various TV genres (feature films, musicals, soap operas, nature documentaries, etc.) and at developing a digital technology to provide AD soundtrack to every household.

Four methods were employed in the project to elicit feedback. A questionnaire about viewing habits and difficulties in following programme content was distributed by the Royal National Institute for the Blind among the visually impaired in the UK. Later, 200 respondents “of all ages and levels of visual impairment from around Britain” (ITC Guidance 2000: 3) participated in viewing sessions and expressed their opinions about examples of presented audio described programmes. Additionally, a focus group was established for in-depth critiques of the audio described content and a trial TV service was launched for a few months in 1994 for peak-time ITV and BBC programmes. 100 special set-top receivers were distributed among respondents who could watch 7-10 hours of audio described content per week and were periodically interviewed to elicit feedback about any aspect of the service.

Based on the feedback from this wide-ranging and exhaustive study, the guidelines address a plethora of issues in AD, from the most crucial principles illustrated by examples (usage of tenses, prioritising information, use of adjectives and adverbs, descriptions of logos, opening titles, cast lists and credits, etc.) to descriptions of various programme categories (including even sport and live events, sexually explicit or violent programmes and advertisements) (ITC Guidance 2000).

Another important large-scale reception study project conducted in the UK was “Bollywood for all” instigated by the Royal National Institute of Blind People (RNIB) in 2009. The study aimed at identifying the existence of potential demand for audio described Bollywood films in the UK and India and the best medium to offer accessibility to such productions and at understanding “the best ways of reaching out to this audience to create awareness about the availability of this service” (Rai 2009: 3). Being “the world’s most prolific film industry” with approx. 700 new productions per year, Bollywood was chosen so that “every single person, including people of Asian origin with sight loss, get[s] that chance to sit back and enjoy a Bollywood film of their choice, independently” (Rai 2009: 3). The report from the project entitled “Bollywood for all: the demand for audio described Bollywood films” was to serve as a reference tool for the industry and AD providers by identifying key findings, conclusions and eight major recommendations pertaining to the study objectives. As mentioned above, since this paper focuses on methodology we will neither describe the results nor evaluate the project outcomes.

The project was conducted by a research team including AD activists and sociologists. It included a quantitative research study with an impressive number of 260 visually impaired respondents of Asian descent living in the UK. The sample of respondents was representative in terms of age, sight condition and exposure to Bollywood films (Rai 2009: 6). It also encompassed a qualitative research study in the form of in-depth interviews with 25 people with sight problems.

The quantitative study involved both face to face and telephone interviews conducted by 23 interviewers. The high number of interviewers and interviewees

highlights the wide scale of the project, which can rarely be achieved by scholars working individually or in small teams. In fact, one of the major hurdles in reception research including blind and partially sighted participants is getting access to them. The Bollywood project reached a high number of respondents by contacting 50 organisations (including charities, statutory organisations, eye hospitals and local authorities), of which 20 were willing to assist by identifying scheduled gatherings and meetings with the visually impaired the interviewers could attend. In fact, many respondents were identified through various “approaches to generate contacts including the use of social networks of friends and relatives as well as visiting local shops, schools, colleges, hospitals and housing estates” (sic!) (Rai 2009: 29). It is quite telling that only about 33% of interviews took place in community centres while the majority was conducted at the homes of the respondents. All this shows that reaching study participants is a difficult, time- and effort-consuming activity that requires flexibility.

The interviews in the quantitative study took approx. 10 minutes each. There were 14 questions pertaining to the respondent’s age, vision dysfunction, TV/DVD/cinema viewing habits, preferences regarding Bollywood films and factors making watching such films difficult, exposure to AD and sources of information about products designed for the blind and partially sighted.

The questions about audio description were preceded by a short introduction of the topic, i.e. the interviewer defined AD and only later asked about the respondent’s experience. All questions were multiple choice questions. The ones about vision dysfunction included descriptions equivalent to vision impairments (e.g. “In a room during daytime, I can tell by the light where the windows are” or “I can see the shapes of furniture in a room”) rather than specific medical names of various dysfunctions (congenital blindness or Age-related Macular Degeneration) to make them more accessible to all participants. Surprisingly, the interviewers did not ask when the respondents’ impairments originated (congenital or acquired dysfunctions), which is a key factor (alongside blindness vs. partial sight difference) contributing to the heterogeneous character of the AD target group and specific preferences.

The in-depth interviews conducted in the qualitative study lasted for approx. 35 minutes. The respondents were informed about confidentiality and anonymity of their responses. The questionnaire included 23 questions, many similar to those in the quantitative survey. The difference was that the respondents of this study were additionally shown three 3 minute clips from a Bollywood film – without AD, with English AD and with Hindi AD. They were then asked detailed questions about the clip (location, actions of characters, whether they found it enjoyable, whether AD improved their experience and what additional items they understood with the second audio described film). This means that each respondent saw each clip three times and the methodological weakness here is that the answers obtained after the third viewing may not have only been influenced by the Hindi AD only but also by yet another exposure to the

experimental clip (the issue of cross-balancing experimental materials and participants will be discussed further in section 4).

In general, however, the Bollywood project is impressive in its range and provides excellent contribution both to the reception research on AD and to the practice of audio description.

A small-scale reception study in AD was conducted by Cabeza-Cáceres (2011) to examine the influence of narration speed, explicitation and intonation on the comprehension of AD. This study, together with the qualitative reception study of the Audetel project constitutes another type of reception experiments. Respondents are not only simply asked about their preferences, they are also presented with examples of audio descriptions with manipulated variables (lack of AD, language of AD in Audetel; speed, explicitation and intonation in the study by Cabeza-Cáceres). The dependent variable is thus comprehension measured by a series of questions about the content of presented video clips.

Cabeza-Cáceres managed to interview 30 visually impaired respondents recruited via ONCE (Spanish Association of the Blind) and ACIC (Catalan Association for the Integration of the Blind). They were divided into groups of 10 and each group was presented with AD with a different level of the independent variable (for instance in the case of intonation: monotonous, neutral and dramatised). The study is methodologically well designed but the groups are actually quite small which may obscure the statistical analysis of the results.

It is interesting to see how Cabeza-Cáceres coped with the problem of measuring comprehension, which is yet another methodological difficulty in the studies of such type. He analysed his video material on the basis of the narrative schema comprehension model by Branigan (1992) including such stages of narration as orientation, initiating event, goal, complicating action, resolution and epilogue (Cabeza-Cáceres 2011). He then identified the following items in his clip: events (10 items), changes in the two characters (6 items) and emotional context (2 items). He decided on the relevance of these items for comprehension and assigned weights (30%/50%/20% respectively), thus arriving at a measurable index of 100%. This is an elegant method to quantify comprehension in order to plot it as a dependent variable against other data (respondents' profiles and independent variables).

According to Iglesias Fernández (2010: 216), "reception studies in AD should not stop at users' preconceived preferences but further expose them to actual AD products for assessment." She strongly supports reception studies that would, as in the case of the previously discussed studies, include samples of AD to make it easier for respondents to create their own opinions and express them directly. The visually impaired people's quality expectations regarding audio description may differ from the actual assessment of AD products and hence the usefulness of confronting these expectations with actual samples in the reception studies.

Iglesias Fernández et al. (2011) were interested in the role of paralinguistic aural stimuli in the perception and assessment of AD. They posited that more

congruent stimuli (the audio describer's voice congruent with the visual stimuli) may contribute to increased comprehension and aesthetic experience. They recruited 12 visually impaired participants affiliated to ONCE (from Madrid and Granada) to complete a set of questionnaires regarding their expectations, voice sonority assessment and AD quality quality assessment. The participants were divided into two equal groups and presented with two experimental clips with AD. The audio describer's voice was the independent variable with two levels: either congruent or incongruent with the "emotional landscape and the character's mental state" (Iglesias Fernández et al. 2011). The stimuli were counter-balanced across the participants, i.e. the first group of participants watched the first clip with the congruent voice to identify emotions and the second clip with the incongruent voice to assess quality while the second group of participants watched the first clip with the incongruent voice and the second clip with the congruent voice to perform the same tasks respectively.

The voice ratings were done by choosing appropriate sonority and emotional correlates (for instance, lax vs. tense for the former correlates and courageous vs. fearful for the latter) or by using a 5 point Likert scale. The respondents were also asked directly if the "describer's voice contribute[d] to a better and more congruent understanding of the emotional atmosphere and the character's mental state and intentions" (Iglesias Fernández et al. 2011) and which version of AD they preferred.

The sample size in this study is a serious weakness barring any statistical analysis. However, the authors offer an interesting design and ways to elicit feedback about the respondents' preferences in a structured and quantifiable way.

Yet another small-scale reception study was conducted by Mączyńska and Szarkowska to find out if the visually impaired audience would accept text-to-speech audio description (i.e. AD read out by a speech synthesizer rather than pre-recorded or delivered live by a human) to a documentary. The authors overcame the usual problem of finding participants by organising an open screening in a café and by conducting an online survey. They managed to reach 54 visually impaired respondents. In an online survey the respondents first watched a 30 minute audio described film ("La Soufriere" by Werner Herzog) and then completed an online questionnaire with 15 questions, including the usual questions about the respondent's age, visual impairment, exposure to AD and some more specific questions about the use of speech synthesis software, AD preference, the selection of voice and the acceptance of text-to-speech AD. There were three open-ended questions inviting the respondents to provide comments about text-to-speech AD and synthetic voices. The survey was a typical reception study aiming at eliciting preferences. The application of an online survey turned out to be a good solution and helped the authors reach a wider group of respondents.

The AD-Verba research project is part of a research grant awarded by the Polish Ministry of Science and Higher Education and its aim is to develop Polish AD standards that would reflect the preferences of the visually impaired audience. Our goal is to interview approximately 100-120 VIPs in order to learn about their viewing habits and AD preferences. The end of the project is scheduled for the end of 2011. Prior to conducting the reception study discussed in this article, we also conducted a pilot study, the results of which we used to design the study concerned (we discuss some aspects of the pilot study in the Methodological considerations and hurdles section below).

### RESEARCH DESIGN

The reception study discussed here was in the form of a face-to-face questionnaire-based interview lasting approximately 30 minutes. Each interview was scheduled individually with a given participant. The final version of the questionnaire (see the Appendix; also see the Methodological considerations and hurdles section below) consisted of three parts: a pre-questionnaire, the questionnaire proper and a post-questionnaire. The pre-questionnaire comprised 13 questions and was aimed at eliciting general information about the respondents, such as age, gender, education, type of visual impairment, prior experience with AD, general viewing habits and preferences (i.e. how much time they devote to watching films or television and what their preferred programmes and mass media are, how they watch films or programmes, etc.), as well as the sources of information on facilities for the blind they consult most often (for details see the Appendix; we also discuss the question on the types of visual impairment in greater detail in the Methodological considerations and hurdles section).

In the questionnaire proper we showed the respondents three clips with AD. All of the clips are from Polish films and represent various genres: the first clip is from a drama/film noir *Rewers* by Borys Lankosz (2009), the second one is from a romantic comedy *Tylko mnie kochaj* by Ryszard Zatorski and Adam Iwiński (2006), and the third one is from a comedy *Testosteron* by Andrzej Saramonowicz and Tomasz Konecki (2007). The clips last 2.19, 1.28 and 2.35 minutes, respectively. For each clip we recorded two alternative audio descriptions (versions A and B), so there were six clips altogether. The alternative AD differed in the solutions we were testing. Where necessary for comprehending a given scene, we presented a context for this scene. We also asked the participants whether they had seen a given film. Each clip was then followed by 11 comprehension and/or preference questions (for examples see The reception study proper – tested items section below). The clips were counterbalanced across participants, i.e. each participant was presented with either clip A or B, rather than with both clips, in order to

rule out the influence of the first viewing on the responses obtained for the second viewing (cf. research design of Bollywood for all study discussed in the AD reception studies – state of the art section above; see also Methodological considerations and hurdles below).

The final part of the questionnaire was a post-questionnaire in which we asked the participants questions about their general AD preferences, such as the gender of the voice talent reading AD, objective vs. subjective descriptions (e.g. ‘an attractive singer’ vs. ‘a long-legged singer in a miniskirt’), and AD comprehension of description of gestures (‘Konstanty bids farewell to Stefan with a hand gesture’ vs. ‘Konstanty raises his fist and spreads his fingers in a farewell gesture’). We also elicited their opinion (using a Likert-type scale) on the quality of AD delivery (e.g. overlaps with dialogues, speed, pauses) and on various descriptive elements (e.g. the use of colours, evaluative adjectives, similes, reading out credits or the name of the audio describer by the voice talent). We also asked them what they thought about AD usefulness and potential (for details of the post-questionnaire see the Appendix).

#### THE RECEPTION STUDY PROPER – TESTED ITEMS

In this section we discuss some of the items we tested in the reception study proper. In the first clip (*Rewers*) Sabina, the main female protagonist, is being mugged by two thugs and is saved by a handsome passer-by, Bronisław. He then walks her home, they talk a bit and decide to meet for tea some time. We then see Sabina in the lift smiling softly, her eyes dreamy.

Facial expressions pose a great challenge to audio describers and there is a heated debate underway whether such expressions should be described objectively (e.g. “She raises her eyebrows”) or whether some degree of interpretation should be allowed for ease of comprehension or because of time constraints (e.g. “She looks surprised”) (see e.g. Dosch and Benecke 2004: 24; Mazur and Chmiel in press; Orero and Vercauteren, manuscript in preparation). We therefore wanted to test whether our respondents would have better comprehension of the scene in the case of a more objective or rather a more subjective, interpretative description. For the scene in question we provided two alternative ADs: ‘Her pupils are dilated. She squints her eyes’ (objective) vs. ‘She has dreamy eyes’ (interpretative). We then asked a comprehension question: “What was Sabina’s facial expression after meeting Bronisław, while in the lift?” Along similar lines, we tested the preference of our respondents for other objective vs. interpretative descriptions, such as: ‘Her eyes are wide open. She shifts her weight from foot to foot’ vs. ‘She looks frightened’ or ‘He blows out the smoke. His hands in the pockets. His legs wide apart’ vs. ‘Self confident and nonchalant, he blows out the smoke’ (for more examples see the Appendix).

In this clip we also tested the naming of characters. The prevailing AD standards are not unanimous when it comes to when characters should be named. For example, according to UK guidelines characters can be named the first time they appear on screen, unless their identity is to be kept secret (ITC Guidance 2000: 16), whereas in the German tradition it is preferred to name the characters when their names first appear in the film. Until that time the characters are identified by their characteristic feature(s), for example ‘a man with a moustache’ (Dosch and Benecke 2004: 23). As regards the latter solution, problematic may be cases where characters are named very late in the film or when there are a lot of relevant characters and the scenes change quickly (cf. Benecke 2011). In such cases more elaborate descriptions could be too time consuming and/or too taxing for the audience, and thus naming characters immediately would be preferred. In the scene at hand, the identity of the characters does not need to be kept secret, but the two main protagonists introduce themselves early on in the film. In one of the two alternative ADs for the clip we named the characters the first time they appeared on screen, and in the other the characters were named when they first introduced themselves to each other. Having shown either of the two versions we then asked our respondents which solution they preferred (naming the characters immediately vs. naming them when they are named in the film). In the case where the latter option was chosen, we asked a follow-up question: “What if the character’s name is introduced as late as in the middle of the film?”

The second clip shown by us (*Tylko mnie kochaj*) featured a little girl, Michalina, and her alleged father, Michał, who takes her out on a shopping spree. We see Michalina trying on different outfits, dancing and presenting herself to Michał, who sits in an armchair and non-verbally expresses his opinion about the clothes. The scene is very much reminiscent of the famous scene in *Pretty Woman* where Julia Roberts presents herself in different outfits to Richard Gere who either approves or disapproves of them. We therefore wanted to see whether such an intertextual reference in the AD of the scene would facilitate its visualisation/comprehension (at least for those who lost their sight later on in life and have seen the film). So one of our two descriptions read as follows: ‘Michalina, just like Julia Roberts in *Pretty Woman*, presents to Michał her outfits: a coat and a beret, two summer dresses’. We then asked the participants whether this reference helped them visualise the scene.

Another item tested in this clip was the amount of detail in AD: one of the descriptions was very precise, describing all of Michalina’s outfits in great detail, for example: ‘Michalina in a beret, a red dress and a coat presents herself to Michał. (...) Michalina in a blue cotton dress (...) Michalina presents a white dress with the pattern of green apples, red cherries and strawberries, yellow lemons’. The other description was much more general: ‘Michalina (...) presents to Michał her outfits: a beret and a coat, two summer dresses’. We then asked the respondents to choose their preferred option. We expect here some gender-

related differences, with women preferring more elaborate descriptions, while men going for leaner ones.

The third and final clip in our study (*Testosteron*) featured a group scene with seven men of a similar age, and thus similar-sounding voices (except for one, the father of one of the characters, who is older). There is a lot of dialogue in the scene, the characters speak fast and their lines are quite short. One of the solutions in such a case would be to introduce each line with the name of the character saying it. We were interested whether our respondents would welcome such an option. We thus drafted and recorded two ADs for the clip: in one the dialogue in the group scene was uninterrupted, in the other each line was introduced with a character's name. We then asked the respondents about their preference.

Another item that we tested was explicitation. In the clip there is a scene where the main protagonist – in a state of shock – is giving a lecture to sheep in a pen. He addresses them as if they were humans, starting with “Ladies and Gentlemen”. Although the bleating of the sheep is audible, we thought that VIPs could find the scene confusing. So in one of the versions of AD we explicitly said: ‘Kornel in the pen addressing the sheep’, whereas in the other we gave a regular, general description: ‘Kornel in the pen’. Having presented either of the two versions of AD we then tested comprehension of the scene by asking: “Who did Kornel address his lecture in the pen to?”

With this clip we also tested comprehension of a scene audio described objectively and in a more interpretative manner. The final scene of the clip features one of the characters, Tretyn, leaning over Kornel, who is hallucinating and sees Tretyn as a werewolf. The two ADs proposed by us read: ‘Tretyn’s face turns bluish, his irises shine. He bares his teeth’ and ‘Tretyn’s face appears to Kornel as a werewolf’s muzzle’. We then asked how Kornel perceives Tretyn leaning over him.

## METHODOLOGICAL CONSIDERATIONS AND HURDLES

Before conducting the reception study discussed in this article, we carried out a pilot study in which we interviewed 18 participants. In principle, the questionnaire-based study was in the form of face-to-face interviews, which took place directly after two screenings of audio described films (see below). We also printed the questionnaires in large print for low vision respondents who were willing to fill them in by themselves, however later it turned out that some questions or even whole pages were left unanswered, so in our subsequent study we abandoned this method of data collection in favour of face-to-face interviews (for detailed methodology and results of the pilot study see Chmiel and Mazur 2011).

In the pilot study, in addition to the standard questions about the age, gender, education and type of visual impairment, we asked the respondents about AD usefulness, the major obstacles in AD reception and their preferences as regards

the description of places or characters. Another question concerned subjective interpretation in AD. We thought that a question on subjective interpretation in AD should be included in the survey, especially given the heated subjectivity vs. objectivity debate among AD practitioners and researchers and the prevailing opinion that VIPs generally mind subjectivity in AD (for an overview of the debate see Mazur and Chmiel in press). We therefore asked the respondents whether they considered the following descriptions to be subjective interpretation. Although in this article about methodology we refrain from presenting the results, we will make an exception in this case to justify subsequent changes in the questionnaire. The table below presents the percentages obtained for this question.

Description	<i>Is this subjective interpretation?</i>	
	Yes	No
he watches him with concern	31%	69%
elegantly dresses	23%	77%
they exchange surprised looks	31%	69%
attractive singer	46%	54%
sexy brunettes	62%	38%
worried Stefan	38%	62%
taps him on the shoulder good-naturedly	38%	62%

Table 1. Results concerning subjective interpretation in AD (pilot study).

As one can see, with the exception of ‘sexy brunettes’, all of the other descriptions were not considered subjective interpretation by the majority of the respondents. However, what we observed while conducting the interviews was that some of the respondents, not “tainted” with the on-going subjectivity vs. objectivity debate, were quite perplexed by this question and not sure what to answer. We thus thought that in our reception study proper it would be better to elicit responses concerning objectivity or subjectivity indirectly (implicitly) rather than directly (explicitly) as was the case in the pilot study. We then reworked this question so that instead a direct ‘yes’ or ‘no’ question, we asked the respondents to choose one of two possible descriptions (e.g. ‘an attractive singer’ vs. ‘a long-legged singer in a miniskirt’) that they found more appealing. Along similar lines, in the pilot study there was a question concerning clarity in AD (“Is this description clear to you?”), where we had descriptions such as ‘a run-down house’. Again, we replaced this direct question with a preference one and the respondents were asked to choose between e.g. ‘a run-down house’ and ‘a house with plaster peeling off the walls’ (with a note that the latter is only one of the many features that make the house look run-down, but there is no time to describe them all).

Another important observation that we made during the pilot study was that while preference questions are by all means necessary, what is also needed in a

reception study (and what can perhaps tell us more about the effectiveness of a given AD solution) are comprehension questions. Given this consideration, in the reception study at hand we decided to include AD samples (clips) with some proposed AD solutions in two versions (often concerning debatable issues, such as naming of characters, explicitness vs. implicitness in AD, subjectivity vs. objectivity) and test which AD is more comprehensible by the viewers (for details see the previous section where we discussed the clips and the tested items; on methodological problems with comprehension questions see below).

Despite the fact that when drafting the questionnaire for the reception study at hand we were heavily drawing on our experience gained during the pilot study, we revised the questionnaire six times. Each time, having interviewed a batch of new respondents, sometimes following their feedback, we saw that some things needed changing, while other needed to be made more specific. We also saw the need for new questions, while we thought that other could be done away with. We would then revise the questionnaire accordingly. We are aware that the final version may be far from perfect, nonetheless it has been given a lot of thought and consideration on our part and includes on-going feedback from our respondents. Below we present some of the methodological aspects of the questionnaire as well as methodological hurdles to overcome.

In the pre-questionnaire we decided to make the question concerning the type of visual impairment more specific. In the original version of this part of the survey we had the following options for this question: a) a blind person; b) a low vision person, c) other; d) comments. It was followed by a question about the age when the impairment occurred in order to establish whether the participants had any residual visual memory. It soon turned out that such information may not be enough to interpret the obtained results, for example as regards the usefulness of AD, the use of colours or detail-orientation of description, since some of the participants reported that with their visual impairment they could notice certain things on the screen. We thus decided to make this question more specific, by asking the respondents to mark statements – which were largely based on the ones in the Bollywood for All survey (Rai 2009) – that best describe their impairment, such as “I can read a newspaper headline” or “I can see the shapes of furniture in a room” (note that questions 9 and 10 in the pre-questionnaire were also modelled on the Bollywood for all survey; see the Appendix). Also in the pre-questionnaire, we decided to move questions concerning age and education to the end, so that the participants would feel more at ease and ‘warmed-up’ when answering them.

When it comes to the reception study proper, the major change that we introduced was to show either of the two versions of AD of a clip to a respondent. Originally, we showed both versions (i.e. altogether six clips to one person). This turned out to be counterproductive, for example, for comprehension questions, because when answering a question after the second version of AD was shown, the participants had already been influenced by the contents of the first version,

so comparing answers in such a case is problematic to say the least (see the ‘werewolf’ example discussed in the section on The reception study proper – tested items). Now we simply compare answers given by respondents who saw version A with those provided by respondents who were shown version B, in which case a real comparison is possible. Irrespective of this new design, it must be noted however that it is always problematic to measure the respondent’s comprehension after watching a film/clip because comprehension is difficult to quantify and because there is an important confounding variable that is difficult to rule out, i.e. memory. When exposing study participants to video clips with or without AD the experimenter is never sure to what extent the responses to clip content questions are influenced by the participants’ memory span and to what extent they are aided by AD. Also, some respondents may find the very situation of an interview stressful and thus lack proper concentration. This is a major methodological limitation that has to be taken into consideration when interpreting the study results.

As for the post-questionnaire, we moved here two questions from the pre-questionnaire, namely “Do you find AD useful?” and “Do you prefer AD to be read by a male voice, a female voice or it doesn’t matter?”<sup>1</sup> This was because some of the respondents have had very little experience with AD before or have not seen films with AD at all. Asking the first of the two questions in the post-questionnaire made more sense, as such persons have at least seen the three clips and could make some comments on AD usefulness. The same goes for the other question – once they have heard AD read out by a male and female voice, they could have an opinion as to which option they prefer.

We also introduced a number of modifications to the preference questions in the post-questionnaire. For example, we added questions concerning the use of filmic language in AD (such as ‘the camera freezes’), short vs. long sentences in description as well as ones pertaining to a more technical side of AD (such as acknowledging the audio describer, the voice talent or audio describing the logo of the film studio). In the final version of the post-questionnaire we used a seven-level Likert-type scale (with 7 meaning “I strongly agree” and 1 meaning “I strongly disagree”) to elicit most of the opinions of the respondents about AD in general. The Likert scale had replaced “Yes”, “No”, “I don’t know” answers used by us in the previous versions of the post-questionnaire, as we thought the seven-level scale would be more precise.

In addition to questionnaire-related challenges, we also faced some other, more general methodological hurdles. Everyone who has ever conducted a reception study with VIPs knows that it is quite difficult to reach such persons. Some researchers contact associations or foundations for the blind (cf. studies discussed in section on AD reception studies – state of the art), which is a very good option, as such organisations can help get in touch with quite a few respondents. We used this option for our pilot study when we contacted the Association of Friends of the Blind and Visually Impaired in Pozna . The association helped us

organise an AD screening, after which we held the interviews. Similarly, we conducted some interviews after a screening of two documentaries with AD during a film festival devoted to the topic of disability. Another good way of reaching prospective respondents is to join an event organised with VIPs in mind. In this way, one can have access to a large number of VIPs in one place and conduct the interviews in an efficient manner. For example, for the purpose of our study we met with blind participants of a dance festival in Poznań as well as of an orientation meeting organised by our university for disabled students. We are also considering conducting the survey online (cf. Mączyńska and Szarkowska 2011), though we fear that, given the substantial length of the questionnaire, the respondents will get discouraged and leave some questions unanswered (cf. large print questionnaires discussed above). A middle-of-the-road solution could be to make the questionnaire and clips available online to the VIPs and then have an interviewer conduct the interview over the phone or Skype.

Another difficulty in reaching appropriate respondents lies in the fact that in order for their feedback to be of value they should have adequate educational or even mental development level (vision dysfunctions are sometimes associated with other disabilities, including mental ones). Often it is not possible to verify such things before the interview, which can cost a lot of time and effort on the part of the researcher, as despite the time spent the data obtained are not valid. What is more, it is often the case that in a country such as Poland where AD is still quite a novelty, some of the respondents know very little about this form of audiovisual transfer and do not know yet what to expect from it (Barbara Szymańska – personal communication). Nonetheless, we believe that an opinion obtained from a blind or low vision person can be valid and valuable, despite the fact that such a person knows little about AD, as such feedback gives us better insight into how VIPs ‘see’ the world and visualise and comprehend what they hear.

## CONCLUSIONS

If we treat visually impaired persons as customers of audio description, and quality as what the customer wants and expects, then it seems that the best way to provide quality AD is to develop standards or guidelines that would reflect the preferences of the customer. In our opinion the best way to learn about the preferences of target recipients of AD is through reception studies.

In this paper we reported on two such studies – a pilot one involving 18 participants, and an on-going research project AD-Verba aimed at development of Polish AD standards that would reflect the preferences of Polish VIPs. In the article we focused on methodological issues, without discussing the results, which will be the subject matter of a separate article. Both studies were principally in the form of questionnaire-based face-to-face interviews, and when drawing up the AD-Verba questionnaire we heavily drew on our experiences gained while

conducting the pilot study. Nonetheless, we revised the questionnaire several times in order to reflect suggestions for improvement made by the VIPs during the interviews. Other changes were the result of our observations and reflections while interviewing VIPs.

After conducting the pilot study we noted that in addition to having preference questions it would be good to test the respondents' comprehension of descriptions, as this could tell us more about the effectiveness of certain AD solutions. In AD-Verba we thus included clips followed by comprehension questions. We also learned that it was more efficient to cross balance the clips and questions across the interviewees to make their comprehension more comparable. However, a methodological difficulty that we encountered here was that some of the answers could have been influenced by the respondents' lack of concentration or stress or, as the case may be, by a confounding variable that is memory. So sometimes it was difficult to discern whether we were in fact testing comprehension or the subject's memory span. After the pilot study we also learned that it is more effective to elicit responses concerning objectivity or subjectivity of descriptions indirectly (implicitly) rather than directly (explicitly).

When conducting reception studies it is usually a challenge to access a sufficient number of VIPs that could provide valuable feedback. We believe that it is useful to contact organisations and associations for the blind as well as organisers of events for VIPs, where larger groups of respondents can be interviewed. Also, our experience shows that it is ineffective to provide low vision persons with questionnaires in large print, as some questions were left unanswered. For this reason, we think that although having the questionnaire and clips accessible online by the VIPs is an option to consider, we are quite sceptical that a fairly long questionnaire would be filled in completely without the assistance of an interviewer over the phone or Skype.

All in all, despite some criticism of the usefulness of reception studies, we believe that even in countries such as Poland where AD is still at the crawling stage and VIPs may not know what to expect from it and what it can give them, conducting such studies can give us a lot of valuable insight into how blind and low vision persons understand and visualise different descriptions and thus perceive the world. Such knowledge could then be used to create quality AD based on guidelines that reflect what the customer wants.

## NOTES

\* Research discussed in this article was funded by the Polish Ministry of Science and Higher Education as part of a research grant no. N N104 178236.

1 Please note that originally all of our ADs were recorded by a female voice, but given this question we decided to use a male voice as well. So in the final version of the study AD to one of the clips is recorded by a male voice.

## REFERENCES

- Benecke, B. 2011. "Creating Audio Description: The Use of Character Fixation and Set Fixation." Conference paper presented at 4<sup>th</sup> Media for All Conference, London.
- Branning, E. 1992. *Narrative Comprehension and Film*. London & New York: Routledge.
- Cabeza-Cáceres, C. 2011. "Intonation in AD: Does It Affect Users' Comprehension?" Conference paper presented at 4<sup>th</sup> Media For All Conference, London.
- Chmiel, A. and Mazur, I. 2011. "Audiodeskrypcja jako intersemiotyczne przekład audiowizualny – percepcja produktu i ocena jakości" In *Przekład jako produkt i kontekst jego odbioru*, Kasperska, I. and Żuchelkowska, A. (eds.), 13-30. Poznań: Wydawnictwo Rys.
- Dosch, E. and Benecke, B. 2004. *Wenn aus Bildern Worte Werden. Durch Audio-Description zum Hörfilm*. München: Bayerischer Rundfunk.
- Fresno, N. 2011. "Going Cognitive. The Role of Memory in Audio Description." Conference paper presented at 4<sup>th</sup> Media For All Conference, London.
- Iglesias Fernández, E. 2010. "La dimensión paralingüística de la audiodescripción: un acercamiento multidisciplinar." In *Un corpus de cine: Fundamentos teóricos y aplicados de la audiodescripción*. Jiménez Hurtado, C., Seibel C. and Rodríguez A. (eds.), Granada: Tragacanto.
- Iglesias Fernández, E., Martínez Martínez, S. and Chica Núñez, A. 2011. "Reception Studies in Audio Description and Interpreting Quality Assessment: A Case of Cross-Fertilization." Conference paper presented at 4<sup>th</sup> Media For All Conference, London.
- International Organization for Standardization. [www.iso.org](http://www.iso.org), 12.09.2011.
- "ITC Guidance on Standards for Audio Description." 2000. [http://www.ofcom.org.uk/static/archive/itc/uploads/ITC\\_Guidance\\_On\\_Standards\\_for\\_Audio\\_Description.doc](http://www.ofcom.org.uk/static/archive/itc/uploads/ITC_Guidance_On_Standards_for_Audio_Description.doc), 3.9.2011.
- Mazur, I. and Chmiel, A. (in press). "Audio Description Mad to Measure: Reflections on Interpretation in AD Based on the Pear Tree Project data." In *Media for All 3*. Carroll M., Orero P. and Remael A. (eds.). Amsterdam: Rodopi.
- Mączyńska, M. and Szarkowska, A. 2011. "Text-to-Speech Audio Description with Audio Subtitling to a Non-Fiction Film *La Soufriere* by Werner Herzog." Conference paper presented at 4<sup>th</sup> Media For All Conference, London.

Orero, P. and Vercauteren, G.  
(manuscript in preparation).  
“The Importance of Adequate  
Descriptions of Facial Behaviour:  
Cognition beyond Description.”

Rai, S. 2009. “Bollywood for All.”  
[http://www.rnib.org.uk/aboutus/  
Research/reports/accessibility/  
Documents/2009\\_09\\_  
Bollywood\\_AD\\_report.pdf](http://www.rnib.org.uk/aboutus/Research/reports/accessibility/Documents/2009_09_Bollywood_AD_report.pdf),  
14.12.2009.

Van der Heijden, M. 2007. “Film en  
televisie toegankelijk voor blinden  
en slechtzienden. Een onderzoek  
naar de functionaliteit van geluid  
bij audiodescriptie.” Unpublished  
MA paper. Hogeschool voor de  
Kunsten Utrecht, Faculteit Kunst,  
Media & Technologie.

APPENDIX<sup>1</sup>

I. PRE-QUESTIONNAIRE

<b>1. Gender</b>	F <input type="checkbox"/>	M <input type="checkbox"/>
<b>2. Type of visual impairment:</b>		
a) blind	b) partially sighted	
comments:		
<b>3. When did you lose vision? / Since when have you had vision impairment? (in the case of partially sighted respondents)</b>		
a) I have been blind since birth / I've had vision impairment since birth	b) since I was _____ years old	
<b>4. Which statement best describes what you are able to see (when wearing glasses or contact lenses if you use them). You can select more than one answer. I can see well enough to:</b>		
a) read a newspaper	b) read a newspaper in large print	
c) read a newspaper headline	d) recognise a friend across the road	
e) recognise a friend across a room	f) recognise a friend who is at arms length away	
g) recognise a friend if I get close to his or her face	h) see the shapes of the furniture in a room	
i) In a room during daytime, I can tell by the light where the windows are	j) I cannot see anything at all	
<b>4. Have you heard of audio description?</b>		
a) Yes, I have but I don't know what it is	b) Yes, I have and I know what it is	
c) Yes, I have and I have used it	d) I've never heard of it.	
<b>5. How much audio described material have you seen? (if answer c to question 4)</b>		
a) less than 5 hours	b) 5-10 hours	
c) 10-20 hours	d) over 20 hours	
<b>6. How much time do you spend daily on watching TV/films?</b>		
a) less than an hour	b) 1-3 hours	
c) more than 3 hours	d) none	
<b>7. What films and series do you watch most often?</b>		
a) Polish films and series	b) foreign films and series with Polish dubbing	

<sup>1</sup> Please note that this is one of the two versions of the questionnaire that we were using in our reception study. The other version includes some minor changes particularly as regards comprehension questions for clips with different AD.

c) foreign films and series with Polish voiceover	d) foreign films and series with Polish subtitles read by speech synthesizer
<b>8. Where do you watch programmes or films most often?</b>	
a) on TV	b) on DVD
c) on the Internet	d) in cinema
<b>9. When watching TV:</b>	
a) I do not find that I have any difficulty following what is going on the screen	b) I have difficulty seeing the picture on the TV screen
c) I have difficulty seeing details on the TV screen	d) I have difficulty seeing text on the TV screen
e) I am able to see the light of the TV screen	f) I cannot see anything
<b>10. How do you watch programmes and films? (choose all the options that apply to you)</b>	
a) I wear special stronger glasses	b) I get closer to the TV screen
c) I use a magnifier	d) I ask someone to assist me by explaining what happens on the screen
e) I just try to pick up as much as I can from the sound of the film or programme	f) other
<b>11. What are your sources of information for products specially designed for the visually impaired?</b>	
a) specialist magazines for people with disabilities	b) organisations for people with disabilities, e.g. the Polish Association of the Blind
c) the Internet	d) press, radio, TV
e) family and friends	e) other
<b>12. Age</b>	
a) 18-24	b) 25-34
c) 35-44	d) 45-54
e) 55-64	f) 65-74
g) 75+	
<b>13. Education:</b>	
a) primary	b) vocational
c) secondary	d) student
e) university degree	

## II. QUESTIONNAIRE PROPER

You will watch three audio described videos. After watching them I will ask you some questions about their content or audio description.

*Rewers [The reverse]*

Have you seen the film? YES / NO

[REWERS A]

What was Sabina's facial expression after meeting Bronisław, while in the lift?

--

1. Should the characters be named immediately (*Sabina, Bronisław*) or only when they are named in the film?

Immediately		Later	
-------------	--	-------	--

2. What if the character's name is introduced as late as in the middle of the film?

Immediately		In the middle of the film	
-------------	--	---------------------------	--

3. Which solutions do you prefer?

<i>Her pupils are dilated. She squints her eyes</i>		<i>She has dreamy eyes</i>	
<i>Her eyes are wide open. She shifts her weight from foot to foot</i>		<i>She looks frightened</i>	
<i>He blows out the smoke. His hands in the pockets. His legs wide apart</i>		<i>Self confident and nonchalant, he blows out the smoke</i>	
<i>in a beret, loose-fitting jacket, too long skirt and flat shoes</i>		<i>a modestly dressed grey mouse</i>	
<i>a young dark-haired man in a suit and a trenchcoat</i>		<i>handsome like a heart-throb</i>	

*Tylko mnie kochaj [Just love me]*

Have you seen the film? YES / NO

CONTEXT: The film is set in Warsaw. The main character is Michał, a businessman and a young bachelor. One day, seven-year old Michalina knocks on his door and claims to be his daughter. Michał thinks it's a joke but the little girl doesn't give up and moves in for five days to convince him that he can love her.

**[ZAKUPY B]**

1. Which solutions do you prefer?

<i>Michalina in a beret, a red dress and a coat presents herself to Michał. She turns from side to side</i>	<i>Michalina (...) presents to Michał her outfits: a beret and a coat, two summer dresses. She dances and fools around</i>	
<i>Michalina in a blue cotton dress. She lowers her shoulder straps. Michał nods. Michalina presents a white dress with the pattern of green apples, red cherries and strawberries, yellow lemons. Michalina dances. She pirouettes</i>	<i>Michał evaluates her outfits</i>	

2. Has the reference to *Pretty Woman* helped you imagine the described scene? **YES / NO**

*Testosteron*

Have you seen this film? **YES / NO**

CONTEXT: Alicja, a bride standing before the altar tells the groom, Kornel, that she's in love with another man. She kisses one of the guests (Tretyn) in the church and runs away. Stavros, the father of the groom, Robal, Fistach, Tretyn, and groom Kornel (weakened after a fight in the church) and his brother Janis arrive at the inn at the lake where the wedding party was to take place. Waiter Tytus joins the six men.

**[TESTOSTERON A]**

Who did Kornel address his lecture in the pen to?

How does Kornel perceive Tretyn leaning over him?

1. Do you prefer when the names of characters are announced in a group scene?

Yes		No	
-----	--	----	--

2. Which solutions do you prefer?

<i>Tretyn's face appears to Kornel as werewolf's muzzle</i>	<i>Tretyn's face turns bluish, his irises shine. He bares his teeth</i>	
<i>Kornel in the pen</i>	<i>Kornel in the pen addressing the sheep</i>	

### III. POST-QUESTIONNAIRE

<b>1. I prefer when audio description is read by:</b>							
a) a man				b) a woman			
c) depends on the film/programme				d) it doesn't matter to me			
<b>2. Which solution do you prefer?</b>							
attractive singer				long-legged singer in a miniskirt			
Stefan watches the cat that continues walking on the display cases knocking off plastic ice-cream spoons				Stefan watches the cat. The cat continues walking on the display cases. It knocks off plastic ice-cream spoons			
a house with plaster peeling off the walls				a run-down house [i.e. peeling plaster, destroyed roof, derelict windows and doors, but no time for such a description]			
autumn sun				weak sunlight			
he winks raising his left lip corner				he winks meaningfully			
Robert at the railway station. He approaches the escalator. The camera freezes showing the top stairs. Red heels, black tights. The camera moves up to show a woman's face				Robert at the railway station. He approaches the escalator. He looks at the top stairs. A woman in black clothes and red high heels goes down			
<b>3. Please, show the described gesture.</b>							
Konstanty bids farewell to Stefan with a hand gesture.				Konstanty raises his fist and spreads his fingers in a farewell gesture			
<b>4. To what extent do you agree with the statements below? (7 – strongly agree; 1 – strongly disagree)</b>							
1) Audio description helps in film reception	1	2	3	4	5	6	7
2) Aspects that are irrelevant to the plot should be described	1	2	3	4	5	6	7
<b>3) It bothers me in audio description when:</b>							
a) the text is read too fast	1	2	3	4	5	6	7
b) the text is read too slowly and includes unnatural pauses	1	2	3	4	5	6	7
c) the text fills all pauses between dialogues	1	2	3	4	5	6	7
d) the text overlaps with dialogues	1	2	3	4	5	6	7
e) the text is not synchronised with the picture (is read before or after what it describes)	1	2	3	4	5	6	7
<b>4) Audio description should include</b>							
a) evaluative adjectives (e.g. beautiful, ugly)	1	2	3	4	5	6	7

b) colours	1	2	3	4	5	6	7
c) similes, e.g.							
- a building as tall as ten elephants put one on top of another	1	2	3	4	5	6	7
- moves his lips like a fish gasping for air	1	2	3	4	5	6	7
- bog with a colour and consistency of a tomato sauce	1	2	3	4	5	6	7
d) logo of the film studio	1	2	3	4	5	6	7
e) all opening titles, time permitting	1	2	3	4	5	6	7
f) all credits, time permitting	1	2	3	4	5	6	7
g) the audio describer's name	1	2	3	4	5	6	7
h) the voice talent's name	1	2	3	4	5	6	7
<b>5) I would always use audio description services if it was available on TV</b>	1	2	3	4	5	6	7

# Text-to-speech audio description of voiced-over films. A case study of audio described *Volver* in Polish

AGNIESZKA SZARKOWSKA  
University of Warsaw, Poland

ANNA JANKOWSKA  
Jagiellonian University, Krakow, Poland

## ABSTRACT

*Given that the production (esp. recording) of AD is quite costly, there are not very many audio described films available on the Polish market. Moreover, there is practically no audio description to foreign films in Poland since it has been assumed that blind and partially sighted audiences will not manage to assimilate multiple soundtracks (original soundtrack in foreign language, voiceover and audio description). In order to overcome the cost hurdle, we propose text-to-speech audio description (TTS AD) as a cheaper alternative to traditionally produced AD. We will demonstrate how TTS AD can be combined with voice-over to produce AD to foreign films on the example of *Volver* by Pedro Almodovar. We will also present the results of a survey conducted among a group of blind and partially sighted audience after a screening of voiced-over *Volver* with TTS AD. The results of the survey demonstrate that the participants are quite open to the idea of TTS AD both as an interim solution – until there are more audio described films available – and as a permanent solution.*

## INTRODUCTION

Back in the 1990s, the map of audiovisual Europe was divided into dubbing, subtitling and voice-over countries (Gottlieb 1998). With subtitles in the cinemas and television voice-over for both fiction and non-fiction audiovisual productions, Poland seems to be an exception in Europe's audiovisual landscape.

Television voice-over is at the same time an advantage and a disadvantage for the blind and partially sighted audience. On the one hand, unlike subtitles, voice-over provides them with the translation of the foreign dialogue. Unfortunately, at the same time it has been assumed that due to multiple soundtracks (the original soundtrack in a foreign language, the voice-over and the audio description), it is pointless – if not impossible – to combine voiced-over foreign films with audio description (AD). Therefore, there is practically no audio description to foreign films in Poland. This stands in stark contrast with the number of foreign language films screened in cinemas, broadcast on TV or released on DVD/Blue-ray. Last but not least, as declared on many occasions, blind and partially sighted people, just like a sighted audience, want to watch foreign films.

In order to audio describe a foreign film, one needs to combine the AD script with either audio subtitling or voice-over. Since an overwhelming majority of films in Poland are voiced-over, in our paper we will present how this AVT modality can be combined with audio description. Given that the production (esp. recording) of AD is quite costly, there are not very many audio described films available on the Polish market. In order to overcome the cost hurdle, we propose text-to-speech audio description (TTS AD) as a cheaper alternative to traditionally produced AD. We will demonstrate how TTS AD can be combined with voice-over to produce AD for foreign films on the example of *Volver* by Pedro Almodóvar. Finally, we will also present the results of a survey conducted among a group of blind and partially sighted audience after a screening of voiced-over *Volver* with TTS AD.

## POLISH AUDIOVISUAL LANDSCAPE

Until recently, Europe has traditionally been divided into dubbing, subtitling and voice-over countries (Gottlieb 1998). However, the recent findings of the Media Consulting Group (2007) show that this classical division is a simplification of a far more complex situation. We no longer can, if we ever could, talk about strictly dubbing, subtitling or voice-over countries since most of them employ all three methods in various contexts.

With dubbing, subtitling and voice-over existing side by side, Poland seems to be an excellent example of this complexity. The most popular AVT modality in Polish cinemas is subtitling. For young audience, however, dubbing is preferred. Nevertheless, movies considered to be suitable for audiences of any age (e.g.

*Alice in Wonderland*, *Harry Potter*, *Shrek*) are often made available both dubbed and subtitled. Voice-over is without any doubt the predominant modality in television broadcasts. It is used for both fiction and non-fiction audiovisual products, the only exception being animated and non-animated productions for children and some animated movies for the general audience (e.g. *Shrek*). The choice, however, is not always obvious – the BBC television series *The Chronicles of Narnia* was broadcast with voice-over, whereas the motion picture *The Chronicles of Narnia* by Walt Disney Pictures was dubbed. DVDs and Blu-ray usually contain both voice-over and subtitles, or dubbing and subtitles.

Such an audiovisual landscape influences many areas – one of them is accessibility for the blind and partially sighted. While cinema with subtitles is virtually inaccessible to audiences with visual impairments, one could argue that television voice-over has the advantage of at least providing them with the translation of the foreign dialogue. In the long run, however, this seems to be a drawback rather than an advantage for the blind and partially sighted. Currently there is no audio description on Polish television mostly because the digital turn is yet to come.<sup>2</sup> Unfortunately, the chances of making television fully accessible for the visually impaired are scarce since it has been assumed that it is impossible to provide audio description for voiced-over programs. In a country with a high import of foreign television programmes, where most – if not all – of them are voiced-over, this virtually means that television will not be accessible, or it will be accessible in a very limited way, via audio described programs produced in Polish.

#### AUDIO DESCRIPTION IN POLAND – DEVELOPMENT AND CHALLENGES

Though to many it may come as a surprise, audio description was launched in Poland already in the late 1990s. As described by Jankowska (2008: 242), the project was initiated by Andrzej Woch, a blind employee of the Jagiellonian University in Krakow, and funded by the Central Library of the Polish Association of the Blind (*Polski Związek Niewidomych*, PZN). The so-called *typhlo-films* involved providing additional commentary for twenty films on VHS tapes, making them available to the blind and partially sighted in the Central Library. As opposed to today's audio description, freezing the image whenever the additional commentary did not fit in the gap between the dialogues was a common practice in *typhlo-films*. It is highly probable that due to this feature the idea of describing films for the blind and partially sighted did not manage to force its way to broad social consciousness until some years later. Nevertheless, it should be noted that within the *typhlo-film* project additional commentary for the blind was provided also for foreign voiced-over films. This, until recently, was the only attempt at combining audio description with voiced-over films.

The first public audio described cinema screening, which took place on 27 November 2006 in Białystok, can be regarded either as a revival or as the true

beginning of audio description in Poland. The audio description for this first screening was read live by an audio describer who was sitting with the audience and reading the script with a microphone in his hand, so that AD could be heard by all the spectators. Ever since, the only more or less regular screenings of audio described films take place in Warsaw as a part of the project *Cinema beyond silence and darkness* carried out by the *Foundation for Children Help on Time (Fundacja Dzieciom Zdażyć z Pomocą)*.<sup>1</sup> Irregular screenings take place in other larger cities, but they are rather one-off events very often organized by enthusiasts and unfortunately almost as often by amateurs.

Until very recently audio described programs have been available only online. Polish public television (TVP) audio described some of its programs (about 74 hours) and made them available online on its website. The audio described programs include two feature films, five TV series and one series for children. In order to access audio described films, one needs to receive a special password which can be obtained from the Polish Association of the Blind – free of charge for its members or for a fee for non-members. As to DVDs, at the moment of writing this article there are only 12 discs with AD available on the Polish market.

This situation will hopefully improve since on 1 July 2011 a new law obliging all broadcasters to provide AD and SDH came to life. Unfortunately at this point it is really hard to predict when and if the broadcasters will provide accessibility services. The law stipulates that 5% of quarterly broadcast (commercials and tv-shopping excluded) in 2011 and starting from 2012 about 10% of quarterly broadcast (10% relates to AD, SDH and sign language interpreting taken together)<sup>2</sup>. It also allows the broadcasters to apply to the National Broadcasting Council for individual permission to lower the percentage of accessibility services. In the case of not providing the services, the law provides for a fine up to 10% of yearly income. From what can be observed now, broadcasters are becoming interested in providing accessibility services. Although because of financial reasons they are more willing to provide SDH than AD due to their programmes. At the time of writing this article, the only broadcaster to provide AD was TVP – however it should be mentioned that they screened one of the series that has been made available on-line since at least 2008.

It should be stressed that – apart from typhlo-films – all the above mentioned films and TV series were either originally filmed in Polish or were dubbed into Polish. As we have already explained, it has been assumed that due to multiple soundtracks (the original soundtrack in a foreign language, the voice-over/audio subtitling and the audio description), audio description cannot be combined with foreign films screened with voice-over or audio subtitles. As a result, for the time being there are no audio described foreign language films, which is especially striking when compared with the amount of foreign language productions present on TV or released on DVD/Blu-ray. Above all, it also collides with the preferences and capabilities of the blind and partially sighted people who, as declared on many occasions, want to watch foreign films just like the sighted audience.

This is also confirmed by our findings from a pilot study conducted prior to launching the TTS AD project. In February 2010 we asked a group of 17 pupils, aged 12 to 18, from the Special Educational Centre for Blind and Partially Sighted Children in Krakow, about their television viewing preferences as well as about their opinion on combining audio description with voiced-over/audio subtitled foreign programs. When asked about the audio described films and series they would like to watch, 53% preferred foreign productions, 29% opted for Polish and 18% did not prefer either one of the options. The views on combining audio description with voice-over or audio subtitling were especially interesting, as 87% of the interviewed pupils declared that multiple soundtracks would not be an obstacle to film enjoyment.

Unfortunately, reluctance towards providing audio description to foreign films is not the only challenge that needs to be faced in Poland. After the initial enthusiasm, the lack of financing inhibits the implementation of audio description.

#### PREVIOUS STUDIES ON THE USE OF SYNTHETIC SPEECH BY THE VISUALLY IMPAIRED

The invention of synthetic speech has been an important milestone in the everyday lives of many visually impaired people, who can now benefit from a host of text-to-speech (TTS) applications both in their work and for leisure activities. There are numerous applications harnessing the power of text-to-speech systems for people with visual impairments: from GPS-based mobility aids, screen reading software for web browsing, email, etc., educational tools, such as TTS dictionaries and textbooks, to entertainment, for instance audio subtitles in audiovisual materials (see Freitas and Kouroupetroglou 2008 in Cryer and Home 2008: 5).

Synthetic speech enables visually impaired people to access information without relying on other people reading it to them or waiting for it to be brailled (Garcia 2004; Llisterri, Fernández, Gudayol, Poyatos and Martí 1993), thus allowing for more independence. This is particularly important in the context of receiving financial information (Thompson, Reeves and Masters 1999).

It is thanks to a relatively low cost of synthetic speech that the number of materials made accessible to the blind and partially sighted people is on the rise. A good example is the RNIB service known as *Talking Books*.<sup>4</sup> RNIB research on user attitudes towards synthetic speech in *Talking Books* reveals that while most users prefer a human narrator for leisure reading, they felt synthetic voice “would be acceptable for reference, instructional and non-fiction books” (Cryer and Home 2009: 5). It has also been found that the attitude towards synthetic speech varied greatly and largely depending on previous experience with TTS applications. Many users stressed that – as opposed to a human narrator – synthetic speech allows them to choose their own accent-free voices with

neutral emphasis; this, in turn, makes it possible for users to “add their own interpretation” (Cryer and Home 2009: 6).

Synthetic speech has also been used to read daily newspapers to blind and partially sighted people, as reported in a study by Hjelmquist, Jansson and Torrell (1990). The four-month study revealed that “whilst initially many users had doubts about the system, and found synthesised speech difficult to understand, all reported ‘getting used to it’ after a few hours of listening” (cited in Cryer and Home 2008: 6). In fact, the more experience users have with synthetic speech, the better their comprehension performance is (Rhyne 1982; Venkatagiri 1994).

Apart from previous experience with synthetic speech, other factors that may influence comprehension of text-to-speech messages include the presentation rate (Hjelmquist, Dahlstrand and Hedelin 1992; Koul 2003), the voice intelligibility (Papadopoulos et al. 2009) and the presence of background noise (Koul and Allen 1993). Last but not least, it is the quality of synthetic speech – especially its intelligibility and naturalness – that plays a crucial role both in comprehension and attitude towards text-to-speech systems. It needs to be noted that some of the studies cited here were conducted several years ago and that the quality of synthetic speech has improved considerably since then.

#### TEXT-TO-SPEECH AUDIO DESCRIPTION

Over the past few years audio description became fashionable. Sadly, despite the growing enthusiasm and demand for audio description, the number of audiovisual products accessible to the visually impaired is hardly sufficient.

Text-to-speech audio description (TTS AD) was proposed as a possible option to increase the output of audio described programs through cost reduction. As noted by Szarkowska (2011: 145) “TTS AD offers unequalled cost-effectiveness in terms of AD production in comparison with conventional methods of producing audio description as it does not require the recording of the AD script (for pre-recorded AD) nor does it incur any human labour costs for the reading out of the AD script (for live AD)”. In fact, the human labour costs mentioned above are reduced to the cost of creating and adapting the audio description script to the demand of TTS AD. Apart from all the pitfalls of traditional AD script writing, the TTS AD process is not excessively complicated (see Szarkowska 2011 for details).

In order to create text-to-speech audio description, first an AD script is written and then, using subtitling software, it is synchronised with the film. Put simply, synchronizing the AD script with a film means preparing a list of AD chunks in the form of subtitles, each consisting of text and time codes. From the technical point of view, the process differs from preparing the traditional subtitles only by the fact that the text is to appear in the gaps between the dialogues and not simultaneously with them. Later on, the text file is read by speech synthesis software while the film is played on a multimedia player, or the AD script can be recorded and mixed with the multimedia file.<sup>5</sup>

## THE TTS AD PROJECT

The scarcity of audio description concerns not only the amount but also the range of audiovisual products. On the one hand, due to the costly production process, there are simply not enough audio described products. On the other hand, the reluctance towards audio describing films with voice-over/audio subtitling results in a lack of foreign films with audio description.

The goal of the TTS AD project is to research the possibility of increasing both the number and the range of audio described films through investigating the feasibility of combining TS AD with audiovisual programs broadcast with dubbing, subtitling and voice-over as well as programs filmed originally in Polish.

The TTS AD project is being developed at two Polish universities – the University of Warsaw and the Jagiellonian University of Krakow. The five-stage study includes the implementation of TTS AD in:

- 1) a Polish feature film (see Szarkowska 2011),
- 2) a foreign feature film with Polish dubbing (unpublished MA thesis and PhD dissertation)
- 3) a foreign feature film in English combined with voice-over in Polish (discussed in this paper)
- 4) a foreign feature film in English combined with audio subtitles in Polish (work in progress),
- 5) a documentary in English combined with voice-over in Polish (work in progress).

It is our contention that the ultimate purpose of audio description is to serve the blind and the partially sighted community, therefore reception studies are the core of our research. In order to find out the opinion of the blind and partially sighted viewers, we conduct a survey after every screening. One of the most important research issues is whether visually impaired viewers would accept TTS AD either as a permanent or interim solution. Having completed three out of the five stages, we can say that so far TTS AD has proved to be an acceptable solution for the visually impaired community, as 94% of our viewers accept TTS AD as an interim solution and 63% are willing to accept it as permanent solution.

## RATIONALE

The film selected for the purposes of this research was *Volver*, a 2006 Spanish drama directed by Pedro Almodóvar. This is how Peter Bradshaw described it in his *Guardian* review:

*Volver* (in English, *Coming Home* or *Coming Back*) is a gripping melodrama inspired by the trash TV that is a soundtrack to its characters' lives. Penélope Cruz is Raimunda, a hard-working woman with a teenage daughter, Paula, and a feckless, layabout husband. [...] Raimunda's family life shatters with one terrible act of violence, and there

is a secret about her late mother Irene that surfaces when Irene returns from beyond the grave to make contact with her astonished daughters.

The film rendered itself quite well for audio description. It nevertheless presented the audio describer with a number of problematic issues, such as the vivid imagery of picture composition, insufficient time to convey important information, culture specificity of certain elements presented on screen (food, places or signs in Spanish) and the intensity of emotions expressed in the characters' faces, especially in close-ups. Given that several scenes featured at least two female characters, whose original voice was covered by voice-over, the issue of speaker identification was particularly pertinent. In order to solve this problem, the names of the characters were incorporated in the AD script whenever it was deemed necessary.

As the film was already available on DVD with Polish voice-over translation, it was decided that this version would be used in the study, complemented with the AD script read by text-to-speech software. For the project, the speech synthesiser Ivona (Ivo Software) was used together with the synthetic voice Krzysztof (Loquendo).

#### AUDIO SUBTITLING VS. VOICE-OVER

*Volver* was one of the few foreign (i.e. non-English speaking) films audio described in the UK and released on DVD, so the AD script had to be accompanied by a translation of dialogues, which was done through audio subtitles read out by a female narrator. The choice of the female voice most probably stemmed from the nature of the film, where women play the most important characters. The AD script, in contrast, was read by a male voice talent. This solution enabled the audience not to confuse the AD script with the dialogues. However, the presence of one female voice for all the (mainly female) characters and the poor quality of the recording of audio subtitles, which drowned out the original Spanish voices so that they were hardly audible, resulted in viewers having difficulties recognizing which character was speaking as many scenes in the film feature a few women talking. As a result, the overall quality of the AD was perceived as poor and the audio described film met with fierce criticism from the British visually impaired community (Leen Petre from RNIB – personal communication).

It is worth noting at this point that the British audience is not used to hearing a translation of a film being read out to them on top of the original voices. Polish viewers, in contrast, have had many years of experience of listening to the voice-over translation of film dialogues on television, which makes them more accustomed to this AVT modality.

Audio subtitling and voice-over seem to be two audiovisual translation modalities which have a lot in common. First of all, they both consist of a translation of the dialogue list of a foreign or multilingual film. Secondly, the

translation is read out to the target audience – the main difference being that in the case of voice-over, the target audience is simply conceived of as mainstream sighted population, whereas in audio subtitling it comprises a much smaller group of visually impaired people. Thirdly, the translation is usually read out by one voice talent (typically male in Poland), while the voices of the original actors can still be heard in the background though their volume has been turned down. Polish voice-over is done in a professional recording studio, which usually guarantees good sound quality. Finally, apart from the different target audiences envisaged at the production stage, audio subtitling is created together with the AD script and thus allows for some flexibility in combining the two tracks, whereas in the case of Poland, AD would be added to a voiced-over film at a later stage, which makes it virtually impossible to introduce changes to the pre-recorded voice-over so that it can be seamlessly interwoven with the AD script.

## RESEARCH QUESTIONS

The key objective of the present study was to determine whether visually impaired viewers would find it acceptable for text-to-speech software to read AD scripts to voiced-over feature films. To address this objective, the following three research questions were formulated:

- 1) Which AD voice would the visually impaired prefer if they had a choice between a human voice and a synthetic voice?
- 2) Would TTS AD be acceptable as an interim solution, until a system has been agreed to have a human voice reading out the AD?
- 3) Would TTS AD be acceptable as a permanent solution, next to AD read by a human voice?

## PROCEDURE

The screening took place at an informal meeting for blind and partially sighted people organised by the Foundation Chance for the Blind (*Fundacja Szansa dla Niewidomych*) in Jachranka, near Warsaw, on 24 April 2010.

The audience was first invited to watch the film and after the projection they were asked to answer 13 questions, which were read out by sighted volunteers. The first part of the questionnaire was meant to establish participants' age, education, degree and type of sight loss (congenital/acquired). The second part aimed to find out their views on the use of speech synthesis in AD, their previous experience with audio description and their familiarity with text-to-speech software.

## SAMPLE

After the screening, a total of 20 people were interviewed: 14 women (70%) and 6 men (30%). As shown in Table 1, five of them were blind (25%), 13 were partially sighted (65%) and two of them (10%) were sighted.

Table 1. Participants by age and degree of sight loss

Age bracket	Blind	Partially sighted	Sighted
18-25	-	3	-
26-39	2	7	1
40-59	3	1	-
60-74	-	2	1
Total	5	13	2

Most participants (12 people, 67%) had a congenital sight loss, while one in third (8 people, 33%) acquired their sight loss at a later stage in life. Both the degree and type of sight loss was determined based on self-declarations of the participants.

12 out of 20 participants (67%) said they use text-to-speech software regularly, either at home or at work. Only 11 people (55%) had seen some films with audio description before, while nine of them had no prior experience of AD.

## RESULTS

When asked about what voice they would prefer to read AD scripts, half of the participants (10 people, 50%) declared their preference for a human voice. Perhaps somewhat surprisingly, one person preferred a synthetic voice to read AD, whereas many others stated that this depends on the type of programme (6 people, 30%). Three participants (15%) were not sure and would have preferred to have more experience with AD to make a more informed choice.

In terms of accepting TTS AD as either an interim or permanent solution, most participants were in favour of both (see Table 2). Some expressed their concerns whether the introduction of TTS AD would not result in eliminating human voices and substituting them completely with synthetic voices.

Table 2. The acceptance of TTS AD as an interim or permanent solution

	Interim	Permanent
Yes	95%	70%
No	-	15%
Don't know	5%	15%

n=20

Overall, all participants apart from one were in favour of introducing TTS AD as an interim solution, especially if it meant more audio described programmes accessible to people with visual impairments. The participants were slightly more sceptical, however, about the introduction of TTS AD as a permanent solution: while 70% of them support the idea, one in three is either against or unsure.

A closer examination of the preferences for TTS AD as an interim or permanent solution based on the degree of sight loss has shown a slight tendency on the part of blind participants to be more supportive of the idea (see Table 3).

Table 3. The acceptance of TTS AD as an interim or permanent solution by degree of sight loss

	TTS AD as interim			TTS AD as permanent		
	yes	no	don't know	Yes	No	don't know
Blind	100%	-	-	80%	-	20%
Partially sighted	92%	-	8%	70%	15%	15%

n=20

This may be due to the fact that partially sighted viewers can to some extent rely on their eyesight to watch films, while blind viewers are more dependent on AD and thus want more audio described films regardless of the voice reading the script.

In terms of gender, it is female participants who appear to be more inclined to accept TTS AD than men (Table 4).

Table 4. The acceptance of TTS AD as an interim or permanent solution by gender

Gender	TTS AD as interim			TTS AD as permanent		
	yes	no	don't know	Yes	no	don't know
Women	100%	-	-	79%	14%	7%
Men	83%	-	17%	50%	16%	34%

n=20

There seem to be no clearly discernible correlations between the preference for TTS AD as an interim or permanent solution in terms of participants' age (Table 5).

Table 5. The acceptance of TTS AD as an interim or permanent solution by age

Age	TTS AD as interim			TTS AD as permanent		
	yes	no	don't know	Yes	no	don't know
18-25	100%	-	-	67%	33%	-
26-39	90%	-	10%	60%	20%	20%
40-59	100%	-	-	75%	-	25%
60-74	100%	-	-	100%	-	-

n=20

Interestingly, participants from elder age groups seem to be slightly more willing to see TTS AD as a permanent solution than those from younger age groups in the study. Naturally, the sample is too small to draw any further reaching conclusions.

As stated above, previous studies on synthetic speech revealed that the experience and exposure to text-to-speech software may positively influence the attitude towards it. This pattern seems to be confirmed in our study (Table 6).

Table 6. The acceptance of TTS AD as an interim or permanent solution by the use of TTS software

	TTS AD as interim			TTS AD as permanent		
	yes	no	don't know	Yes	no	don't know
TTS users	100%	-	-	75%	8%	17%
Not TTS users	86%	-	14%	57%	29%	14%

n=20

TTS users are more likely to accept TTS AD both as an interim and as a permanent solution when compared to those who have had no regular experience with speech synthesis software. This pattern is more noticeable with regard to TTS AD as a permanent solution.

A similar trend can be observed when it comes to the preference for either human or text-to-speech narrator (Table 7).

Table 7. The preference for human/synthetic audio describers by the use of TTS software

	TTS users	Not TTS users
Human	33%	64%
Synthetic	-	12%
Depends on the programme	42%	12%
Don't know	25%	12%
Total	12 people	8 people

n=20

In the study, people who do not habitually use text-to-speech software were more likely to prefer human narrators, while regular TTS users were more open to the idea that TTS AD may be a good solution for some types of programmes, but not for all (42% stated the choice of the human/synthetic voice depended on the programme). This issue will be pursued in further stages of our research when we investigate the application of TTS AD in non-fiction genres, such as a documentary and an educational programme.

## CONCLUSION

Although we acknowledge that further research involving a wider range of films and more varied viewers is needed all in all, the present study not only reveals a considerable potential of text-to-speech audio description but also the acceptance of visually impaired viewers of audio description for voiced-over films. Together with the results of other parts of the project, it seems clear that TTS AD is acceptable for the majority of the visually impaired viewers interviewed with a surprisingly high number of people also accepting it as a permanent solution.

## ACKNOWLEDGEMENTS

This work has been supported by the research grant ‘Text-to-speech audio description’ no. N N104 148038 of the Polish Ministry of Science and Higher Education for the years 2010-2011 and research grant ‘Audio description with text-to-speech software’ from the Faculty of Philology of the Jagiellonian University in Krakow.

Many thanks to Marzena Chrobak, the Polish translator of the *Volver* screenplay published by the ZNAK publishing house, for letting us use several fragments of her translation in the AD script. We also thank Joanna Pietrulewicz and Agnieszka Walczak for their help in administering the questionnaire after the screening. Last but not least, we owe our warmest thanks to Marek Kalbarczyk for inviting us to Jachranka and allowing us to carry out the study.

## NOTES

1 Special screenings used to take place twice a month, on selected days. So far 21 films, with both audio description and subtitles for the deaf and hard of hearing have been presented, however the project was suspended for some months due to lack of financing.

2 Things in Poland have just changed and AD has been introduced on Polish television right after writing this article.

3 10% includes audio description, subtitling for the hearing impaired and sign language interpreting taken together.

4 In 2009, the RNIB Talking Books library had over 16,000 titles. According to Cryer and Home (2009: 8), "it takes an average of five days to record a

Talking Book, and around 100 new Talking Books are added to the library each month". It is believed that synthetic speech "may significantly reduce the time required to produce new books, which may mean more books could be produced" (*ibidem*).

5 Initially it was impossible to change the reading speed and the volume, which were set at the same level throughout the film. Because of that at times audio description was not audible or it did not convey as much information as needed. Now those problems are solved thanks to some advanced features of the text-to-speech software which allows the control and the adjustment of the volume and the reading speed to the on-screen action.

## REFERENCES

- Cryer, H. and Home, S. 2008. "Exploring the Use of Synthetic Speech by Blind and Partially Sighted People." RNIB Centre for Accessible Information, Birmingham: Literature review #2.
- Cryer, H. and Home, S. 2009. "User Attitudes towards Synthetic Speech for Talking Books". RNIB Centre for Accessible Information, Birmingham: Research report #7.
- Bradshaw, P. 2006. "Volver." *The Guardian*, 25 August 2006. Available at: <http://www.guardian.co.uk/film/2006/aug/25/drama.pedroalmodovar>
- Cryer, H. and Home, S. 2009. "User Attitudes towards Synthetic Speech for Talking Books." RNIB Centre for Accessible Information, Birmingham: Research report #7.
- Freitas, D. and Kouroupetroglou, G. 2008. "Speech Technologies for Blind and Low Vision Persons." *Technology and Disability* 20, 135-156.
- Garcia, L.G. 2004. "Assessment of Text Reading Comprehension by Spanish-Speaking Blind Persons." *British Journal of Visual Impairment* 22 (1): 4-12.
- Gottlieb, H. 1998. "Subtitling". In *Routledge Encyclopedia of Translation Studies*, Baker M. (ed.), 244-248. London and New York: Routledge.
- Hjelmquist, E., Jansson, B. and Torell, G. 1990. "Computer-Oriented Technology for Blind Readers." *Journal of Visual Impairment and Blindness* 17: 210-215.
- Jankowska, A. 2008. "Audiodeskrypcja, szczytny cel w tłumaczeniu." *Między Oryginałem z Przekładem* (14): 225-246.
- Media Consulting Group. 2007. "Study on Dubbing and Subtitling Needs and Practices in the European Audiovisual Industry. Final Report", Media Consulting Group, Paris / Peacefulfish, London. [http://ec.europa.eu/information\\_society/media/overview/evaluation/studies/index\\_en.htm](http://ec.europa.eu/information_society/media/overview/evaluation/studies/index_en.htm)
- Papadopoulos, K. et al. 2009. "Perception of Synthetic and Natural Speech by Adults with

Visual Impairments." *Journal of Visual Impairment and Blindness* 103 (7): 403-414.

Rhyne, J.M. 1982. "Comprehension of Synthetic Speech by Blind Children." *Journal of Visual Impairment and Blindness* 10 (10): 313-316.

Szarkowska, A. 2011 "Text-to-Speech Audio Description. Towards wider availability of AD." *Journal of Specialised Translation*, 142-162.

Thompson, L., Reeves, C. and Masters, K. 1999. "In the Balance: Making Financial Information Accessible." *British Journal of Visual Impairment* 17 (2): 65-70.

Venkatagiri, H. S. 1994. "Effect of Sentence Length and Exposure on the Intelligibility of Synthesised Speech." *Augmentative and Alternative Communication* 10: 96-104.

## APPENDIX 1

### The questionnaire

1. Age
  - 18-25
  - 26-39
  - 40-59
  - 60-74
  - 75+
2. Gender
  - Female
  - Male
3. Are you
  - Blind
  - Partially sighted
  - Sighted
4. Type of sight loss
  - Congenital
  - Acquired
5. Do you use text-to-speech software regularly at home or at work?
  - Yes
  - No
  - Don't know
6. Have you seen any films with audio description?
  - Yes
  - No
  - Don't know
7. If you had a choice, which AD voice would you prefer?
  - A human voice
  - A synthetic voice
  - Depends on the film/programme
  - Don't know / doesn't matter
8. Would you accept TTS AD as an interim solution, until a system has been agreed to have a human voice reading out the AD?
  - Yes
  - No
  - Don't know
9. Would you accept TTS AD as a permanent solution, as an alternative to a human voice?
  - Yes
  - No
  - Don't know

10. Did you like the fact that the AD script was read by a male voice – just as the *lektor's* voice?
- Yes, I liked it
  - No, I'd prefer to have a female voice
  - Don't know

Why?

11. Was it easy to decide who was speaking?
- Yes, it was quite easy
  - No, it was quite difficult
  - Sometimes it was difficult
  - Don't know

Why?

12. Did you like the fact that the AD script included fragments of Almodovar's screenplay and that AD was not only a description of what was going on on the screen, but it also included extra information, such as the description of the main characters at the beginning of the film?
- Yes
  - No
  - Don't know

Why?

13. What did you like and what did you not like in the AD you have just seen? Do you have any comments?

# Audio description and audio subtitling in a dubbing country: Case studies

BERND BENECKE

Bavarian Broadcasting, Munich, Germany

## ABSTRACT

*In many European countries foreign films are not dubbed but subtitled. An audio describer has to include all the written subtitles in his script and try to make the description fit in between. Dubbing countries like Spain, Italy and Germany are also used to combining audio description and audio subtitling – for different reasons. This presentation shows how audio subtitling affects the work of describers in a dubbing country like Germany. It will present examples from daily work to show how many different ways are used to deal with the subtitles.*

## INTRODUCTION

A new focus in the research on audio description is the interaction with the field of audio subtitling. In many European countries, foreign films (mainly with English dialogues) are not dubbed, but rather subtitled. In such a case, the work of the audio describer becomes more complicated, for he has to include all the written subtitles in his script and try to make the description fit in between. In the production process, sometimes more than one narrator is needed to make a distinction between what is subtitle and what is description. From time to time,

the describer has to introduce the name of a character being subtitled, in order to make clear who is speaking. However, audio subtitling is not a common practice only in subtitled countries: dubbing countries like Spain, Italy and Germany are also used to combining audio description and audio subtitling. Even when the (Spanish, Italian or German) dubbed version of the film is described, subtitles may appear. By analyzing four specific cases, this article will demonstrate where, why and how describers can deal with it.

#### CASE 1: A PROTAGONIST OF A FILM SPEAKS IN SIGN LANGUAGE

This is usually a rather easy problem to handle, because there is (exceptions are, of course, possible) no sound in the film when the subtitled sign language appears. To make what is subtitle and what is description clear, you may choose a second voice for the subtitles – normally of the same sex as the subtitled character. With more than one person talking together in sign language, a third or fourth speaker may be necessary – if the budget makes it possible. Otherwise one voice has to do more than one character. The following example from *The piano* (J. Champion, 1993, New Zealand) shows how this works. This is the English translation of the German audio description and subtitling:

##### DESCRIPTION:

The woman enters the tent and talks in sign language:

##### SUBTITLING:

And the wind said: Remember how we played earlier. So the wind took her by the hands and said: Come with me. But she refused.

##### DESCRIPTION:

The girl too talks in sign language.

##### SUBTITLING:

Mom, I have thought about it.

##### SPOKEN:

I don't call him Dad. I don't call him anything.

##### DESCRIPTION:

The woman touches her daughter's cheek.

In this example, the problem is that the girl first uses sign language and then – from one second to another – switches to spoken language. As the subtitled sign language is spoken by another person and not by the girl's voice in the film, blind and visually impaired people may get confused and not realize that both sentences are uttered by just one character.

## CASE 2: A PROTAGONIST IN A DOCUMENTARY SPEAKS IN A FOREIGN LANGUAGE

A useful example is provided by a scene from *Am Limit* (*To the limit*, P. Danquart, 2007, Germany), a German documentary about two mountain climbers trying to break the speed record in climbing up *El Capitan* in Yosemite National Park, USA. They meet two other climbers and talk to them in English, which is subtitled. In the described version all subtitles had to be spoken by one voice (different from the voice doing the audio description). To make this understandable, in some cases the describing voice had to announce who was actually speaking:

DESCRIPTION:

Two elderly mountain climbers appear. Thomas:

SUBTITLING:

American Climbers of the old fashioned style

DESCRIPTION:

Alexander:

SUBTITLING:

Your hands look like hard work.

DESCRIPTION:

One of the old guys:

SUBTITLING:

This man went up with me. He is disabled. He called me and said: I wanna go down the El Capitan. I said: But we can't go down on a wheelchair.

SPOKEN:

(All four are laughing)

SUBTITLING:

Then he went down with me! Four times.

In the editing and mixing of that scene, the audio description team tried to connect parts of the original English dialogue with the German audio subtitling, for some people may be able to understand the English and therefore recognize that the translation really corresponds to what is said in the original – something that is often done in documentaries with the voice-over. I define these connections between the original soundtrack and the audio subtitling *audio connectors*. Examples of these audio connectors in the case described are 'Disabled' (Original) – *Behindert* (Subtitles) or 'Four times' (Original) – *Vier Mal* (Subtitles).

## CASE 3: A SONG IN A DUBBED FILM IS KEPT IN THE ORIGINAL LANGUAGE

This kind of audio subtitling appears very often in dubbed films. Only in children's movies – for example Walt Disney movies – are songs newly recorded

in the dubbing language. But an intro-song of a *James Bond* movie or a song in a musical are usually kept in the original language, which sometimes – if the lyrics are important for the understanding – lead to subtitles which have to be spoken in the audio described version. But as the song itself is important, the blind and visually impaired audience wants to hear as much as possible of it and the audio subtitling has to be edited very precisely into the gaps between the song lyrics or over repeated lines.

Let's take as example two films by Pedro Almodóvar: *Volver* (2006, Spain) and *Hable con ella* (*Talk to her*, 2002, Spain). Both films include a very important song for the story that in the German dubbed version is kept in Spanish and therefore audio subtitled. In *Volver* the female describing voice also does the audio subtitling, which is obviously not the best way to do it. Although she tries to switch between a more neutral describing voice and an emotional subtitling voice, it is sometimes hard for the audience to follow – especially as the song is not only covered with audio subtitling but also with descriptions of what is going on while the woman is singing her song. In the editing and mixing, an audio connector is established with *Volver/Zurückkehren*, which also refers to the title of the film.

In *Habla con ella*, the audio subtitling is presented by a second female voice, whereas a man does the describing. This is much easier to understand, blind and visually impaired people can easily distinguish the description from the audio subtitling. We find audio connectors in the words *cantar/singen* and *Paloma/Taube*.

#### CASE 4: THE USE OF A SECOND LANGUAGE IS A MAJOR TOPIC OF THE FILM

Examples of this growing number of films are *Inglorious Basterds* (Q. Tarantino, 2009, USA) or *Babel* (A. González Iñárritu, 2006, USA). Even if the English in *Inglorious Basterds* is dubbed, the French, German and Italian parts have to be subtitled, otherwise the whole story would not make sense. *Babel*, with its blend of English, Spanish and Japanese is another example of this kind: the title-giving Babylonian language-mix has to be kept and only one language (usually the English) can be dubbed.

A further example is the German-British film *Desert Flower* (S. Horman, 2009), where flashbacks are kept in the original Somali language, that the main protagonist used to speak as a child. Sometimes it gets hard for the audio subtitler, because the four people who talk to each other in Somali need to be audio subtitled. The only solution to this problem is the use of two additional voices, a female voice for the women speaking and a male voice for the men. If it becomes too difficult to identify the character speaking, e.g. when two women or two men are speaking to each other, the describing voice has to specify who is actually speaking. In these cases audio connectors do not seem very useful, as only few people might understand the original Somali.

Another very interesting example is the German film *Im Juli* (*In July*, 2000) directed by the German-Turkish director Fatih Akin. In many of his films, German and Turkish dialogues are mixed together and Turkish dialogues are subtitled in German for the German. The audio subtitling is quite challenging because of the very frequent cases of language-mixing, e.g. speakers mix the two languages in one sentence (starting in German and ending in Turkish) or change the language from question to answer.

Combining all this with the description is a very complicated task, and it is hard for an audience to understand.

In the following example there is a man called Isa speaking Turkish and German, a man called Daniel speaking only German, and a border guard speaking only Turkish. Isa and Daniel are sitting in a car at the Turkish border and Isa realizes that Daniel hasn't got his passport:

DESCRIPTION:

Isa and Daniel in the car:

SPOKEN:

I: Get Out

D: I am not getting out

I: I told you to get out

DESCRIPTION:

A border guard:

SUBTITLING:

Are you blind? Why don't you drive forward? Get out!

DESCRIPTION:

Isa gets out

SUBTITLING:

G: He has to get out too!

SPOKEN:

I: You have to get out!

SUBTITLING:

G: What did you say?

I: That he has to get out!

G: That's what I said

I: He is German. He does not understand.

G: Ah! Deutsch, deutsch... Passport and registration papers!

DESCRIPTION:

Isa gives him his passport and the registration papers.

SUBTITLING:

I: His passport was stolen.

G: Tell this to your grandma

SPOKEN:

D: What is the problem?

I: Halt's Maul!

SUBTITLING:

G: What?

I: I told him to Shut Up!

SPOKEN:

G: Halt's Maul!

D: I didn't say anything!

SUBTITLING:

G: He can't get through without papers! Open the trunk!

DESCRIPTION:

Isa opens the trunk. The guard discovers the dead body.

## CONCLUSIONS

All these examples show that audio subtitles may greatly affect the work of describers in a dubbing country. However, if the describers succeed in achieving a fine and balanced interaction between description and subtitling, this could bring a new dimension into the listening to audio described films.

# Environmental description

RIITTA LAHTINEN

PhD, Head of Communication Unit,  
the Finnish Deafblind Association

RUSS PALMER

International Music Therapist

## ABSTRACT

*Environmental description is the description of general, physical, personal and social space and action, where visual, auditory and other sensory information is shared with the receiver in spoken, written or sign language, either vocally, i.e. producing sounds or in another form (pointing, touching, drawing). It can be divided into the expression of basic characteristics, basic description, precise and extended description, and it can be carried out physically on the spot (close description) or far away from the target (distant description). Description can be carried out spontaneously in real time, in joint action systematically i.e. pre-prepared description or it can be recorded beforehand as a text format, or as consecutive i.e. a postponed description after the event. The target group may be one person or a group. In addition to verbal description, environmental description can be produced with various sounds, such as vocalization without words or other sources of sounds e.g. musical instruments. Interaction in a situation between the describer and the receiver may be a one-way description or a dialogue. It can further be divided into functional dialogue, reciprocal description supporting sensory perceptions, telling and pointing in front of the target, reciprocal description by drawing or through movements and the exploration of objects. In detailed descriptions the main subjects are followed by details. Description can be classified according to the size of the space that is extensive, large, in a room or nearby.*

*The purpose of this article is to present the possibilities and different areas of description more widely, from the point of view of visually impaired and dual sensory impaired/deafblind persons receiving environmental description. Description may be classified in various ways, for instance, in accordance with possibilities for interaction (one-way or reciprocal) or according to the spatial perception of the visually impaired person. Description methods are also classified. Drawing on the back is presented as an example of how to use the body in receiving description.*

## BACKGROUND OF DESCRIPTION

Description has always been used for persons with visual impairment. Various professional groups use description in their work in order to clarify their actions and support their instructions: they verbalize their visual view. For instance, mobility instructors describe space and routes, IT instructors describe the computer screen, keyboard and other equipment, while physical education instructors describe body movements. Description and verbal information are often needed and can be useful even for sighted people.

Description is included in the curriculum of various professional education fields such as in sign language interpreter education, programmes for 'interpretation for various client groups' and in audio describer education. A 64-hour audio describer course was first introduced in Finland in 2005 by the Cultural Services for the Visually Impaired with the support of the Ministry of Education. By 2011, three groups had already received an audio describer education (Turunen 2005, [www.kulttuuripalvelu.fi](http://www.kulttuuripalvelu.fi)).

Traditionally, description for people with visual impairment includes audio description of visual arts, movies and theatre. This method has made culture available for people with visual impairment. Interest towards description has been increasing, due, for instance, to descriptions (narration tracks) inserted in movies (DVD). In addition, further university studies of audio description have been made of how to describe movies using sound tracks. However, there are no studies on the wider, multimodal use of the senses as part of receiving description.

In sign language interpretation for people with dual sensory impairment or deafblindness, environmental description is one of the areas of interpretation. It supports language interpretation for the deafblind people (for instance signing in free space, tactile i.e. hands-on signing, repeating speech) and moving (guiding a person), or it can be a separate action, such as the description of a target. Thus, in the 1980's, description was already included in sign language interpreter education.

## WHAT IS DESCRIPTION?

What is description? Is it just the verbalization of visual information? And how does the describer choose how and what to describe? If the receiver has a hearing impairment in addition to a visual impairment, the concept of description becomes wider, thus also including the description of auditory information.

The overall aim of multiple environmental descriptions using various senses of deafblind people give them access to information. The aim of the study is to reflect description widely from the points of view of the describer and receiver. Thus, for instance, possibilities to use various senses (sense of touch, movement) during description must be included, as well as how personal sensory perceptions can be a basis for description. If the person reacts to a change in the surroundings, such as a smell or a vibration, the describer can then describe the action which caused this environmental change.

In the field of deafblindness and dual sensory impairment, the description is a wide description of general, physical, personal and social space and action, where visual (sight), auditory (hearing) and other sensory information is shared with the receiver either by language (spoken, written, sign language), vocally or in another agreed form (pointing, touch, drawing, movement, etc.) (Lahtinen 2004, Lahtinen & Palmer 1996, 1997; Raanes 2004). Information shared by touching the other person's body can be used as a support to language description or independently.

During description, there are at least two people present, one who produces and another one (or a group) who is the receiver. Description can be carried out in real time and close to the target or it can be recorded or written beforehand. In a theatre description, the describer may physically be elsewhere, even though the description is being listened to in real time through headphones. Description may also be carried out consecutively after the event. In this case, though it is not possible to carry out the description in real time during the event, the information still needs to be understood. More generally, even haptics, i.e. touch messages, and their haptemes, i.e. the grammar of touch, are closely related to description. Lahtinen (2008) describes extensively how various things can be conveyed and described on the body by touch. The study is based on practical observations carried out over 20 years, while working with both deafblind and blind persons with hearing.

## ENVIRONMENTAL DESCRIPTION - WHY AND FOR WHOM?

Environmental description supports the multimodal sensory perception of a person with visual impairment. With the help of description, it is possible to support visual information for partially sighted and blind persons, auditory information for hearing impaired persons (movies with visual text description),

as well as both visual and auditory information for deafblind persons. These can be connected with exploring targets haptically, such as feeling various objects (statues, paintings etc.). Receivers of environmental description can be deafblind, dual sensory impaired, hearing impaired and visually impaired children, youth and adults. In addition, the elderly and people of different ages with learning problems will also benefit from this type of instructions and descriptions.

#### EXTENSION OF DESCRIPTION

The extension of description may vary depending on the extent of the target's contents. It can be naming a target, using one word, or an extended description. Lahtinen, Palmer and Lahtinen (2010: 22) divide description into four entities. Description of basic characteristics is carried out almost by everybody on a daily basis, for example when you name a target and say something about it. Basic description includes the expression of basic information on the target. A more detailed description means a more precise description of the target. An extended description can even comprise of information on the target that is not immediately visible, such as the background of an artist, a family history or events during different eras. Description can vary a lot with respect to time. When preparing a description, it is important to consider the extension of the verbal description situations that can be successfully received by the receivers.

#### AREAS OF DESCRIPTION

Description can be divided into different areas. In Figure 1, description is divided according to language and methods. Language description includes description by spoken language, sign language and written text. Descriptions by spoken and sign language differ in that language structures work in different ways. Even though the content of the description is the same, the expression via spoken/written language may be longer due to their linear structure. When naming in sign language, the places and relations are also embedded, so that they don't need to be indicated separately. Vocal description includes description produced by different sounds. The vocal environment is mainly an amendment used when describing the atmosphere, like the humming of the wind when describing visual arts, the sound of steps etc.

The concept of haptics related to exploring the target includes the sense of touch and exploring the environment, where skin, muscles and joints are part of the data collection system. There is no specific sense for that. Then a reciprocal interaction is experienced in contact with the environment (Gibson 1966, 1983). In exploring the target model, haptic information is used, when the explorer can, for instance, distinguish size differences, forms, surface qualities and materials

through his hands (Klatzky & Lederman 1993: 603-604). Haptics, as a part of description, refers either to the receiver's exploration of targets by himself or together with the describer, when the person's perception is combined with his touch and movement information.

To support the description, miniatures may be constructed if the exploration of the target by hand is almost impossible, such as in the case of large buildings, Eiffel Tower. There are recent technical solutions that, with technical support, enable the conveying description. For instance, when moving, you can receive information related to your own whereabouts through a mobile phone (place names, route instructions).

Drawing on the back and body has given new and different possibilities to experience, for instance, arts. Targets are drawn on the person's back (paintings, drawings, pictures). A large surface, the back allows for the description of even big targets compared with the small size of a palm of a hand (Lahtinen et. al 2010).

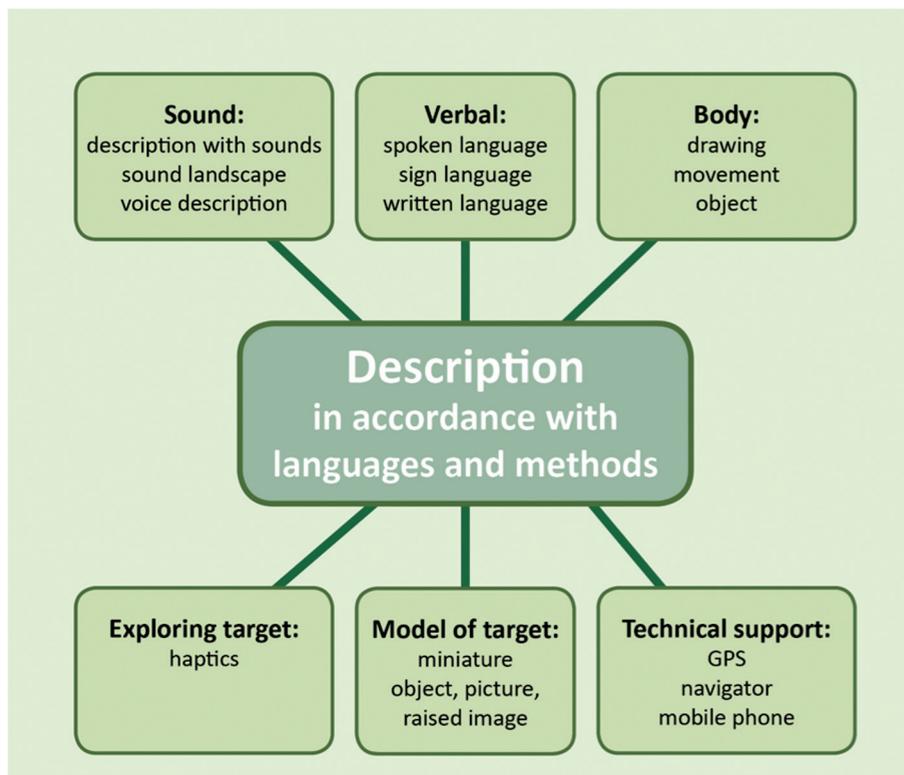


Figure 1. Description according to languages and methods (Lahtinen et al. 2010: 43).

## ONE-WAY OR INTERACTIVE DESCRIPTION?

Description can also be examined as an interactive process (Figure 2). When a describer or an audio describer describes a target, e.g. a theatre play (narrated in real time with technical equipment) or as a pre-recorded part of a movie, then it is a question of a verbal, one-way description. In that case, the receivers cannot influence the situation, for instance, by asking more precise questions. When the receiver is an active person who can ask questions and act, the description becomes interactive. A functional description gives the receiver an opportunity to move and make the described movement with his own body. Teaching physical exercises is an example of active listening to a description while carrying out a certain performance. The describer may also connect his own body movement into a description and let the receiver feel the movements.

There are various possibilities in finding the target in description situations: it can be found with the help of verbal clues, sounds (e.g. audible beacon) or it can be pointed at with the leading hand. Then the describer will physically lead the hand of the receiver onto the target (tracking). These physical methods are influenced by the receiver's wishes regarding the ways of proceeding, the acceptance of various methods and the size of the receiver group.

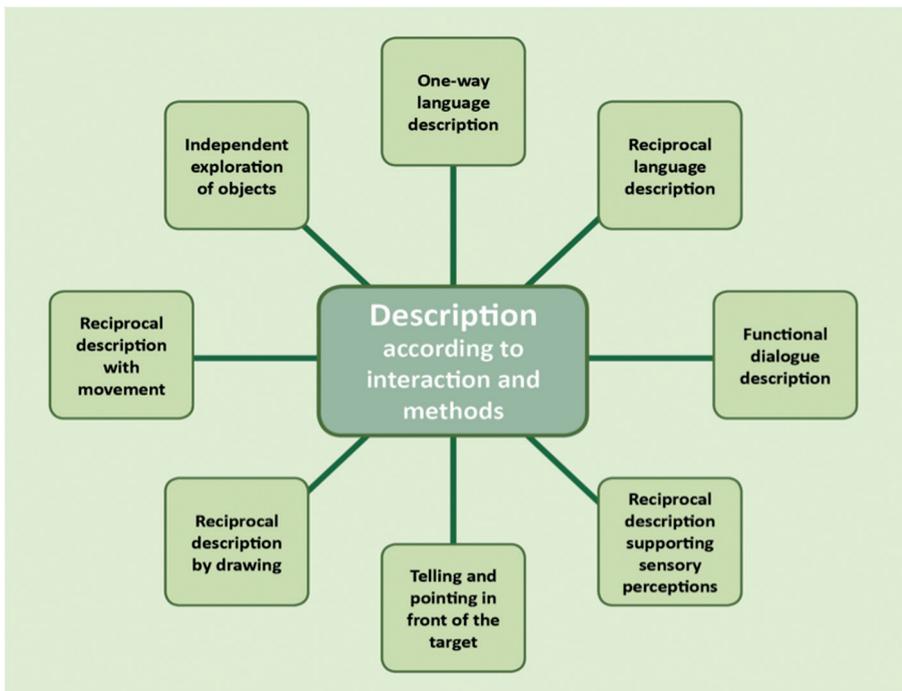


Figure 2. Interaction and methods of description (Lahtinen et al. 2010: 44).

While working with people with visual impairment, it is good to know about their spatial orientation and how different senses are used on the basis of mobility (Törrönen & Onnela 1999). A basic element for the independent moving of people with visual impairment is their awareness of their own position in the space, how to get there and how to leave it. Perceiving space is a process that demands thinking and logic. For a hearing person with visual impairment, the sounds in space tell and give clues on, for instance, the size of the room, materials, activities and distances (Hirn et al. 2009).

Figure 3 shows how the space can be divided according to its width and distances. A description of a general wide space means describing a space without borders. When standing on the top of a hill, we have a 360-degree view around us and we can see far away, whereas an extensive space has frames, there are certain structures around, such as, for example, in an art exhibition hall. When description is carried out in a room, it is possible to name the walls and contours of the space. When the receiver is in a certain location within the space, for instance sitting on a bench, there is a so-called near space/personal space around him. He can feel and explore with his hands, feet and body. The social space around him includes other people and their actions and changes of actions. Details, actions and deviations can be picked out from the space to be described (Lahtinen et al. 2005).

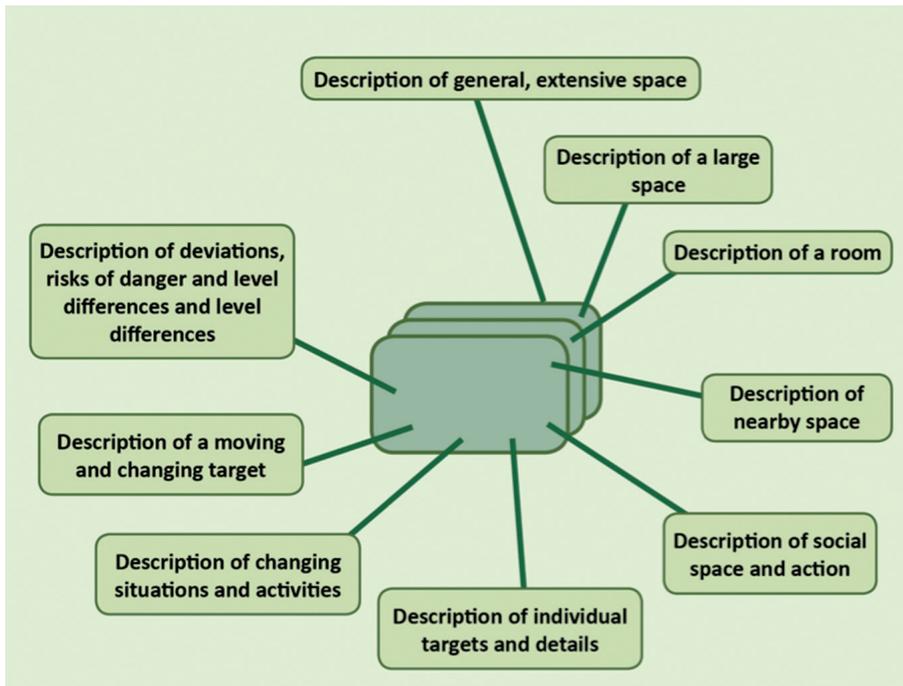
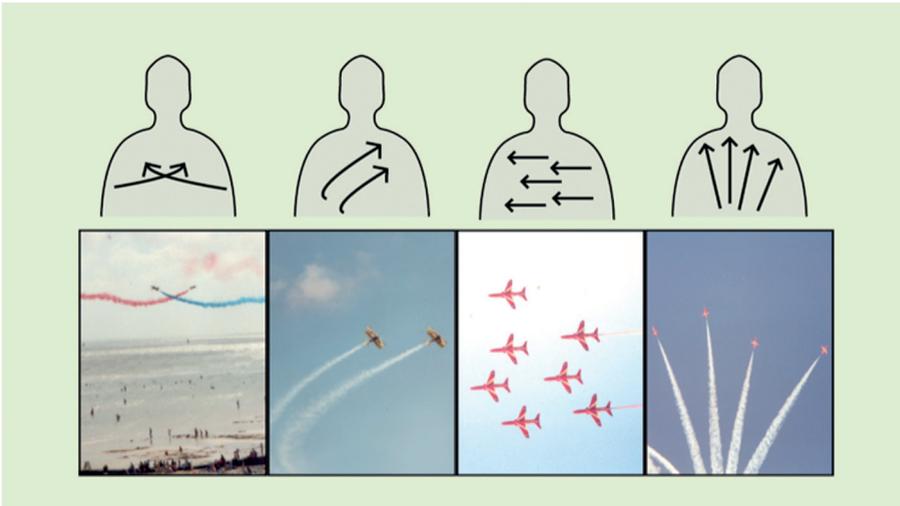


Figure 3. From wide description into details (Lahtinen et al. 2010: 49).

## USING THE BACK AS ONE POSSIBILITY FOR DESCRIPTION

An interesting method for drawing and describing targets is the use of the receiver's back. Studies are being carried out on how many different things can be described on the back. The example in picture 1 shows how interests, such as rapid and changing movements of airplanes during an air show can be indicated simply by drawing with the index finger. Then the back works as a scene, the flying area of the airplanes. Drawing is most often done on the upper part of the back. By drawing, the flying routes, the speed of the movement (by adjusting the speed of the airplanes) and the number of the planes (by drawing with both hands - one, two or more) can be described. The receiver will even perceive the forms of the flight paths, directions (from left to right, right to left, downwards, upwards), changes in directions (same direction, different directions) and sizes. Those who use drawing on the back usually agree on the method beforehand.

Picture 1 is an example of visual flying routes, describing directions on the back (Lahtinen et al. 2010: 131).



Picture 1. Ways of describing flying routes on the back of a person.

## FUTURE REFLECTIONS

The development of audio describing will give people with visual impairment totally new possibilities for similar services of accessing visual arts through description. However, just as the availability of description and audio describers is still in its initial phase, so the availability of accessible services for persons with visual impairment still needs to be fostered by society at large.

In the future, interesting aims of description studies include, for instance, how two and three dimensional targets or pictures can be drawn on the body,

how shared movements are produced and interpreted by another person and what kind of grammatical structures are connected with these.

Accordingly, description is more than an output expressed by spoken or written language. When it comes to targets to be described, other methods can be used than traditional cultural activities. In principle, any phenomenon may be found interesting by a person with visual impairment - even Jupiter and its many moons.

## REFERENCES

- Gibson, J. J. 1966, 1983. *The Senses Considered as Perceptual Systems*. Boston, USA: Prospect Heights, I. L. Waveland Press, Houghton Mifflin Co.
- Hirn, H. 2009. *Pre-Maps: an Educational Programme for Reading Tactile Maps*. Research report 302. Helsinki University. Helsinki: Yliopistopaino.
- Klatzky, R. L. and Lederman, S. J. 1993. "Toward a Computational Model of Constraint-Driven Exploration and Haptic Object Identification." *Perception* 22: 603-604.
- Lahtinen, R. 2005. *Social Quick Messages*. Finnish Deafblind Association Tampere: Cityoffset Oy.
- Lahtinen, R. 2004. *Kuvailu*. Sujuva tulkkaus kuurosokeille. Selvitys kuurosokeiden tulkinkäytön kokemuksista. Julkaisuja B2, raportti 2. Finnish Deafblind Association. Tampere: Cityoffset Oy.
- Lahtinen, R. 2008. *Haptices and Haptemes. A Case Study of Development in Social-Haptic Communication of Acquired Deafblind People*. Helsinki University, Faculty of Behavioural Sciences. Doctoral Dissertation. Tampere: Cityoffset Oy.
- Lahtinen, R. and Palmer, R. 1996. *Holistic Family Communication, Spoken Language by Touch is More than Just Words*. 4<sup>th</sup> European Deaf-blind Conference Espoo, Finland.
- Lahtinen, R. and Palmer, R. 1997. *Theoretical Basis of Holistic Communication for Dual-Sensory Impaired People & Family Members*. Madrid, Spain: EUSSG & DBI proceedings.
- Lahtinen, R., Palmer, R. and Lahtinen, M. 2010. *Environmental Description*. Helsinki: Art-Print Oy.
- Raanes, E. *Kuvailu*. 2004. Työteksti nro 37, (Beskrivelse 2001), 13. NUD: Denmark.
- Turunen, T. 2005. "Kuvailutulkkit valmistuivat!" *Kajastus* 6, 12-14.
- Törrönen, S. and Onnela, J. 1999. *Vapaus tulla, vapaus mennä! Sokean ja heikkonäköisen liikkumistaitojen kehittäminen*. Arlainstituutin julkaisuja 2. Gummerus Kirjapaino Oy, Saarijärvi.
- Kulttuuritapahtuman kuvailu näkövammaiselle. Saatavissa: <http://www.kulttuuripalvelu.fi/id=9> (accessed 25 August 2010).
- Kuvailu. Turun ammattikorkeakoulu. Viittomakielentulkkipolustus. Saatavissa: [http://www.diak.fi/av\\_tarjonta10?yks=Turun%20toimipaikka#5](http://www.diak.fi/av_tarjonta10?yks=Turun%20toimipaikka#5) (accessed 25 August 2010).

# Bionotes Authors

SAVERIA ARMA is a free lance conference interpreter, translator and respeaker. She is completing a PhD in English for Special Purposes with a corpus-based analysis of language in filmic audio description. She actively works in the field of live and pre-recorded subtitling and audio description.

BERND BENECKE is the head of Audio-Description at Bayerischer Rundfunk - the public broadcaster for Southern Germany. He started with Audio Description in 1989 when the first movie was described in Germany and worked as an Audio Description author and narrator before he built up the AD department at BR in 1997. Today he mainly revises Audio Descriptions and directs the narrating and sound mixing process.

AGNIESZKA CHMIEL is Assistant Professor at the Department of Translation Studies, School of English, Adam Mickiewicz University (AMU) in Poznan, Poland. Her research interests include conference interpreting, audio description, audio-visual translation, cognitive studies, memory and visual imagery in interpreting. [www.aga.e-chmiel.pl](http://www.aga.e-chmiel.pl)

ANNA JANKOWSKA: A graduate of Spanish Philology at Jagiellonian University in Krakow, she is now an assistant at the UNESCO Chair for Translation Studies and Intercultural Communication there. Interested in audiovisual translation, in particular AD and SDH, she is preparing her doctoral thesis on the possibility of translating audio description from English into Polish. She teaches audiovisual translation and works in this domain as well as in the one of general translation.

RIITTA LAHTINEN is a Communication Researcher Consultant . She has defended Haptices and Haptemes, her PhD thesis, in 2008. The thesis has been awarded an innovative prize in Finland. She is involved in running International Workshops on Communication Methods for deafblind people, couples, interpreters and professionals in Europe and Scandinavia, focusing on the needs of deafblind people, and training interpreters, planning University academic courses and publications. <http://www.kolumbus.fi/riitta.lahtinen/>

NATHALIE MÄLZER-SEMLINGER works as Wissenschaftliche Mitarbeiterin at the Institute for Translation Studies and Technical Communication at Universität Hildesheim. She has completed her PhD on “The Transfer of French Literature to Germany between 1871 and 1933” in 2009. Her research interests are: AVT, dialogue studies and transmediality.

IWONA MAZUR is Assistant Professor at the Department of Translation Studies, School of English, Adam Mickiewicz University (AMU) in Poznan, Poland. Her research interests include audiovisual translation in general and audio description in particular. Her audio description research focuses on reception as well as perception (eye-tracking) studies.

More information: [http://ifa.amu.edu.pl/fa/Mazur\\_\\_Iwona](http://ifa.amu.edu.pl/fa/Mazur__Iwona)

PILAR ORERO. PhD (UMIST). Works in the CAIAC Research Centre (Universitat Autònoma de Barcelona, Spain). Director of the European eMA in Audiovisual Translation at UAB, and now is the director of the Online European MA in Audiovisual Translation <http://mem.uab.es/metav/>. Leader of numerous research projects funded by the Spanish and Catalan Gov. Leads TransMedia Catalonia <http://grupsderecerca.uab.cat/transmediacatalonia>

RUSS PALMER is an International Music Therapist, songwriter, musician and writer. His philosophy is called Feeling the Music. Russ has expressed his Feeling the Music philosophy by releasing his new CD Warm Summer Days. Through musical expression and therapy he applies a concept and understanding of how people who are losing their hearing and sight can develop other senses in their bodies and how to interpret rhythm and vibrations. These interpretations can be relayed further using the methods of Environmental Description. <http://www.kolumbus.fi/riitta.lahtinen/>

AGNIESZKA SZARKOWSKA, PhD, is an Assistant Professor at the Institute of Applied Linguistics at the University of Warsaw. Her main research interests are audio description and subtitling for the deaf and the hard of hearing. Her current research projects include text-to-speech audio description, subtitling for the deaf and hard of hearing in multilingual films, audio description in education and the application of eye tracking to subtitling. She is a member of ESIST and an honorary member of the Polish Association of Audiovisual Translators (STAW).