

# 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.

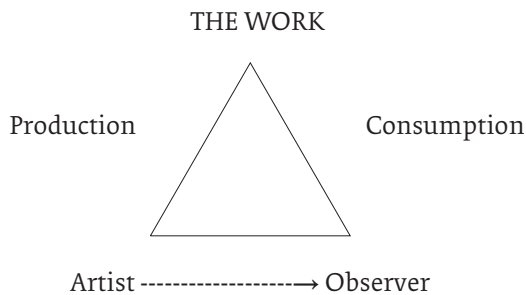


Figure 1 . Triangle of the artistic experience (Monaco 1977: 34).

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.

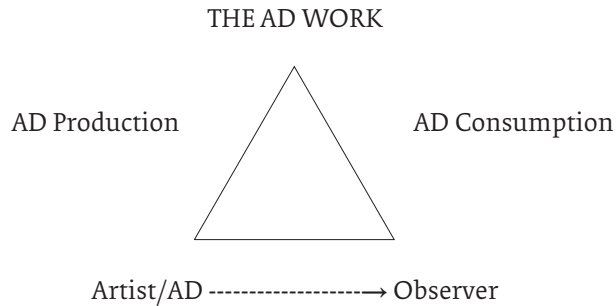


Figure 2 . Triangle of the AD artistic experience.

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.

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.

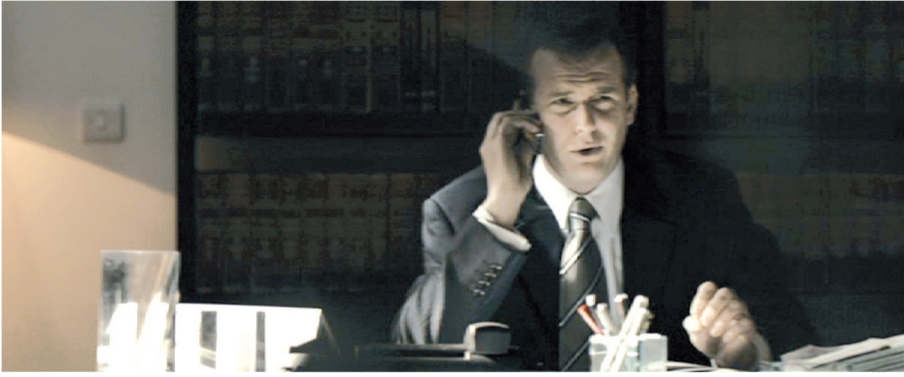


Figure 4. Frame depicting the lawyer.



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).

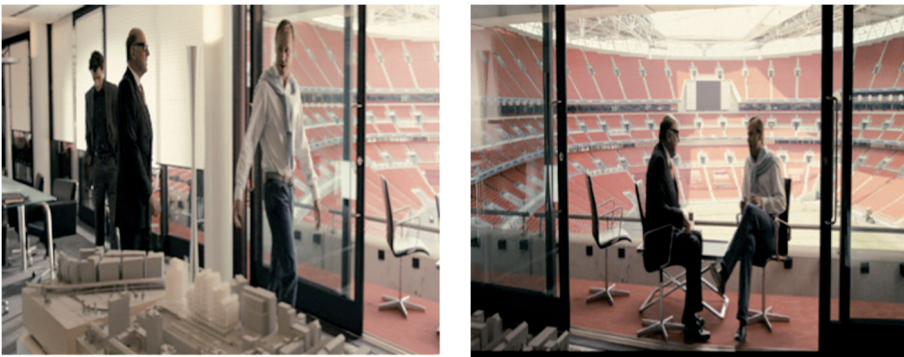


Figure 8. Frames depicting the Emirates Stadium in London.

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.

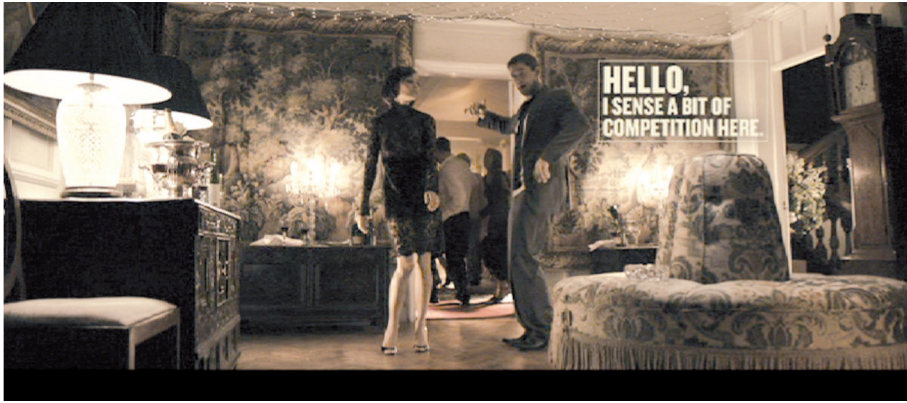


Figure 12. This frame recreates the front cover of the *Hello* magazine.

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.