
This is the **published version** of the book part:

Fresno, Nazaret; Castellà, Judit; Soler-Vilageliu, Olga. «‘What Should I Say?’ Tentative Criteria to Prioritize Information in the Audio Description of Film Characters». A: Researching Audio Description New Approaches. 2016, p. 143-167.

This version is available at <https://ddd.uab.cat/record/275936>

under the terms of the  IN COPYRIGHT license

What Should I Say? Tentative Criteria to Prioritize Information in the Audio

Description of Film Characters

Nazaret Fresno, Judit Castellà, and Olga Soler-Vilageliu

1. Introduction

Toda mi vida me ha encantado el cine y, cuando el médico me dijo que me quedaría ciego, pensé que sería una de las cosas que más echaría de menos: ver películas. Pero gracias a la audiodescripción, todavía disfruto del cine. Diferente de como lo hacía antes, pero aún lo disfruto¹ (Participant 24 in our test).

Audio description (AD) allows users to comprehend and enjoy audiovisual products. Explored in academia only since 2000, the last fifteen years have provided extensive research that has served as the ground basis for the discipline. Descriptive studies have focused on the filmic aspects of AD (for example, Pérez Payá, 2007; Maszerowska, 2012, 2013; Fryer and Romero-Fresco, 2014), on its linguistic characteristics (for example, Piety, 2004; Bourne and Jiménez, 2007; Matamala and Rami, 2009; Arma, 2012 and the research obtained from the corpus studies used in TIWO and TRACCE research projects) and on its narrative dimension (for example, Braun, 2007; Kruger, 2010; Braun, 2011; Remael, 2012; Vercauteren, 2012). These descriptive approaches have been complemented, especially in the last five years, with experimental research that has analysed empirically different aspects of the reception of filmic products by both sighted spectators (for example, all the research gathered in Mazur and Kruger, 2012; Orero and Vilaró, 2012, 2014; Vilaró and Orero, 2013) and blind and visually impaired (BVI) audiences.

At first, the studies within the latter category were scarce and analysed general users' preferences (for example, Rai, 2009; Chmiel and Mazur, 2012). However, the scope of the empirical research has progressively expanded to cover specific areas of interest, such as technical issues (for example, Szarkowska, 2011, Szarkowska and Jankowska, 2012; Matamala, Fernández and Ortiz-Boix, 2013) or filmic aspects of AD (for example, Fryer and Freeman, 2013; Romero-Fresco and Fryer, 2013). More recently, experimental research has also moved closer to psychology and cognition in order to explore how users receive, comprehend and experience audio described products (for example, Fryer and Freeman, 2012; Cabeza-Cáceres, 2013; Ramos, 2015).

Following this approach, Fresno, Castellà and Soler-Vilageliu (2014) conducted an experiment which departed from the premise that, just like sighted viewers, BVI audiences achieve film comprehension thanks to their memory. However, working memory, responsible for the brief storage and manipulation of information while performing complex cognitive tasks, is capacity-limited, and that might be the reason why not all the details received from an audio described film can be remembered by its addressees. In relation to memory, research in the fields of cognitive psychology, media studies and education suggest a number of findings relevant to AD.

To start with, our recall of visual information is more robust than that of auditory details. Viewers can rapidly identify the gist of complex visual scenes and they are able to recall them with details, even after being exposed to the visual materials for a very brief time (for example, Shepard, 1967; Standing, 1973; Brady et al., 2008). However, performance decreases when auditory materials are used (Cohen et al., 2009). Also, video-based materials are more effectively processed and recalled than audio-based information (Graber, 1990; Basil, 1992) and audio/video redundancy seems to favour the receivers' memory (Lang, 1995; Fox, 2004). Furthermore, as predicted by Paivio (1986) in his Dual Coding Theory, in learning contexts students obtain better results

when instructional materials combine words and images, rather than when they present words alone (Mayer, 2001; Eilam and Poyas, 2008; Butcher, 2014). All these insights could have important implications for AD, in which audio describers convey visual information in a fragmented verbal narration that is received by the addressees through the auditory channel. Nevertheless, AD users, that is BVI individuals, are believed to compensate for their lack of vision through a better development of other senses (sensory compensation), and they are also thought to possess a better memory for auditorily transmitted materials. However, empirical research exploring this hypothesis indicates that it might only apply in the case of congenitally blind individuals (Röder, Rösler and Neville, 2001; Amedi et al., 2003), who are a very small group within the potential users of AD. Therefore, even if their memory superiority is real, it should not be taken as a valid reference for AD, since the vast majority of the BVI audiences were born sighted, acquired blindness at different stages of their lives and show a memory performance comparable to that of sighted viewers.

Taking into account this theoretical framework, Fresno et al. (2014) explored the reception of film characters in AD from a cognitive perspective. Acknowledging that BVI audiences create and update mental models of characters in their attempt to understand filmic plots (Fresno, forthcoming), and that working memory is involved in those processes, they explored the effect that the amount of information included in the AD and its presentation had on the recall and reception of characters by BVI audiences. The results of their quantitative analysis showed that memory is indeed affected by both factors, since more information was recalled and recognized when short or segmented ADs were delivered, as opposed to long and unsegmented descriptions, as will be later described in Section 2.

In this paper, an expansion of the aforementioned research will be presented. In order to complement our previous quantitative study, qualitative analyses exploring the

nature of the information which is more frequently recalled and recognized were carried out. The following sections will describe the experiment conducted, the analyses performed, the results obtained and the implications that those might have for the professional practice of AD.

2. The Current Study

This work was carried out as part of a greater research project described in Fresno et al. (2014). It departs from the premise that characters are prominent elements within filmic narratives and, hence, they should be audio described in order to provide BVI audiences with the information regarding their physical appearance. Several AD guidelines offer recommendations on what should be described. For instance, the Irish standards point out that, provided there is enough time, ‘dress, physical attributes, facial expressions, body language, ethnic background (if relevant to the storyline) and age should be audio described’ (Broadcasting Authority of Ireland, 2012: 1). Also, Ofcom (2012: 17) states that ‘when describing characters, aspects such as dress, physical characteristics, facial expression, body language, ethnicity and age may be significant.’ At the same time, however, AD should not be ‘exhausting’ or ‘irritating’ (The ITC, 2000: 14), and should not ‘provocar cansancio en el oyente discapacitado visual’ (AENOR, 2005: 7).² Therefore, audio describers should find a balance and create ADs which are sufficiently informative and evocative as to allow the audience to imagine film characters, but which do not provide excessive details in order to avoid tiredness.

Keeping this in mind, Fresno et al. (2014) measured quantitatively the amount of information that BVI users recalled and recognized after listening to audio described self-contained excerpts from films and TV series which contained long and short unsegmented and segmented character descriptions. Long ADs included eight traits of

the characters, whereas short ADs included four. Unsegmented descriptions were delivered as a single block of information, as opposed to segmented ADs, which were split into two blocks of four traits that were presented at different stages of the script. This experiment showed that short and segmented ADs were better recalled than long and unsegmented descriptions, respectively. However, even in the best condition, average free recall rates barely surpassed 50 per cent and recognition rates did not exceed 70 per cent, which seems to suggest two ideas: first, that the recall of audio described characters by BVI audiences is not comprehensive and, second, that creating long descriptions of characters does not seem the best strategy in terms of users' memory. Therefore, even if the film allows for detailed ADs, it might be advisable to keep character descriptions short or, at least, not very long, so that receivers have a better chance to remember them.

The fact that long descriptions are not properly remembered highlights the need to establish criteria that allow for information prioritization in AD scripts. Audio describers are supposed to select the information that they provide to their audience according to its relevance in the plot. Nevertheless, this is not as easy as it may sound when it comes to characters, since not all of them have what could be called 'prominent features.' In the case of very prototypical genre-specific characters, it may be easier to select the most relevant information. For instance, when audio describing the Joker from the film *The Dark Knight* (Nolan, 2008), audio describers will refer to the white make-up that tries to hide the character's scarred face, to the asymmetric black paint that surrounds his eyes as if imitating a deadly clown, and to the nightmarishly smudged red make-up around his mouth that looks more like a rictus than an actual smile. Those three traits alone are able to provide a fair idea of the physical appearance (and, indirectly, also of the psychological condition) of the Joker. However, in the case of more neutral characters whose features are not so extreme, it may sometimes be hard to

prioritize information. For instance, John Watson, from *Sherlock* (Gatiss, Moffat and Vertue, 2010), is characterized as a regular man who wears regular clothes. None of his traits is marked in the TV series as more relevant than the rest and, hence, different audio describers might possibly include different traits in their scripts according to what they consider more informative to users. In these cases, the responsibility for choosing the most relevant information lies solely with the subjectivity of the scripts' creators, but it is our belief that moving the focus towards the audience could be of use. An alternative approach that explores what users remember of characters after listening to audio described films could provide some valuable insights to understand how audiences 'picture' characters in their minds in their attempt to reach film comprehension. Perhaps, some features of characters tend to be more salient than others in the sense that they occupy a more prominent position within the mental model of the characters created by AD addressees. If that were the case, those distinctive traits would be more frequently recalled and, hence, exploring users' memory for character descriptions could be of use in order to prioritize the information to be included in AD scripts. Cuing on this argument, the first hypothesis for our test assumed the following:

Hypothesis 1 (H1): Some physical traits of the characters are recalled and recognized better than others.

In order to test H1, a specific research question was posed:

Research Question 1 (RQ1): Which are the most frequently recalled and recognized types of traits of characters?

On the other hand, Fresno et al. (2014) found evidence indicating that the presentation of the AD affected its reception. Specifically, their research showed that dividing the character ADs in short bites of information delivered at different stages of the film contributed to the recall of a greater amount of information. Taking this finding into account, our second hypothesis foresaw the following:

Hypothesis 2 (H2): Segmentation of character ADs favours the recall of a more varied typology of information.

In order to test H2, the following research question was explored:

Research Question 2 (RQ2): Are more categories of traits correctly recalled and recognized when character ADs are delivered in a segmented manner?

Finally, receivers of written narratives are supposed to put more effort into outlining the mental models of the main characters (Schneider, 2001). Acknowledging their prominent role within the situation model also in audiovisual narratives, Magliano, Taylor and Kim (2005) assessed how film spectators monitor for certain mental states (specifically, goals) of several characters in the same filmic experience and found out that spectators observed the most prominent characters more closely. Departing from these ideas, our third hypothesis stated the following:

Hypothesis 3 (H3): As addressees of filmic narratives, AD users might pay closer attention to the most prominent characters, which could lead to differences in the recall and recognition between them and secondary characters in films.

In order to test H3, the following research question was posed:

Research Question 3 (RQ3): Are more categories of traits correctly recalled and recognized for main characters than for secondary characters?

2.1 Methods

The aforementioned hypotheses and research questions were explored by means of an experiment aimed at studying AD users' memory. In order to assess which were the features of characters most frequently recalled and recognized, they were classified into five categories: age, height and weight, facial features, hair, and clothes and other items. This selection of categories was preferred over other possibilities because of its unambiguity (each trait could be attributed to only one category), and balance (each category included a comparable amount of features to be analysed). For all of the categories in the experiment, both the information correctly recalled by the participants and their false recalls (features wrongly ascribed to each character) were analysed.

2.2 Participants

A total of 44 BVI participants took part in the experiment, 21 males and 23 females, aged 18 to 76 years ($M=48.43$; $SD=13.72$). We aimed at conducting a naturalistic experiment, which would reproduce a real AD context as closely as possible and, thus, the age of the subjects in the sample was not restricted. Representative subjects of all ages took part in the test, just like real addressees of all ages are potential AD users. 40 of the participants in the experiment were blind according to the World Health Organization standards (either they had an acuity minor to 0.05 or a visual field minor

to 10°) and four of them suffered from low vision (they had an acuity between 0.3 and 0.05 or a visual field minor to 10°).

2.3 Materials

For comprehension purposes, this section will offer a basic explanation of the materials used. A more comprehensive description can be found in Fresno et al. (2014).

2.3.1 The corpus

The corpus was created from several audiovisual materials: a self-contained excerpt (CAN) from the Spanish film *Caníbal* (Martín Cuenca, Bovaira, De Santiago and Hernández, 2013), a self-contained excerpt (PMS) from the Spanish-dubbed film *Pequeña Miss Sunshine* (Friendly, Dayton and Faris, 2006), and two self-contained excerpts (BB1 and BB2) from three episodes of the Spanish-dubbed version of the television series *Breaking Bad* (Gilligan and Cranston, 2010; Gilligan and McKay, 2008; Gould and Bernstein, 2010). All of them showed five characters on screen, and they were very similar in length (about nine minutes-long), number of words in the dialogues and speed of their utterance.

Four versions of AD were created for each clip (x 1+, x 2+, x 1-, x 2-). Two of them (x 1+ and x 2+) included long descriptions of characters, which mentioned eight physical traits and differed only in its presentation: one was unsegmented (x 1+) and the other included segmented character descriptions (x 2+). The other two versions of AD (x 1- and x 2-) showed short descriptions of characters which included four traits for each of them. One of those ADs was delivered in an unsegmented manner (x 1-) whereas the other one was segmented (x 2-). The rest of the ADs (that is, those parts of

the video descriptions in which the appearance of characters was not described) remained the same in the four AD versions of each clip.

Once the ADs were ready, they were recorded by a voice talent and mixed in a professional studio to obtain the final audio clips (.wav) that formed the corpus for the test. During the recording, the speed of delivery in all the ADs was controlled. According to Cabeza-Cáceres (2013), users' comprehension is comparable to that of sighted viewers when the AD is delivered at 14 characters per second. However, if the speed is increased, comprehension rates decrease. Therefore, the AD delivery in our experiment was controlled and limited to 14 characters per second (around three words per second).

Even though 16 audio clips were created and used in the experiment detailed in Fresno et al. (2014), only eight were considered for the current analysis: those showing conditions x 1+ and x 2+ of each audio clip. This is due to the fact that the main aim of the present experiment was to study the recall and recognition of the physical features of characters in order to find criteria that might help in prioritizing information. Our focus was to explore which categories of traits were better recalled and recognized by BVI users and, thus, we needed descriptions long enough to allow for a consistent analysis. That is the reason why we limited our scope to the long character ADs (x 1+ and x 2+), which included eight traits belonging to several categories, instead of short AD, in which only four features classified into less categories were mentioned.

2.3.2 Instruments

A questionnaire was designed by our team to assess participants' free recall and recognition of the physical features of audio described characters.

The free recall part included three questions aimed at assessing three issues: the participants' perception of their own comprehension; if they had been able to imagine the characters after receiving the auditory information in the clips together with their AD, and what physical traits of the characters they recalled freely.

In contrast, the recognition part of the questionnaire consisted of 'yes or no questions.' Participants could also answer 'I do not remember', but they were instructed to avoid this option if possible. Half of the questions in the recognition task presented the real physical traits explicitly mentioned in the ADs of the characters, whereas the other half of the questions mentioned invented features or traits included in the ADs of other characters. To distract participants from the real aim of the study and avoid them foreseeing what they would be asked about in coming clips, some more questions about other issues unrelated to characters were included in this part of the questionnaire (mainly questions about the settings or objects described in each clip).

2.4 Results

RQ1, RQ2 and RQ3 aimed at exploring three issues. Firstly, the nature of the information that BVI tended to remember more frequently from audio described characters. Secondly, whether the segmentation of the descriptions had a positive effect on their recall and recognition, and finally whether more categories of physical traits were remembered for primary characters than for secondary characters. In order to assess the results, the participants' answers in the two parts of the questionnaires (free recall and recognition) were treated separately.

Firstly, data obtained from the free recall questions was analysed. An ANOVA on recall proportion was conducted with Category and Block (1 vs 2) as within-subjects factors and Character (main or secondary) as between-subjects. Only Category and

Block effects were significant ($F(4,64)=3.702$; $p<.009$ and $F(1,16)=8.507$; $p<.01$, respectively).

Pairwise comparisons on Category were also performed, showing significant differences between 'age' and the rest of categories (all $p<.05$) except 'height and weight.' Figure 8.1 shows the mean proportion of correct recall as a function of category in the free recall task. Recall was significantly higher when information was presented in 2 blocks ($F(1,16)=8.507$, $p<.001$ (1 block, $M=.387$; 2 blocks, $M=.487$)).

<INSERT FIGURE 8.1 HERE>

Figure 8.1 Mean proportion of correct recall as a function of category in the free recall task

Data obtained in the recognition questions was then analysed. An ANOVA on correct recognition proportion was conducted with Category and Block (1 vs 2) as within-subjects factors and Character (main or secondary) as between-subjects. A significant main effect of block was found ($F(1,18)=4.284$; $p<.053$), showing better recognition when information was presented in 2 blocks as compared to one block ($M=.736$ vs $M=.664$). Category effect was also significant ($F(4,72)=15.318$; $p<.000$) but it was mediated by a significant interaction between Character and Category ($F(4,72)=2.854$ $p=.044$), that was due to 'hair' being better recognized for the main characters ($M=.749$ vs $M=.529$). As in the recall results, Character main effect was non-significant, suggesting that categories were equally recalled for main and secondary characters.

Pairwise comparisons on Category showed significant differences between 'age' and the rest of categories (all $p<.001$). Also, 'height and weight' was significantly

higher than ‘facial features’ ($p < .041$). Figure 8.2 shows the mean proportion of correct recall as a function of category in the recognition task.

An ANOVA was also conducted on recognition errors and ‘no recognition’ scores. Only Category effects were significant in both measures ($F(4,72) = 3.844$; $p < .007$ and $F(4,72) = 21.471$; $p < .000$, respectively). ‘Age’ had significantly less recognition errors and less ‘no recognition’ scores than the other categories (all $p < .05$). Moreover, ‘height and weight’ also had significantly less ‘no recognition’ scores than the rest of the categories.

<FIGURE 8.2 ABOUT HERE>

Figure 8.2 Mean proportion of correct recall as a function of category in the recognition task

Some considerations, aside from the statistical analysis, are worth noting at this stage. Firstly, the free recall of characters was very poor, even in the best condition of the experiment (x 2+). For instance, the average free recall of ‘age’, the category best recalled by the participants, was only slightly over 60 per cent. From the rest of the categories explored, only ‘height and weight’ surpassed 50 per cent. As per recognition, the average rates were higher, exceeding 90 per cent in the case of ‘age’ and reaching almost 75 per cent for ‘height and weight.’ The rest of categories showed worse memory performance and a higher percentage of ‘errors’ and ‘no recognitions.’ Figures 8.3 and 8.4 show the free recall and recognition results in the best condition, respectively.

<INSERT FIGURE 8.3 HERE>

Figure 8.3 Average free recall of categories in the best condition of the test (x 2+)

<FIGURE 8.4 ABOUT HERE>

Figure 8.4 Average recognition of categories in the best condition of the test (x 2+)

3. Discussion

The present experiment has explored empirically the recall of audio described characters by a BVI audience. We addressed three research questions: what kind of physical features were more frequently recalled and recognized from audio described characters (RQ1), whether the segmentation of the descriptions affected their recall and recognition (RQ2), and whether more categories of traits were remembered for primary characters than for secondary characters (RQ3). The results obtained in the free recall and recognition tasks will first be discussed in relation to RQ1, then to RQ2, and finally to RQ3.

In order to find an answer to RQ1, statistical analyses were carried out, which showed differences in the recall of the categories of traits explored. Namely, the free recall of ‘age’ was significantly higher than that of all the rest of the categories, except ‘height and weight.’ Also, ‘age’ was statistically better recognized than all the other categories analysed, and it showed considerably less errors and ‘no recognitions’ (that is, participants committed less mistakes and answered ‘I don’t remember’ less frequently when asked to recognize the age of the characters in comparison to the rest of the categories). The fact that ‘age’ was consistently better recalled and recognized than

the rest of categories could be due to two reasons related to each other: firstly, to the fact that the information regarding the age of characters is reinforced by their voice in the film and, secondly, to the fact that the voice is the only trait to which AD users are exposed in a continuous manner. Even though listening to the voices of characters does not mean being able to guess their exact age, it allows BVI audiences to locate characters within a limited age range. Also, the fact of hearing the voices throughout the film as opposed to only once (as is usually the case with other traits which are mentioned in the script), possibly contributes to integrate this information in the character mental model and ends up leading to a better recall and recognition. However, there might still be another important reason which could help explain the superior memory for this category: age does not provide strictly visual information in the sense that it allows AD users to infer further details, not limited to the visual ones, but expandable to more semantic domains. By knowing the age of characters, the BVI can activate their prior knowledge, access valuable information associated with that particular age range and activate a number of expectations in their attempt to understand film characters. Perhaps this evocation of a more semantic meaning, rather than a purely visual one, could also have a positive impact upon users' memory.

Pairwise comparisons also indicated that participants recognized the category 'height and weight' with considerably more efficiency than the category 'facial features' and, as was the case with 'age', the analyses highlighted that significantly less errors and 'no recognitions' were ascribed to 'height and weight' as opposed to the rest of the categories explored. A possible interpretation of these results might have to do with the fact that height and weight provide the most basic information about characters, meaning that those are the two traits which help picture the most schematic image of them, and those which provide BVI audiences with the minimal information necessary to outline a simple sketch. It is convenient to mention at this stage that, even

though we did not take into account for our analysis the participants' preferences, when asked generally about which traits of the characters they preferred the AD to deliver in those cases where time constraints applied, almost 75 per cent of the participants in our test answered height and weight, 15 per cent mentioned the age, around 5 per cent preferred to know about their clothing and other items, and 5 per cent mentioned that they appreciated details on the way characters looked at each other because that kind of information (more related to the characters' actions than to their physical description) allowed them to understand the characters' psychology. It is interesting to note that even though 'age' was the category which boasted the best free recall and recognition, it was not what users preferred. Indeed, participants pointed out that they could guess an approximate age of the characters through their voices and, therefore, they found 'height and weight' to be of more use within the script.

Finally, pairwise comparisons showed no statistical differences in the recall or recognition of the categories 'hair', 'facial features' and 'clothes and other items.' Three figures are worth noting: free recall of the hair did not surpass 50 per cent in the best condition of the experiment (x 2+). Regarding facial features, free recall was below 40 per cent, and as per clothing and other relevant items, the highest scores of free recall were around 40 per cent. A higher performance was observed for recognition, even though the amount of errors and 'no recognition' was consistently around 30-35 per cent for each of those categories in the best condition of the test. Despite the fact that participants had difficulties in recalling and recognizing these specific traits efficiently, they frequently remembered the general idea to which they pointed. For instance, when asked about Richard, one of the characters in PMS, 60 per cent of the BVI in our sample who had not recalled his blue shirt or grey trousers, remembered that he was 'well-dressed' or 'elegant.' Similarly (and also more strikingly) most of the participants remembered that Flynn, in BB1, 'had some kind of physical or mental problem' and

that Frank, also in PMS, ‘suffered from a deep depression’, but less of them recalled or recognized Flynn’s crutches and Frank’s bandaged wrists. In the case of these two characters, this is surprising because both the crutches and the bandaged wrists were very distinctive features which could be expected to be properly recalled and recognized. However, even though their recall and recognition was superior to the average, it was still far from ideal (that is, free recall of Flynn’s crutches was only around 65 per cent). The fact that the participants did not recall the specific details but were able to mention general ideas associated to them could be interpreted as an attempt to extract semantic meaning from the visual information. BVI audiences could be using the visual descriptions as the ground from which to create more complex mental models of characters, which might rely not only on visual features but, especially, on elaborations of those that serve as indications of more meaningful information: the characters’ social, professional and economic status, as well as their mental states.

As per RQ2, results showed that both free recall and recognition were consistently better when information was segmented and divided into two blocks which were presented at different stages of each clip. This might be due to the fact that providing longer descriptions imposes more cognitive load on the AD user, who needs to handle more information within working memory at a time. Since working memory is a capacity-limited system, offering short descriptions of characters including less categories of physical features may help addressees remember them. These results are consistent with Wong et al. (2012), who, under the light of the Cognitive Load Theory (Sweller, 1988, 1989; Chandler and Sweller, 1991), suggested that long segments of transient information demand more cognitive resources to receivers than shorter segments. This is also aligned with Fresno et al. (2014), who found that more information was remembered by BVI audiences when segmented descriptions were provided. These findings together could be interpreted as an indication that segmenting

the AD of characters favours users' memory both quantitatively and qualitatively, since more and more varied information is remembered and recognized.

Regarding our last research question (RQ3), no differences in the categories of traits recalled or recognized were observed for the primary and secondary characters in our corpus. Even though Fresno et al. (2014) found that more information was recalled and recognized for the most prominent characters in the clips, the fact that the same categories were recalled for main and secondary characters suggests that BVI audiences might follow the same approach in their attempt to understand characters, regardless of their prominence within the filmic plot. That is to say, AD users might put more effort into outlining the mental model of the primary characters, as has been also proposed for readers of written narratives (Schneider, 2001) and for film spectators (Magliano, Taylor and Kim, 2005), but they might take into account the same categories during the process of creating and updating the mental model of all characters. It might also be interesting to note that an interaction was found in our analysis between the prominence of the characters and the categories explored: specifically, the category 'hair' was statistically better recognized for the main characters. This could perhaps suggest that when characters are perceived as more important to the plot, more cognitive effort is devoted to the creation and updates of their mental model, and more attention is paid to the purely visual features, such as the hair.

Taken together, the aforementioned results seem to confirm our hypotheses that some categories of traits are recalled and recognized more efficiently than others, and that segmenting the descriptions of characters leads to the recall and recognition of a greater variety of information. Nevertheless, they refute our hypothesis that more categories of physical traits would be recalled and recognized for the most prominent characters in the plot. Even though they provide some interesting insights, these results should be understood as preliminary since our study had some limitations, especially

related to the materials used and to the size of the sample. For our test, brief audio clips were used as opposed to complete films, which could have an impact on its reception. Probably, the cognitive load imposed on users varies from a short lab test lasting about ten minutes to a real filmic context lasting around one-and-a-half-hours, and this might influence what addressees remember about characters after each experience. In addition, since we needed our results to be as comparable as possible, the four clips selected as the basis for our corpus were very similar in terms of genre, dialogue density and narrative complexity, which could also have obvious effects on their reception. These three elements are the ultimate reason why certain films become easier to understand than others, since they determine its intrinsic cognitive load (Sweller, 1988, 1989; Chandler and Sweller, 1991) by posing more or less demands on the audience. As per the size of our sample, we had to work with a limited number of participants because, as has already been pointed out in previous experimental research (Chmiel and Mazur, 2012; Cabeza-Cáceres, 2013; Ramos, 2013), finding a substantial amount of BVI AD users willing to take part in these kinds of projects is complicated. Finally, since a naturalistic environment was sought for the test, the age of the participants was not restricted and we worked with a sample of subjects aged 18 to 76. Their results were analysed as a whole, considering them as representatives of all potential AD users. However, age is a variable which might have an effect on memory for films and, hence, conducting further research in which participants are classified in different age ranges and their results are analysed accordingly might yield very interesting findings that could complement and expand our current results.

Despite these limitations, some interesting implications might be drawn from the results obtained in this experiment. First of all, the fact that not all categories of physical features are recalled and recognized with the same efficiency allows for some tentative criteria to prioritize information in the AD scripts, at least for those ‘regular’ characters

which are not prototypical or strongly marked. The differences observed in the nature of the information remembered by the participants in our test suggest that certain categories of traits play a more important role in the reception process. Therefore, when time constraints apply, audio describers might want to prioritize in their scripts the information which is more frequently remembered by the BVI audiences (age, height and weight), so as to provide them with the details that seem more relevant in cognitive terms.

In contrast, both the recall and recognition of the remaining categories of traits explored was far from ideal. The specific details were poorly remembered, but general abstractions of those were more frequently observed. Through this strategy, BVI addressees seem to try to trespass the merely visual descriptions and extract their semantic meaning in an unconscious inferential process which might have some cognitive cost. It is our belief that a closer monitoring of this inferential process would be highly advisable because it could approach the reception and comprehension of audio described products from an absolutely unexplored perspective. If BVI audiences are not interested in the visual details *per se* and tend to use them as a scaffold upon which to ascribe semantic meaning, providing addressees with ‘semantic ADs’ instead of ‘visual ADs’ could be beneficial in reducing their cognitive effort. If further research confirms our tentative results, a new direction in the AD of characters could be contemplated, in which the degree of semantic or visual elements in the descriptions could be modulated according to two criteria: the prominence of the character within the plot and the intrinsic cognitive load of the film. For those secondary characters with no relevance in the story, descriptions could perhaps be closer to the more visual end of the gradation, since no important inferences would be generated for them. However, in the case of those characters with more weight in the film, ADs could be more semantic so as to favour users’ memory. For instance, Marie, a secondary character in BB1 with very

limited prominence in the clip, could be described as a ‘40-year-old, tall and slim woman’, a pretty basic and cognitively undemanding description including the three best recalled and recognized categories according to our test. In contrast, Richard, a primary character in PMS, is described in English as follows: ‘a neat-looking man in his late thirties.’ This is a perfect example of a very brief and semantic description, in which no specific details of his physical traits or clothing are mentioned. Of course, if time restrictions did not apply, more information could be added to his AD: ‘a neat-looking man in his late thirties wearing a suit.’ Through these kinds of semantically-oriented ADs, users would receive shorter descriptions, which would pose less cognitive demands on them, but which could nevertheless help them imagine the characters; that, eventually, could lead BVI audiences to a better recall and recognition of film characters.

Following the same logic, the more difficult the plot, the more helpful it might be to create semantic ADs. As the reader will have noted, this proposal is not aligned to the traditional conception of AD, which considers that only visual information should be described in order to allow users to make inferences themselves. However, it is not far from the lines of research exploring the narratology of AD (that is, Kruger’s (2010) distinction between audio narration and AD). In the end, it consists in creating descriptions which, instead of delivering purely visual details, transmit the narrative effect of those in order to contribute to a more cognitively-efficient understanding of the filmic plot.

Focusing now on the presentation of the AD and taking into account the results obtained in Fresno et al. (2014), segmentation has proven a valid strategy to favour users’ memory, both quantitatively and qualitatively. More and more varied information is remembered if descriptions are divided into shorter units, hence the current practices that provide complete descriptions of characters the first time they appear on screen

should be reconsidered. Under the light of our results, a more convenient approach would be to provide users with short ‘bites’ of information at different stages of the script, even if this means delivering part of the description later in the film. This would possibly imply sacrificing immediacy, but it would increase the likeliness of remembering character descriptions. Furthermore, the fact that more information is remembered for primary than for secondary characters, but that no differences are found in the categories recalled and recognized for them highlights that BVI approach them similarly in their search for film comprehension. Therefore, criteria to prioritize information in scripts and techniques aimed at enhancing users’ memory for film characters (such as information segmentation) might be applied indistinctively to the AD of all characters, regardless of their prominence within the plot.

4. Conclusions

Our quantitative and qualitative research exploring the recall and recognition of film characters seems to highlight two ideas: the fact that less is more and the primacy of the semantic over the visual. These findings are still tentative but, if confirmed, they could have important implications for AD since they could point to new directions in professional practice. We are still at the beginning of a long road and need further empirical research to broaden our preliminary findings. For instance, in relation to segmentation, it would be convenient to find out whether the benefits observed when descriptions are segmented vary depending on the intrinsic cognitive load of the materials used as a corpus. Perhaps segmentation is not as effective with very simple audiovisual products as it is when narratively complex stories are described.

It is also proposed in this paper that characters with more prominence in the plot be described in a more semantic manner, whereas visual descriptions be left for those

secondary characters who do not lead to important inferences. Future research could explore the effect of semantic ADs as opposed to visual ADs of characters in films with different degrees of narrative complexity. Assessing free recall and recognition in different filmic contexts, and adding the users' preferences to the equation could help provide a wider picture of character reception by BVI audiences.

To conclude, extensive research is still needed to delve into the ideas outlined in this paper and, in general, to deepen our understanding of the reception of audio described products. Approaching our discipline from a cognitive perspective could help identify the real needs of AD users, work towards meeting them and, eventually, produce scripts that allow BVI audiences to comprehend and enjoy audio described films.

Acknowledgements

This research has been supported by Spanish Ministry of Economy and Competitiveness (project FFI-2012-31024) and Catalan government (2014SGR0027) funds.

References

- Amedi, A., Raz, N., Pianka, P., Malach, R., and Zohary, E. (2003) 'Early 'Visual' Cortex Activation Correlates with Superior Verbal Memory Performance in the Blind', *Nature Neuroscience*, 6(7): 758-766.
- Arma, S. (2012) 'Why Can't You Wear Black Shoes Like the Other Mothers? Preliminary Investigation on the Italian Language of Audio Description' in E. Perego (ed.) *Emerging Topics in Translation: Audio Description* (Trieste: EUT Edizioni Università di Trieste).

- Asociación Española de Normalización y Certificación (2005) *Norma UNE 153020: Audiodescripción para personas con discapacidad visual. Requisitos para la audiodescripción y elaboración de audioguías* (Madrid: AENOR).
- Basil, M. D. (1992) 'Attention to and Memory for Audio and Video Information in Television Scenes'. Paper presented at the International Communication Association, Miami, May.
- Bourne, J. and Jiménez, 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 J. Díaz Cintas, P. Orero and A. Remael (eds.) *Media for All: Subtitling for the Deaf, Audio Description and Sign Language* (Amsterdam: Rodopi).
- Brady, T. F., Konkle, T., Álvarez, G. A. and Oliva, A. (2008) 'Visual Long-Term Memory has a Massive Storage Capacity for Object Details', *Proceedings of the National Academy of Science of the United States of America*, 105(38): 14325-14329.
- Braun, S. (2007) 'Audio Description from a Discourse Perspective: A Socially Relevant Framework for Research and Training', *Linguistica Antverpiensia*, 6: 357-369.
- Braun, S. (2011) 'Creating Coherence in Audio Description', *Meta: Journal des Traducteurs/ Meta: Translator's Journal*, 56(3): 645-662.
- Broadcasting Authority of Ireland (BAI) (2012) *Guidelines on Audio Description*, https://ncbi.ie/files/BAI_Access_Rules_Sub_Jan12.doc, date accessed 7 May 2015.
- Butcher, K. R. (2014) 'Multimedia Principle', in R. E. Mayer (ed.) *The Cambridge Handbook of Multimedia Learning*, Second edition (New York: Cambridge University Press).

- Cabeza-Cáceres, C. (2013) *Audiodescripció i recepció. Efecte de la velocitat de narració, l'entonació i l'explicitació en la comprensió fílmica*. Published PhD, <http://www.tdx.cat/handle/10803/113556>, date accessed 7 May 2015.
- Chandler, P. and Sweller, J. (1991) 'Cognitive Load Theory and the Format of Instruction', *Cognition and Instruction*, 8(4): 293-332.
- Chmiel, A. and Mazur, I. (2012) 'AD Reception Research: Some Methodological Considerations', In E. Perego (ed.) *Emerging Topics in Translation: Audio Description* (Trieste: EUT Edizioni Università di Trieste).
- Cohen, M. A., Horowitz, T. S. and Wolfe, J. A. (2009) 'Auditory Recognition Memory is Inferior to Visual Recognition Memory', *Proceedings of the National Academy of Science of the United States of America*, 106(14): 6008-6010.
- Eilam, B. and Poyas, Y. (2008) 'Learning with Multiple Representations: Extending Multimedia Learning Beyond the Lab', *Learning and Instruction*, 18(4): 368-378.
- Fox, J. (2004) 'A Signal Detection Analysis of Audio/Video Redundancy Effects in Television News Video', *Communication Research*, 31: 524-536.
- Fresno, N. (forthcoming) 'Carving Characters in the Mind. A Theoretical Approach to the Reception of Characters in Audio Described Films', *Hermēneus. Revista de Traducción e Interpretación*, 18.
- Fresno, N., Castellà, J. and Soler Vilageliu, O. (2014) 'Less is More. Effects of the Amount of Information and its Presentation in the Recall and Reception of Audio Described Characters', *International Journal of Sciences: Basic and Applied Research*, 14(2): 169-196.
- Friendly, D. T. (Producer) and Dayton, J. and Faris, V. (Directors) (2006) *Pequeña Miss Sunshine* [Motion picture] (United States: Fox Searchlight Pictures).

- Fryer L. and Freeman, J. (2012) 'Presence of Those With and Without Sight: Audio Description and its Potential for Virtual Reality Applications', *Journal of CyberTherapy & Rehabilitation*, 5(1): 15-23.
- Fryer L. and Freeman, J. (2013) 'Cinematic Language and the Description of Film: Keeping AD Users in the Frame', *Perspectives: Studies in Translatology*, 21(3): 412-426.
- Fryer L. and Romero-Fresco, P. (2014) 'Audio introductions' in A. Maszerowska, A. Matamala and P. Orero (eds.) *Audio Description: New Perspectives Illustrated* (Amsterdam: Benjamins).
- Gatiss, M., Moffat, S. and Vertue, B. (Producers) (2010) *Sherlock* [Television series] (United Kingdom: Hartswood Films).
- Gilligan, V. (Writer) and McKay, J. (Director) (2008) 'Cancer man' [Television series episode] in *Breaking Bad* (United States: Gran Via Productions, High Bridge Productions and Sony Pictures Television).
- Gilligan V. (Writer) and Cranston, B. (Director) (2010) 'No más' [Television series episode] in *Breaking Bad* (United States: Gran Via Productions, High Bridge Productions and Sony Pictures Television).
- Gould, P. (Writer) and Bernstein, A. (Director) (2010) 'Caballo sin nombre' [Television series episode] in *Breaking Bad* (United States: Gran Via Productions, High Bridge Productions and Sony Pictures Television).
- Graber, D. A. (1990) 'Seeing is Remembering: How Visuals Contribute to Learning from Television News', *Journal of Communication*, 40: 134-155.
- Kruger, J-L. (2010) 'Audio Narration: Re-narrativising Film', *Perspectives: Studies in Translatology*, 18(3): 231- 249.
- Lang, A. (1995) 'Defining Audio/Video Redundancy: From a Limited-Capacity Information Processing Perspective', *Communication Research*, 22(1): 86-115.

- Magliano, J. P., Taylor, H. A. and Kim, H-J. J. (2005) 'When Goals Collide: Monitoring the Goals of Multiple Characters', *Memory & Cognition*, 33(8): 1357-1367.
- Martín Cuenca, M., Bovaira, F., De Santiago, S. and Hernández, A. (Producers) and Martín Cuenca, M. (Director) (2013) *Caníbal* [Motion picture] (Spain: Golem).
- Maszerowska, A. (2012) 'Casting the Light on Cinema - How Luminance and Contrast Patterns Create Meaning', *Meta: Journal des Traducteurs/ Meta: Translator's Journal*, 4: 65-85.
- Maszerowska, A. (2013) 'Language Without Words: Light and Contrast in Audio Description', *The Journal of Specialised Translation*, 20: 165-180.
- Matamala, A., Fernández, A. and Ortiz-Boix, C. (2013) 'Enhancing Sensorial and Linguistic Accessibility with Technology: Further Developments in the TECNACC and ALST projects'. Paper presented at the 5th International Conference Media for All. Audiovisual Translation: Expanding Borders. Dubrovnik, September.
- Matamala, A. and Rami, N. (2009) 'Comparative Analysis of Spanish and German Audio Description of "Good-bye, Lenin!"', *Hermēneus. Revista de Traducción e Interpretación*, 11: 249-266.
- Mayer, R. E. (2001) *Multimedia Learning* (Cambridge: Cambridge University Press).
- Mazur, I. and Kruger, J.-L. (eds.) (2012) 'Pear Stories and Audio Description: Language, Perception and Cognition across Cultures' [Special issue], *Perspectives: Studies in Translatology*, 20(1).
- Nolan, C., Roven, C. and Thomas, E. (Producers) and Nolan, C. (Director) (2008) *The Dark Knight* [Motion picture] (United States: Legendary Pictures).
- Ofcom (2012) *Guidelines on the Provision of Television Access Services*,

http://stakeholders.ofcom.org.uk/broadcasting/guidance/other-guidance/tv_access_serv/guidelines/, date accessed 7 May 2015.

- Orero, P. and Vilaró, A. (2012) 'Eye Tracking Analysis of Minor Details in Films for Audio Description', *Meta: Journal des Traducteurs/ Meta: Translator's Journal*, 4: 295-319.
- Orero, P. and Vilaró, A. (2014) 'Secondary Elements in Audio Description' in A. Maszerowska, A. Matamala and P. Orero (eds.) *Audio Description: New Perspectives Illustrated* (Amsterdam: Benjamins).
- Paivio, A. (1986) *Mental Representations: A Dual Coding Approach* (New York: Oxford University Press).
- Pérez Payá, M. (2007) 'La audiodescripción: traduciendo el lenguaje de las cámaras' in C. Jiménez (ed.) *Traducción y accesibilidad. Subtitulación para sordos y audiodescripción para ciegos: Nuevas modalidades de traducción audiovisual* (Frankfurt: Peter Lang).
- Piety, P. J. (2004) 'The Language System of Audio Description: An Investigation as a Discursive Process', *Journal of Visual Impairment & Blindness*, 98(8): 453- 468.
- Rai, S. (2009) *Bollywood for All: The Demand for Audio Described Bollywood Films* (London: Royal National Institute of Blind People (RNIB)).
- Ramos, M. (2013) *El impacto emocional de la audiodescripción*. Published PhD, <http://www.tdx.cat/handle/10803/124169>, date accessed 7 May 2015.
- Ramos, M. (2015) 'The Emotional Experience of Films: Does Audio Description Make a Difference?', *The Translator*, 21(1), 68-94.
- Remael, A. (2012) 'For the Use of Sound. Film Sound Analysis for Audio-Description: Some Key Issues', *Meta: Journal des Traducteurs/ Meta: Translator's Journal*, 4: 255- 276.

- Röder, B., Rösler, F. and Neville, H. J. (2001) 'Auditory Memory in Congenitally Blind Adults: A Behavioral Electrophysiological Investigation', *Cognitive Brain Research*, 11: 289-303.
- Romero-Fresco, P. and Fryer, L. (2013) 'Could Audio-Described Films Benefit from Audio Introductions? An Audience Response Study', *Journal of Visual Impairment & Blindness*, 107(4): 287-295.
- Schneider, R. (2001) 'Toward a Cognitive Theory of Literary Character: The Dynamics of Mental-Model Construction', *Style*, 35(4): 607-640.
- Shepard, R. N. (1967) 'Recognition Memory for Words, Sentences, and Pictures', *Journal of Verbal Learning and Verbal Behavior*, 6: 156-163.
- Standing, L. (1973) 'Learning 10,000 Pictures', *Quarterly Journal of Experimental Psychology*, 25: 207-222.
- Sweller, J. (1988) 'Cognitive Load during Problem Solving: Effects on Learning', *Cognitive Science*, 12: 257-285.
- Sweller, J. (1989) 'Cognitive Technology: Some Procedures for Facilitating Learning and Problem Solving in Mathematics and Science', *Journal of Educational Psychology*, 81: 457-466.
- Szarkowska, A. (2011) 'Text-to-Speech Audio Description. Towards Wider Availability of AD', *The Journal of Specialised Translation*, 15: 142-162.
- Szarkowska, A. and Jankowska, A. (2012) 'Text-to-Speech Audio Description of Voiced-Over Films. A Case Study of Audio Described *Volver* in Polish' in E. Perego (ed.) *Emerging Topics in Translation: Audio Description* (Trieste: EUT Edizioni Università di Trieste).
- The Independent Television Commission (2000) *ITC Guidance on Standards for Audio Description*,

http://stakeholders.ofcom.org.uk/binaries/broadcast/guidance/itcguide_sds_audio_desc_word.doc, date accessed 7 May 2015.

- Vercauteren, G. (2012) 'A Narratological Approach to Content Selection in Audio Description. Towards a Strategy for the Description of Narratological Time', *MonTI*, 4: 207-230.
- Vilaró, A. and Orero, P. (2013) 'Leitmotif in Audio Description: Anchoring Information to Optimise Retrieval', *International Journal of Humanities and Social Science*, 3(5): 56-64.
- Wong, A., Leahy, W., Marcus, N. and Sweller, J. (2012) 'Cognitive Load Theory, the Transient Information Effect and e-Learning', *Learning and Instruction*, 22: 449-457.

¹ 'All my life I have been keen on cinema and when my doctor told me I would become blind, I thought that would be one of the things I would miss the most: watching films. However, thanks to audio description, I can still enjoy cinema. Differently than before, but I still enjoy it' [Our translation].

² 'cause fatigue to the visually impaired listener' [Our translation].