APPROACHING THE AUDIO DESCRIPTION OF HUMOUR*

Juan José Martínez-Sierra
Universidad de Murcia

ABSTRACT

The main purpose of this article is to provide an introduction to audio description (AD) and to approach the question of the AD of humour. AD is defined, the current situation of this practice is briefly discussed, and some of the most significant works on AD are mentioned. A case study is then presented, the main objective of which was to analyse (from a descriptive perspective and focusing on visual jokes) the AD of the British comedy film I want Candy (several examples are given to illustrate the way humour was dealt with by the AD agent). Among others, some of the main findings discussed are that 1) although some marginal and punctual instances of description over dialogue were found, the general rule of using the gaps or silences between dialogues to insert the descriptions was observed (this sometimes meant losing potentially humorous elements) and 2) almost two fifths of the visual humorous fragments were not described, most likely due to time restrictions. Finally, some further research is suggested.

KEYWORDS: media accessibility; audio description; humour; audiovisual translation

RESUMEN

Este artículo tiene como propósito principal ofrecer una introducción a la audiodescripción (AD) y realizar una aproximación a la cuestión de la audiodescripción del humor. En primer lugar se incluye una definición de la AD, a la vez que se comenta brevemente la situación actual de esta práctica y se mencionan algunos de los trabajos más representativos en este campo. A continuación se presenta un estudio de caso cuyo principal objetivo consistió en analizar (desde una perspectiva descriptiva y centrándola la atención en los chistes visuales) la AD de la comedia británica I want Candy (se muestran varios ejemplos con objeto de ilustrar cómo se manejó el humor). Algunos de los resultados obtenidos son: 1) si bien se detectaron ciertos momentos puntuales y marginales en los que la descripción se superpuso al diálogo, en términos

* This piece of research was conducted as a part of a research project founded by the Spanish Ministerio de Ciencia e Innovación (Programa José Castillejo).
generales se respetó la regla de usar los espacios o silencios entre diálogos para incluir la descripción (lo que en ocasiones supuso la pérdida de elementos potencialmente humorísticos) y 2) casi dos quintas partes de los fragmentos visuales potencialmente humorísticos no se describieron, probablemente debido a las restricciones temporales. Por último se proponen algunas ideas de investigación futura.

PALABRAS CLAVE: accesibilidad; audiodescripción; humor; traducción audiovisual

1. DEFINITION AND CURRENT SITUATION OF AD

In general terms, AD is a practice the purpose of which is to help the blind and those with some sort of visual impairment have access to audiovisual material such as films, television programmes, and even theatre or opera. Basically, what AD does is to use the absence of dialogue to describe verbally every visual or acoustic detail that is considered relevant. Benecke (2004: 78) defines it as

the technique used for making theatre, movies and TV programmes accessible to blind and visually impaired people: an additional narration describes the action, body language, facial expressions, scenery and costumes. The description fits in between the dialogue and does not interfere with important sound and music effects.

This definition is in agreement with more recent ones, such as that provided by Díaz-Cintas. According to him, “AD consists in transforming visual images into words, which are then spoken during the silent intervals of audiovisual programmes or live performances” (2008: 7).

In both cases, it can be noticed that no allusion to change of language is made, in the sense of translation from a source language to a target language. Still, as Díaz-Cintas points out, audiovisual translation (AVT) scholars and practitioners seem to agree that modes such as subtitling for the deaf and the hard-of-hearing (SDH) and AD “are an integral part of AVT” (2008: 7).

In my opinion, and assuming a wide, flexible conception of what translating means, a priori it seems possible to find some similarities and differences between AD and other AVT modes: for example, 1) both the translator for dubbing or subtitling and the describer\(^1\) perform a translation

\(^1\) The debate narrator / describer lies beyond the scope of this article.
2. HUMOUR IN AUDIOVISUAL TEXTS

Humour is one of the most difficult aspects to handle in the different modes of audiovisual translation. Hence, we could assume it is also so in AD. The restrictions in practices such as dubbing or subtitling are many, and so they seem in AD. Both technical and cultural factors can affect the transfer of humour. For example, in the case of AD, the time to provide the verbal description can be short. Besides, and especially considering the possibility of interlinguistic AD – in the sense that the narration has been translated into

---

2 As Snyder (2008: 196) explains, “Throughout Europe, AD is considered a form of audiovisual translation – a way to translate information that is perceptible in one sense (visual) to a form that is comparably accessible with another (aural).” In similar terms, Bourne and Jiménez-Hurtado (2007: 175) claim that, “from a semiotic point of view, AD may be viewed as the translation of images to words and is therefore a type of intersemiotic translation.”

3 See Agost (1999) for further details.

4 Díaz-Cintas (2008: 7-8) lists a series of volumes, articles, papers, panels, modules, conferences, and the like devoted to SDH and AD.
another language –⁵ the cultural specificity of the humorous segment to be described can seriously complicate the task.

The study of the translation of humour in audiovisual texts has been undertaken by different authors.⁶ On this occasion, I shall allude to Martínez-Sierra (2008) who, broadly speaking, identifies three different ways in which humour can be created in this type of texts. He distinguishes between 1) humour created via language, 2) humour created via sounds, and 3) humour created via images.⁷ The examples provided will show elements of these three categories, although it is on the third type – humour created via images – that I wish to focus this time.

Apart from providing an introduction to AD, the main purpose of this study is to approach the question of the intralinguistic AD – in the sense that the narration has not been translated into another language – of humour. In order to do so, the source, audio described version of the British comedy film I want Candy (Stephen Surjik 2007) was used.⁸ I am aware that this case study is necessarily fractional, since only the AD narration has been considered, purposely ignoring – on this occasion – the dialogues.

3. STRUCTURE OF THE ANALYSIS

In order to structure the analysis, the guidelines given by Vercauteren (2007: 142-147) were considered. He asks four questions:

   a) What should be described?
   b) When should it be described?
   c) How should it be described?
   d) How much should be described?

Let us see the answers that he gives to these four questions.

⁵ As opposed to intralinguistic (see Bourne and Jiménez-Hurtado 2007). It seems rather safe to predict that in a not too distant future importing audio described films for dubbing – and hence for translating the narration – will be usual in countries such as Spain. Additionally, as Díaz-Cintas (2008: 8) explains, “a new type of AD is also taking shape with the development of audio subtitling in countries where most of the programmes are broadcast or shown in a foreign language.”


⁷ For further details and more specific audiovisual humorous elements, see Martínez-Sierra (2008).

⁸ Although all considerations in this work are about comedy films, they can be clearly applicable to other audiovisual genres such as television sitcoms, animated shows, or even on stage performances.
WHAT SHOULD BE DESCRIBED?

According to Vercauteren, the AD agent should describe:

1) Images: where, when, what, who, and how
2) Sounds: “sound effects of the programme (that are difficult to identify), song lyrics and languages used other than the source language of the programme”
3) On-screen texts: “logos, opening titles, cast list, credits and text on signs that might be shown on-screen as subtitles.” According to Chaume (2004: 25), written language can also be perceived visually. In Martínez-Sierra (2004 and 2008), inserts are conceived as possible carriers of potentially humorous graphic elements. Regardless of their label (on-screen texts, written language on screen, or inserts), it is in this last sense that these elements will be understood here.

What is noteworthy about Vercauteren’s proposal is that, unlike other authors who stick to the visual dimension, he makes explicit reference to the AD of sound.

WHEN SHOULD IT BE DESCRIBED?

There seems to be a general rule: “descriptions are inserted in the gaps or silences between dialogues.”

HOW SHOULD IT BE DESCRIBED?

Vercauteren makes some suggestions in this respect:

• Use clear, precise, appropriate, and varied language
• Sound natural and avoid unusual vocabulary
• Use simple sentences
• Be objective
• Try to match the style of the film and the audience

HOW MUCH SHOULD BE DESCRIBED?

Vercauteren acknowledges that further research is needed to give a precise answer to this question. In any case, a starting point will be to describe everything that is considered relevant.
4. HYPOTHESIS

The main hypothesis of this study was that, to keep the humorous element, most of the description work would be done with the images, since dialogues – language – and music, noises, etc. – sounds – can be heard. For my purposes, on-screen texts were also taken into account, given that, due to their visual nature, they need to be read or described. Thus, it was considered that images – including on-screen texts – would call for a description so that the humorous elements could be kept.

5. ANALYSIS

As mentioned before, the study was conducted using the film *I Want Candy*, the plot of which can be summarized as follows: two film students – Joe and Baggy – are desperate to make their first motion picture, for which they have just completed a script. In their search for their big break, they contact Doug, a producer who agrees to finance their project only if they convince porn star Candy Fiveways to work in the film, thus turning their initial love story into an adult entertainment movie.

Let us see some examples of the different possibilities that were identified in the analysis of the film. As already pointed out, the examples include elements from the three general groups already referred to (language, sounds, and images). Since we are considering the source, audio described version of the film, language-related humorous elements will only appear for the sake of contextualisation.

The examples are placed in a table with four columns: 1) the joke, 2) the type of joke, 3) whether it was described and how (for obvious reasons, this will not apply to the dialogues), and 4) a brief comment.

**EXAMPLE 1 (TCR 00:00:40 – 00:01:43)**
Scene: At a cemetery (Joe and Baggy are taping a funeral video. Baggy operates the camera).

---

9 The list of examples provided is not meant to be comprehensive and should just be considered as an illustration of the different possibilities found in the film.

10 Typically, they would be crucial in a study on the translation of this film for dubbing or subtitling.
<table>
<thead>
<tr>
<th>Joke</th>
<th>Type</th>
<th>AD?</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joe flirts with a blonde girl</td>
<td>Image</td>
<td>No</td>
<td>The joke is lost</td>
</tr>
<tr>
<td>Priest: Look, I’m sorry. This won’t do.</td>
<td>Language</td>
<td>Not applicable</td>
<td>Not applicable</td>
</tr>
<tr>
<td>This family is grieving. You need to show some more respect. Joe: This family has paid for a high-quality funeral video. A lasting memory of a fond farewell</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When the priest rebukes Joe and Baggy for their behaviour, Joe shows him a business card: “LEATHERHEAD FUNERAL VIDEOS. Making A Drama Out Of A Tragedy. John Bagley. 07700 90003”</td>
<td>On-screen text</td>
<td>Describer: A card for Letterhead Funeral Videos</td>
<td>The joke is lost, since we are told about the card (informative purpose) but not about its content (humorous purpose), and that is where the humour is</td>
</tr>
<tr>
<td>Baggy [to the priest]: I got you as far as, “I always used to look forward to the visits”</td>
<td>Language</td>
<td>Not applicable</td>
<td>Not applicable</td>
</tr>
<tr>
<td>Joe decides that they have enough material and that it is time to leave. To the people: Good job. Thanks, everyone, for your patience. [He claps his hands loudly and gives them a thumb-up]</td>
<td>Sound / Image</td>
<td>No</td>
<td>The sound humorous element is kept, since no description seems to be needed (it is easy to identify the sound of clapping). However, the visual</td>
</tr>
</tbody>
</table>
Since Baggy insists on taping the crying widow, her son punches him in the face.

Example 2 (TCR 00:01:57 – 00:02:00)
Scene: On campus (Joe parks his car and approaches Baggy, who is walking along the sidewalk).

<table>
<thead>
<tr>
<th>Joke</th>
<th>Type</th>
<th>AD?</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baggy walks along the sidewalk. He looks unsmiling and has a black eye</td>
<td>Image</td>
<td>Descriptor: Baggy has a black eye</td>
<td>The joke is kept</td>
</tr>
</tbody>
</table>

Example 3 (TCR 00:01:57 – 00:02:00)
Scene: In the street (as they walk, Joe tries to cheer up Baggy, who is feeling downhearted).

<table>
<thead>
<tr>
<th>Joke</th>
<th>Type</th>
<th>AD?</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joe: Do you know how long it took Attenborough to make Ghandi? It took him ten, long, painful years and we’ve been at it for, like, two and a half hours</td>
<td>Language</td>
<td>Not applicable</td>
<td>Not applicable</td>
</tr>
<tr>
<td>Joe: The good news is we’re not making some long, boring, cry-baby movie about a bald bloke</td>
<td>Language / Image</td>
<td>Not applicable / No</td>
<td>Not applicable / Regarding the visual humorous element, it is</td>
</tr>
</tbody>
</table>
in a nappy. [Baggy stares at him sceptically] Yeah, I know it was a masterpiece, but you get the picture

Joe: Hey, Bags, Bags, Bags, Bags! [He points at a sign on a wall, next to a door. The sign reads “Head Films”] It’s like a sign. Baggy: It is a sign

Example 4 (TCR 00:31:48 – 00:32:20)
Scene: In the garden of Michael de Vere’s mansion (film’s producer Doug tries to give Mr. de Vere an update on the current status of the film in which he has invested a considerable amount of money. Mr. de Vere is holding a jar containing several butterflies and some cotton).

<table>
<thead>
<tr>
<th>Joke</th>
<th>Type</th>
<th>AD?</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mr. de Vere: Many people think the butterfly suffers, but the chloroform dulls its senses and it feels nothing. This way, it won’t struggle when the [he suddenly changes his tone of voice into some sort of sadistic attitude] pin punctures its body</td>
<td>Language (paralinguistic)</td>
<td>Not applicable</td>
<td>Not applicable</td>
</tr>
<tr>
<td>Mr. de Vere: So, how’s my film? Doug: The film? Oh, it’s great.</td>
<td>Language</td>
<td>Not applicable</td>
<td>Not applicable</td>
</tr>
</tbody>
</table>
Yeah. Fully prepped. Friends become enemies and… enemies have sex with each other  

<table>
<thead>
<tr>
<th>Joke</th>
<th>Type</th>
<th>AD?</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>While Mr. de Vere and Doug talk, Robby (Doug’s assistant) tries to catch some butterflies unsuccessfully</td>
<td>Image</td>
<td>No</td>
<td>The joke is lost</td>
</tr>
<tr>
<td>Example 5 (TCR 00:50:39 – 00:50:52)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scene: At Joe’s house (the crew are shooting a scene).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The crew are shooting an erotic scene starring Candy and Christi.</td>
<td>Image</td>
<td>Descriptor: Vlad stares at them wide-eyed and absentmindedly turns the dial on a light dimmer</td>
<td>The joke is kept</td>
</tr>
<tr>
<td>The two actresses are in a bedroom full of light streams across the walls. Vlad, the lighting technician, stares at them</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The lights explode loudly due to an overvoltage. The two scared girls scream and jump</td>
<td>Sound / Image</td>
<td>Descriptor: Suddenly lights pop and explode</td>
<td>The describer seems to focus on the sound (pop) and ignores the girls’ reaction to the explosion (visual)</td>
</tr>
</tbody>
</table>
6. RESULTS

The results of the analysis offer the following figures:

Humour via image (excluding on-screen texts) → 101 potentially humorous moments (or visual jokes):

63 described
38 not described

In the light of these results, it seems clear that the initial hypothesis – that is, that most of the description work would be done with the images – seems confirmed, but only partially, as it is not always possible to describe the images, mainly due to time restrictions. In fact, a noteworthy percentage of the visual jokes is lost (37.6%), since no description is provided.

As far as on-screen texts are concerned, and as already explained, my purpose was to focus on humour via image, but I also paid attention to potentially humorous inserts, due to their visual nature. From a total of six, four were described. In other words, almost 67% of the on-screen humorous texts were described, which is a rather high number.

7. INTERPRETATION OF RESULTS

In order to structure the interpretation of those results, let us go back to Vercauteren’s guidelines.
WHAT SHOULD BE DESCRIBED?

We already saw that Vercauteren says that images, certain sounds, and on-screen texts should be described. In the case of humour, we can be more precise and start looking at the issue from the priorities perspective.\textsuperscript{11} In this sense, the description of humorous images and on-screen texts, on the one hand, and of humorous sounds, on the other, seems to be a priority, since one of the main goals will be to keep the humour – if our purpose (or \textit{skopos}) is to maintain the humorous nature of the audiovisual product.

WHEN SHOULD IT BE DESCRIBED?

We already mentioned the general rule: “descriptions are inserted in the gaps or silences between dialogues” (Vercauteren 2007: 143). This is clearly applicable to the AD of humour, even if it means losing humorous elements. In fact, this rule can become a serious restriction.

HOW SHOULD IT BE DESCRIBED?

Regarding the different suggestions made by Vercauteren as far as the way descriptions should be presented, matching the style of the film and the audience – that is, keeping the film a comedy – seems to be the most relevant one as far as humour is concerned.

HOW MUCH SHOULD BE DESCRIBED?

We already saw that further research seems to be needed and that, for the time being and as some sort of starting parameter, everything that is relevant should be described. From a discursive and functional viewpoint, it seems clear that, in a comedy, humour will be one of the most relevant aspects to be described and therefore maintained.

8. SOME GENERAL CONCLUSIONS

It is possible to draw some general conclusions regarding the AD of humour:

\textsuperscript{11} See Zabalbeascoa’s priorities / restrictions model (1994).
At least in the case of humour, AD is performed mainly with images, although not only. As we saw, authors such as Vercauteren justify the description of sounds when these are not easy to identify. From the point of view of humour, it could be argued that sounds should also be described when humour is based on them and some description is called for not to risk losing it (even if they can be identified).

Keeping the humour can be considered a top priority.

On the other hand, time restrictions will be crucial. The general rule of using the gaps or silences between dialogues to insert the descriptions can entail serious constraints.

In fact, those restrictions can be as strong as to provoke a significant loss of humour (37.6%). However, the high percentage of lost visual jokes should not discourage us from audio describing comedies. Once we adopt a broad conception of a film as an organized whole of interrelated features, the AD of humour proves to be no chimera. In the case of dubbing, for example, all the humorous elements work together and compensate each other, and this same logic seems entirely applicable to the case of AD. After all, the AD audience is exposed not just to a description, but also to sounds and dialogue – that is, to a whole film understood as an audiovisual text. Authors such as Orero and Wharton (2007: 168) suggest that the AD narrative is not part of the film, which makes sense as long as we consider the AD narration as a unit detached from the rest of the elements that compose a film. However, what is implied here is that another view seems possible: the audio described film conceived as an interconnected whole (dialogues, sounds, and description).

9. FURTHER RESEARCH

Quantitatively, AD is in greater demand these days. However, as already mentioned, the aspects of AD waiting to be dealt with are many, humour being one of them. Quantitative and qualitative studies must be conducted so that it becomes a well-defined practice the level of accuracy of which reaches the highest standards. The following are some questions and suggestions for further research on the AD of humour:

12 As Chaves (2000: 59-60) explains, a film is an organized whole in which any operation on any sequence can affect the totality of the text. Similarly, works such as Martínez-Sierra (2008) evidence how humorous elements perform jointly.
Regarding the description of sounds, as we saw authors such as Vercauteren consider that those sounds that are difficult to identify should be described (in the case of humour, especially if they function as humorous elements). Still, a question comes to mind: how can we assess the difficulty of a sound? Does it just depend on sheer common sense? Would it be possible to develop a hierarchical catalogue of such sounds?¹³

It has been said that it is commonly accepted that the AD agent will describe images. It has been also pointed out that authors such as Vercauteren believe that description of sounds can at times be needed. Yet, in cases in which onomatopoeic verbs are used (such as pop; see example 5), to what degree can we say that we are actually describing a humorous image and not a humorous sound?

Authors such as Orero and Wharton (2007: 168) suggest that “the language of AD should be a neutral discourse written to be read aloud and narrated.” But the question is: can humour be neutral? Moreover, they also mention that “The AD narrative is not an intrinsic part of the film, its plot or characters, and should therefore stand distant from it all.” But in those cases in which humour lies in an image, how distant can the description be without losing the humorous factor? Besides, if we accept that 1) the main purpose of a comedy film is to produce laughter as a whole and that 2) part of that laughter will be induced by the visual elements, is it convenient to detach an essential part (the AD of the images) from the audiovisual text and deal with it as an independent message? How can that affect the humorous effectiveness of the text – the film – altogether? Moreover, if AD is to be regarded as part of AVT, should it not conform to the same standards as those of other AVT modes in terms of considering an audio described humorous film as a unit of interrelated features (in short, language, sounds, and audio described images), just as happens in the case of dubbed or subtitled films? Should the answer to this last question be affirmative, it could be useful to develop a taxonomy of humorous elements in AD, just as has been the case in other audiovisual modes such as dubbing.¹⁴ Such a taxonomy could be based on an initial discrimination between:

- Language related elements
- Sound related elements

¹³ As Orero and Wharton claim (referring to the Spanish UNE Standard 153020), “too much is left to personal interpretation and taste” (2007: 167). See also Orero (2005).

¹⁴ See Martínez-Sierra (2008).
Once we had the aforementioned taxonomy, it would be possible to quantify and compare the humorous elements in the AD and non-AD versions of a film, so that further data of how well humour travels in AD could be obtained.

Similarly, reception studies could be conducted to put humorous AD material to the test and to evidence the degree to which such material remains humorous. In this same line, it could be worthwhile to compare the reception of an AD version of a comedy film by an AD audience with the reception of a non-AD version of that same film by a non-AD audience.

These different proposals could be applied to different films and, therefore, to a larger corpus so that it could be possible to identify tendencies or even operational norms in the AD of humour.

10. REFERENCES


VERCAUTEREN, G. (2007): “Toward a European guideline for audio description” in Díaz-Cintas J. et al. (eds.): Media for All: Subtitling for the Deaf, Audio Description and Sign Language. Ámsterdam, Rodopi, 139-149.