Designing a Course on Audio Description and Defining the Main Competences of the Future Professional

Anna Matamala & Pilar Orero
Universitat Autònoma de Barcelona

Considering Audio Description as a type of intersemiotic translation, and being aware of future opportunities in this field once laws are enforced, a course on Audio Description has been developed within the MA in Audiovisual Translation at the Universitat Autònoma de Barcelona (UAB), compelling us to define the main competences a future professional in audio description will have to acquire. In this article, the skills put forward by various academics and professionals are analysed in order to establish the main competences that should be acquired by an audio describer, and sample activities that are used in class are discussed.

1. Audio describers, a new professional profile

Audio Description (AD) for the media has been carried out mainly by volunteers who have acted as social intermediaries in making audiovisual material accessible, transforming images into vivid narration and allowing those with sight problems to gain access to culture. With the increasing percentages of AD that will be implemented either by law or by broadcasters’ goodwill – as is the case for the BBC, Spanish State TVE and Catalan National TV3 –, the need for a new professional profile will arise (Remael & Vercauteren 2007). The Universitat Autònoma de Barcelona (UAB), being aware of this future opportunity and convinced that translators are the best candidates for this specific type of intersemiotic translation, has developed a course on Audio Description within the MA in Audiovisual Translation (MTAV).

In this article, the skills and competences of a good audio describer according to both academics and professionals are presented (§1.1). Then, and after contextualising the course within the MTAV, a selection of competences and a list of learning outcomes for the module are put forward (§2.1), sample activities are shown (§2.2) and some reflections on the best assessment method are included (§2.3).

1.1. Main competences and skills

The training of the person who makes audiovisual material accessible should go beyond academic and professional standards. Those who work with accessibility have to be aware of their roles as social intermediaries since they make available a metaphoric ramp – taking the simile from architectural accessibility – to audiovisual material. As is the case for specialists
in accessibility working in the physical world of architecture, urbanism or transport, the audio describer should have a deep understanding of the world and reality of those who have sensorial difficulties. Bearing this in mind, and before arriving at our own list of skills and abilities, we looked at the many recommendations and definitions provided in previously published literature on the matter.

According to Orero (2005), the skills needed to be an audio describer as stated by the Audio Description Association in conjunction with the Open College Network West and North Yorkshire in the UK are as follows:

- the ability to summarise information accurately and objectively;
- a good command of language;
- a clear and pleasant speaking voice;
- good sight and hearing (whether aided or unaided);
- the ability to work as part of a team;
- the commitment to access for disabled people, and the provision of quality Audio Description to blind and partially sighted people.

Matamala (2006) devotes an article to the linguistic aspects and training challenges in media accessibility and emphasises that, as far as Audio Description is concerned, professionals need various competences which can be summarised in six items:

- the ability to undertake intersemiotic translations (turning images into words), both intralinguistically and interlinguistically, depending on the task commissioned;
- an excellent command of language and conventions;
- the ability to summarise information in order to adapt the text to the limited space available, keeping the original meaning, by means of rewording and by using synonyms;
- the ability to adapt the linguistic style to the target audience and to the product, by mastering different linguistic registers;
- the ability to critically select the most relevant information;
- in live AD, good oral expression and excellent diction.

In the same article—whose conclusions are also portrayed in Ballester et al. (2006)—the Spanish standard AD UNE 153020 (AENOR, 2005) is investigated. The Spanish Standard definition is “el describer, audio-descriptor, o guionista es la persona que escribe el guión de la audiodescripción” [The describer, audio descriptor or scriptwriter is the person who writes the script for the Audio Description] (2005: 5) and the “locutor o narrador”, who is “la persona que realiza la locución del guión de la descripción” (ibid.) [The voice talent is the person who reads the script. (translations PO/AM). It is in the recommendations concerning the writing features of the script where we can deduce the skills which will be needed by the describer: to use adequate vocabulary and precise adjectives, a rich, simple and free-flowing style,
Designing a Course on Audio Description

which should follow normative grammar. And for the voice talent it recommends good vocal skills and highlights that correct and neutral diction is of the essence. Finally, in the specific field of theatre plays, the standard states that audio describers should be familiar with theatre jargon.

Apart from the official standard, some professionals have listed various competences when reflecting on their job. The Spanish audio describer Navarrete (1997a: 71), for example, stresses four items:

- describers should have wide and eclectic cinematographic knowledge;
- good cultural knowledge;
- a great literary ability that allows them to adapt to different styles very quickly, and
- a great sensitivity to audiovisual products.

Moreover, audio describers should have a good knowledge of theatre jargon when describing theatre plays (Navarrete 1997b: 28). The Catalan audio describer Vidal (2004: 31) underlines the ability to summarise information and an excellent command of both synonyms and voice skills as musts. The American describer Snyder (2004) chooses observation, selection of visual information, usage of language and vocal skills as the four pillars of AD, whereas the British audio describer Hyks (2005) agrees with Snyder in stating that describers should have great powers of observation and judgement to decide what should be described, as well as a good command of vocabulary to describe it.

A more extensive review of the competences of audio describers has been put forward by Díaz-Cintas (2007) in an article derived from a report commissioned by the Centro Español de Subtitulado y Audiodescripción (CESyA, www.cesya.es) and the proposals and reactions it raised in its official presentation in Madrid, in November 2006. Díaz-Cintas includes a series of competences under four categories:

- **Linguistic competences:**
  - excellent command of the mother tongue in all its dimensions, especially regarding vocabulary, grammar and syntax;
  - creativity and linguistic sensitivity;
  - ability to edit and revise texts in the mother tongue;
  - knowledge of English.

- **Competences related to the content:**
  - general knowledge of disabilities and accessibility;
  - extensive knowledge of the world of deafness and blindness;
  - knowledge of cinematographic language and the semiotics of the image;
• knowledge of other modalities of media accessibility;
• knowledge of the labour market and laws concerning SDH and AD;
• extensive knowledge of the theory and practice of AD in all its dimensions;
• knowledge of theatre;
• knowledge of the history of the arts.

• Technological or applied competences:
  • computer literacy (general software and Internet);
  • willingness to update and learn new software;
  • ability to research and document for any issue;
  • ability to work with AD software;
  • advanced IT skills;
  • vocal and prosodic skills.

• Personal and general competences:
  • encyclopaedic knowledge;
  • ability to self-learn;
  • ability to analyse, summarise and interpret information;
  • ability to think quickly, to relate ideas and react rapidly;
  • organisation, planning and information management skills;
  • ability to think critically in problem-solving and decision-taking processes;
  • labour flexibility and ability to work under stress and time constraints;
  • willingness to work in a team;
  • ability to act as an expert mediator in multicultural environments.

As far as voice skills are concerned, one should bear in mind that in certain countries – from a dubbing tradition – such as Spain, audio describers do not generally venture into the field of recording, which is traditionally reserved to dubbing actors (Orero 2006, 2007) with powerful unions and also requiring the need of the equivalent of the British Equity card. However, even in such countries, live AD requires a professional with certain voice skills, although not necessarily at the same level as those of a dubbing actor (Matamala 2006, 2007).

Finally, although intralinguistic Audio Description seems to be the norm – i.e. AD created in a certain language for productions in that same language –, a line of investigation, as already pointed out by Orero (2005) and Matamala (2006), is the translation of AD, especially in countries with a large import of audiovisual programmes in foreign languages. Should the translation of ADs become common practice, translation competences would also have to be included in the list. Hyks (2005: 8) considers that “translating and rewording can sometimes take as long if not longer than
starting from scratch” and adds that “the fact that some languages use many more or sometimes fewer words to express an idea, can drastically affect timings and as well as languages not being necessarily comparable, our different cultures are reflected in the AD”. Though this is an issue to be further investigated, the market – as usual – is already practising AD translations. US distributors are now translating English scripts into German. Benecke (2007) explains that this practice may not be widely accepted by the target audience since AD styles may differ from one country or culture to another.

All the previous competences come into play in different contexts which give rise to various types of AD, as suggested by Orero (2006):

- depending on the product: dynamic (e.g., a film) or static (e.g., a picture). Dynamic products can be classified into live products (e.g., theatre plays) and recorded products (TV programmes);
- depending on the broadcasting system: synchronic or pre-recorded;
- depending on the production: live or pre-recorded but simulating a live production;
- depending on the process: production or postproduction;
- depending on the AVT mode: dubbing + AD, audiosubtitling, voice-over + AD;
- depending on the narration: recorded or live.

Matamala (2007) reduces this comprehensive taxonomic classification to a key differentiation between recorded AD and live AD, with three sub-varieties in the latter group: planned ADs, improvised ADs and a third group in between (the audio describer has visualised the product but has not been able to create a written description).

2. Designing a Course on Audio Description

Training people to understand, observe and implement accessibility is the basic objective of most media accessibility courses. The areas of further education in which media and audiovisual accessibility have been implemented consist of courses from various fields (telecommunication engineering, computer science, journalism, teaching, translation) which focus on different aspects depending on their interests (technical issues or content). In Spain, various courses are offered on media accessibility (Badia & Matamala, 2007) but very few prepare students to be audio describers. The module on Audio Description offered within the MA in AVT at the UAB is one of them.

The MTAV at the UAB derives from the postgraduate diploma in AVT created in 2001 which then included different modules such as theoretical
aspects of AVT, subtitling, dubbing, voice-over, multimedia translation, plus a series of seminars (Díaz-Cintas & Orero 2003). In 2003, an online version of this postgraduate course was launched and a year later the MA in its face-to-face format was created, with a wider array of courses. The current version of the MA includes a compulsory module (theoretical aspects of AVT), five basic modules from which students select four (dubbing, subtitling, voice-over, multimedia translation and research project), plus eight optional modules from which students choose three (intralingual subtitling, Catalan for audiovisual translators, introduction to film studies, videogame translation, audiovisual translation from German into Catalan/Spanish, audiovisual translation from French into Catalan/Spanish, audiovisual translation foundation and Audio Description). Audio Description was the last module to be included and was launched in 2006/07, with 19 students.

In the following sections the competences and learning outcomes of this course will be discussed (§2.1) and sample exercises will be presented (§2.2), as well as the assessment methods used (§2.3). We shall not discuss the technological requirements needed since the digital audiovisual material and technologies available are plentiful. There are two challenges we face at universities: the first is related to the speed in which technology is developing and the need to update, the second lies in finding financial resources to meet the latest requirements.

2.1. Competences and learning outcomes

After reviewing the literature and trying to adopt a realistic view of what can be achieved in a 20-hour course, the following competences and skills were prioritised when designing the course:

- excellent command of the mother tongue, especially regarding vocabulary (synonyms) and grammar;
- ability to observe an audiovisual production and select the most critical information;
- ability to summarise information in order to reach synchronisation;
- basic vocal skills;
- ability to adapt the linguistic style to the target audience and to the product, by mastering different linguistic registers;
- ability to translate ADs and synchronise them with the time slots available.

This does not mean that the competences and skills presented in section 1 should be neglected, but some of these can be acquired in other modules within the MA (for example, in modules such as film theory and script writing skills) and others should be a prerequisite for enrolling for this course (knowledge of English, for instance).
Students should learn to apply these competences to all types of ADs presented in the previous section – especially in both live and recorded AD –, and the final learning outcomes of the course would be:

- to know and understand the needs of those who have sight problems and want access to audiovisual media;
- to gain an understanding of the linguistic and cultural issues which are present in AD;
- to be able to identify and exploit the semiotic relationships between original soundtrack and images, which form the visual text when writing an AD;
- to acquire the linguistic skills needed to write an AD script;
- to acquire the technical skills required in all the steps of the process of creating an AD;
- to develop professional criteria to make AD-related stylistic and functional choices based on genre and intended audience;
- to understand and follow conventions and standards of good AD practice;
- to develop specific AD skills in areas such as spotting, editing, location and text compression;
- to develop IT skills to deal with AD software programs;
- to develop specific AD skills in areas such as spotting, editing, location and text compression;
- to be aware of the professional practice and the resources and links for audio describers available on the internet.

2.2. Course Content, Set Up

The course starts by offering an overview on the social aspects of Audio Description:

- intended audience;
- short history of Audio Description;
- guidelines in Spain and other countries;
- audio describer profile.

Though the course is practical, when performing their Audio Description tasks, students are requested to reflect on and argument their decisions. Since we are dealing with a postgraduate course on Audiovisual Translation, the students are also made to reflect on the shifting definitions between Translation Studies concepts and terminology, as applied in Audio Description. This includes issues related to the significance (or lack thereof) of concepts such as:

- fidelity;
- equivalence;
- source and target text;
- synchrony, etc.
Fidelity, for example, should be qualified for Audio Description since the description is not always related to the images which can be seen synchronically. This lack of fidelity is due to many reason, one being the anticipation of an action – a common practice in Audio Description.

Course activities are usually organised in groups of three people working together with one computer. This association is preferred to individual work because

- it helps with the dynamics of observation: not every person focuses on the same issues;
- it helps with relevance: to decide what is important to describe and to help to prioritise;
- it helps with the research, choice of writing style and the syntax;
- from a didactic point of view, it has all the benefits of collaborative learning;
- in some exercises one person can pretend to be blind, which – apart from the possible academic benefits\(^3\) – has been shown to be extremely useful for creating empathy with the possible audience. This is the way Bernd Benecke teaches in his workshops to mirror the way ADs are created in Germany.

This setup has proven to be desirable. Previous experience following a one-student-to-one-computer setup showed poor output of students’ comments since they lacked constant feedback on their production. The simulation of a blind person in a team of three students not only helps to gain insight into the perception of aural information, it also forces the other two students in the process of discrimination of information and description. The perception of information and the choice of what to describe is not standard and universal. Different people will chose different issues as relevant to produce the AD script. If two people work together they will have to reach a consensus on what is relevant and should be audio described.

3. Exercises

Exercises are graded and designed to take students from a superficial awareness of basic descriptive features of everyday objects to the very complex film clip with simultaneous actions and emotions. We start from descriptions of physical objects since it appeared that textures and colours cannot be fully appreciated from a photograph or from an internet image or clip. For transport and manipulation purposes we have focused on small objects which can easily be carried in a bag. We have also taken into consideration a personal affinity to the choice (baby toys, kitchen utensils,
textiles, etc.) but similarly miniature replicas of vintage cars or warhammer figures could be exploited as first objects to be described.

3.1. First Exercises

First of all, we always provide each group of students with a small object that they can handle. First students are given baby rattles in different shapes, forms and textures. Students in these first exercises have to provide long descriptions – compared to the lack of complexity of the form of the object: fifty words which should take into account the size, colour, texture, etc.

Subsequently, students are given a kitchen utensil which is alien to Spanish cookery ware: potato peeler, potato masher, citrus zester, citrus reamer, melon baller, avocado slicer, etc.

After providing the description students are asked to name the object in a maximum of five words. This last request helps students reflect on how objects are defined – by their functions, by their shapes, etc. – and helps them summarise a longer description into a shorter one.

3.2. Follow-up Exercises

As already mentioned, the exercises are chosen because it was decided that students should be able to handle the objects to be described. It was agreed that focusing on pictures of people, characters and costumes, rather than sets– would be more feasible. Research for vocabulary, terminology, syntax, style, etc. could also be carried out most efficiently if the image was as close to the students as possible to avoid any guessing or interpretation. It was decided to start from the basic components and work with a bottom-up approach.

a) Different materials

Each group is handed two square swaths of some 20 cm of different textiles. First, the students have to define the type of material (linen, cotton, silk, etc.) and then, with the help of internet search engines, they have to identify its weaving pattern: chiffon, organza, etc. Providing photograms of films on the computer screens from where to describe costumes, props, etc. did not work as well as students actually handling materials to be described. Though this exercise may sound excessive, in some cases it is the first time students have to investigate a material in such detail. This exercise does simulate real working conditions when any costume has to be described, as is the case when the describer is granted access to opera or theatre costumes.
Example:
Student group is handed a sample of Harris Tweed herringbone

The students will have to identify that it is wool, and has a herringbone pattern. This is a pattern consisting of rows of short, slanted parallel lines with the direction of the slant alternating row by row and used in masonry, parquetry, embroidery and weaving.

![Fig. 1: Close-up of Harris Tweed herringbone material](image)

b) Different textures

This is a difficult issue, but since people with sight problems may have developed a more sophisticated tactile input, we highlight the need for a full description. Again, students are presented with one of the following items and are asked to describe at least five textures: kitchen green scourer, kitchen aluminium scourer, ball of cotton, cork from a wine bottle, a cube of sugar, a walnut and a shawl.

So a cashmere pashmina shawl would be described as: soft, rich, luxurious, springy, warm, silky.

c) Different patterns and colours

This time, students have to describe the pattern found in the swath of material previously given, along with the colour. Students are encouraged to give a lively description of colours, rather than green and brown: avocado green or chocolate brown. Light is usually described as golden yellow, or even enhanced with the use of synaesthesia such as “warm golden” or “cold blue”.

Wallpaper samples can also be used for this exercise. Patterns may vary from the straightforward red and white gingham to a busy African batik, or
the rich William Morris’ Strawberry Thief (Fig. 2), or an intricate Indian paisley (Fig. 3).

Fig. 2: Strawberry Thief by William Morris

Fig. 3: Indian paisley.
d) Different costumes

Once students are aware of the many elements present in costumes, they are given some fashion magazine cuttings. They have to do two complementary tasks. First, they are instructed to describe the costume. Then, they have to provide the style or fashion trend in a maximum of five words.

e) Different faces

The same procedure as for costumes is now exploited in exercises for the different elements present in a human face. Students are shown pictures which have to be described. Make-up and hairstyle are also considered here as elements and students are also given pictures of faces with make-up to be described. At this stage, only facial features are important; emotions are dealt with later when the students start describing clips.

An interesting example is to hand out the following three faces.

![Fig. 4: Three faces](image)

Students are asked to describe the faces, paying special attention to the facial features. After five minutes, the faces are projected onto the classroom screen and students are asked to read aloud their descriptions. It is curious that the three faces are of the same man but with different hair-styles. The issue of objective versus subjective description can be exploited here.

3.3. Exercises with Film Clips

From now on students are presented with film clips, and they have to put into practice all the skills developed by the previous exercises. Time will now be the restriction which they will have to take into account when choosing the relevant information to describe.
Designing a Course on Audio Description

a) The anagram of a film

This exercise is apparently extremely simple, but it is in fact a challenge. Students are asked to look on the internet for the opening screenshot of a distributor or producer, such as Buena Vista, Columbia Pictures, for example the eight seconds of the Metro Goldwyn Mayer lion roaring. They are asked to describe it bearing in mind synchronisation.

As soon as we start working with film clips, the different templates from different studios and countries to produce a script are shown. As with film scripts, Audio Description texts have different formats since there is not consensus on the best or optimal presentation. For this reason, students are shown some different scripts, and they can use the format they like best.

b) Two minutes of a simple linear action

The students now receive a two-minute clip of one action. The clip chosen shows only one action at a time. They have to produce a synchronic Audio Description. In the course we try to use, as much as possible, recent film releases, and an example for this exercise will be the opening of Woody Allen’s *Scoop* (2006) when Joe Strombel, played by Ian Macshane is in Charon’s ferry on the Styx carrying the soul of the dead to Hades.

c) Two minutes of an artistic film

The total running time of the clip to be audio described is two to three minutes. Its content is more complex than the action taking place, and description also has to take into consideration other elements present on the screen. That is, students have to narrate the action which takes place, but also the characters, such as the chase in *I Robot* (Alex Proyas 2004).

d) Two minutes of a complex scene

This clip is chosen because of the many actions taking place in a synchronous way. Students have to take decisions while they comment on why some actions were chosen while some were discarded. The final fight scene from *Kill Bill I* (Quentin Tarantino 2003) is a good example.

e) The opening credits

Opening credits may vary: from the simple typography on a black foreground such as Woody Allen openings, to the artistic opening of *Catch me if you can* by Steven Spielberg (2002). In fact, credits may be one of the most complex AD exercises given the number of variables which may be present (Orero & Wharton 2007), and sometimes in a synchronous way: the titles, the music possibly with meaningful lyrics, some written information,
and a story within a film, or the ambience and costumes. It may be an opening of a film, such as the *Merchant of Venice* (Michael Radford 2005). In this film the plot starts before some credits appear, which forces the production of an AD of the plot, the characters, the location, and very complex costumes. This film is chosen as the last exercise since this clip demands from the students all the skills which have been progressively acquired throughout the course. The AD is written out in full, students are also asked to write any additional comments they wish to. This exercise is assessed for the final mark.

4. Assessment

Student’s work is assessed in each class. Participation, comments and replies are taken into consideration as well as relevant questions asked while performing exercises. The only written work which has to be handed in is the AD of a film opening and the initial credits of a film. Evaluation in this case takes into consideration the choices made, along with the justification given for them. It also takes into consideration the vocabulary, terminology, plus the overall effect of the AD, since it has to be synchronic.

5. Conclusions

Much thought and preparation have gone into the design and implementation of this course. A wide range of exercises have been provided since students change every year and it has been found that what may be a very fast or easy exercise for one group may take a considerable amount of time for others. Student’s affinity to exercises is also important. For this reason, many more exercises that can be performed in 20 hours have been described as guidelines. Different materials, exercises and approaches were chosen and tested during classes and courses given at different universities. Students were asked at the end of each exercise to comment on the usefulness and degree of satisfaction. Using the old fashioned method of trial and error we were able to achieve what has been shown in this article. Students were extremely interested in the course, and some decided to take this speciality as their research topic, and it is hoped they will become successful describers providing high quality scripts to make audiovisual media accessible and enjoyable.

Bibliography


Anna Matamala & Pilar Orero

Benecke, Bernd (2007) "Film AD". Workshop delivered in the Advanced Research Seminar in Audio Description, Barcelona: Universitat Autònoma de Barcelona February 12-16.

Díaz-Cintas, Jorge (2007) "Por una preparación de calidad en accesibilidad audiovisual". Trans 11, 45-60.


Filmography

Catch me if you can (2002) USA
Dir. Steven Spielberg

Merchant of Venice (2005)
USA
Dir. Michael Radford

USA

Scoop (2006)
US & UK, Dir. Woody Allen

Kill Bill I (2003)
USA
Dir. Woody Allen
This article is part of the research project “Subtitling for the Deaf and Hard-of-Hearing and Audio Description: First Scientific Approaches and Their Applications”. Ref: HUM2006-03653FILO, funded by the Spanish Ministerio de Educación.

A module on Audio Description has been included in the MA in Audiovisual Translation (MTAV, www.fti.uab.es/audiovisual/) at the Universitat Autònoma de Barcelona (UAB) in its face-to-face format in the 2006/07 edition. The online format will start from 2008.

The pretend blind person takes the role of the teacher. His/her comments and questions help the other two students to perform a more accurate description. This practice enhances both control and dynamism, and everything is performed in a more relaxed, less formal atmosphere.

http://www.youtube.com/watch?v=XVEflECtfBM. (last visited on 24 February 2007)
http://www.youtube.com/watch?v=8C-pRG_BDNg. (last visited on 24 February 2007)
http://www.youtube.com/watch?v=iVEgK3nCkao. (last visited on 24 February 2007)

Short AD courses were given at Universidade de Vigo and Universitat Jaume I both in Spain during 2007.