

## Audio Describing Sound – What Sounds are Described and How? Results From a Flemish Case Study

 Gert Vercauteren 

University of Antwerp

 Nina Reviers 

University of Antwerp

Citation: Vercauteren, G. & Reviers, N. (2022). Audio Describing Sound – What Sounds are Described and how? Results From a Flemish Case Study. *Journal of Audiovisual Translation*, 5(2), 114–133. <https://doi.org/10.47476/jat.v5i2.2022.232>

Editor(s): M. Carroll & A. Remael

Received: December 13, 2022

Accepted: July 6, 2022

Published: December 21, 2022

**Copyright:** ©2022 Vercauteren & Reviers. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/). This allows for unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### Abstract

Given that the aural meaning-making channels are the main channels that people with sight loss have access to when processing audiovisual products, it is striking that hardly any research has been done on how dialogues, music, and sound effects are integrated in audio description (AD). A possible explanation is the lack of conceptual, theoretical, and methodological frameworks to systematically analyse sound. In the present article we study whether insights from film studies, multimodality research and audionarratology can be used to study sound in AD. After an overview of the state of the art, we present a case study that was carried out within the context of two master's dissertations and for which an analytical framework for sound descriptions was designed and tested on different scenes from the Flemish film *Loft* (Van Looy, 2008). In addition to testing the validity of the framework, we wanted to obtain some concrete insights into what types of sounds were described and what narrative position and functions they occupied. On the one hand, the analysis shows AD usually renders the sounds that are the most prominent and most easily recognised. On the other hand, the analysis also shows that although the framework for the analysis clearly has its merits and can provide valuable insights, determining the narrative functions of sounds remains difficult. Therefore, the article ends with various new research questions and a clear plea for more research in this uncharted territory.

**Key words:** Audio Description, Narratology, Audionarratology, Multimodality, Media Accessibility.

 gert.vercauteren@uantwerpen.be, <https://orcid.org/0000-0001-6711-2005>

 nina.reviers@uantwerpen.be, <https://orcid.org/0000-0003-0271-6662>

## 1. Introduction

In their introduction to the volume *Audionarratology – Interfaces of Sound and Narrative*, Mildorf and Kinzel (2016) describe how the sonic constituents of a narrative text contribute to the creation of the storyworld (Herman, 2009), both by the author and in the minds of the audience. But they also point out that these aural components “are still rarely taken into focus in narratological analysis” (Mildorf & Kinzel, 2016, p. 2) and that more research is warranted to find out how they actually contribute to telling and interpreting stories. The same is true and equally relevant for AD. Given that the aural meaning-making channels are the main channels that people with sight loss have access to in order to process and understand the story, it is striking that there is hardly any research on how dialogues, music, and sound effects contribute to storyworld creation in AD. Most of that research is product-oriented and descriptive in nature, analysing the use of sound in the audiovisual source text (e.g., Remael, 2012; Szarkowska & Orero, 2014) or how sounds contribute to or hinder multimodal cohesion (Reviere, 2018). A small section of studies takes an experimental approach, studying, for example, how sounds affect the target audience’s reception (e.g., Fryer et al., 2013). Finally, there is some research that explores the possibility of replacing or enhancing AD by alternatives in which sound takes a more prominent role or sound features are explicitly enhanced as a function of the AD (e.g., Lopez & Pauletto, 2009; Lopez et al., 2020). The present article contributes to this new line of research into sound in AD with a study that was conducted within the framework of two master’s dissertations supervised by the authors of this article. Drawing on insights from audionarratology, film studies and multimodality research, a model for the analysis of sound in AD was developed and applied to selected scenes of the Flemish film *Loft* (Van Looy, 2008), (a) to test the solidity of this analytical model and see if and where it needed fine-tuning, and (b) to obtain a first glimpse at what types of sounds were described and how. As such, the present article wants to pave the way for further descriptive research as well as for studies looking at the describer’s decision-making process and reception of different ADs (in terms of sound description) by the target audience.

## 2. State of the Art

As mentioned above, very little research has been done on the description and function of sound in AD. The general advice in guidelines remains limited to the observation that AD should not cover “important” sound effects and that describers should recognise “unclear” sounds and decide what the nature of their description should be (Greening et al., 2010; Szarkowska & Orero, 2014). The Ofcom guidelines (ITC, 2000) acknowledge that description is necessary for “any sounds that are not readily identifiable” (p. 18). However, few studies elaborate further on how to identify ambiguous terms such as “important” or “unclear” or “unidentifiable” sounds, or how sounds complement the descriptive units of AD to convey the entire message (Reviere, 2018).

Most of the research on sound and AD is product-oriented and descriptive in nature. Remael (2012), for instance, studied some of the issues involved in the production and reception of film sound, with a view to identifying its challenges for AD. She points out that a careful analysis of the role of sounds

in film is required for the production of coherence in AD, but that reception research with users is an equally important avenue for further research. Szarkowska and Orero (2014) illustrate the importance of sound for AD and offer general advice on how to describe sound. Fryer (2010) studies the role of sound in AD from the point of view of audio drama, underlining the importance of the coherence between sound and AD and proposing a classification of the effects of sounds. Building on Fryer's classification, Reviers (2018) studied multimodal cohesion between sound and AD and identified the different ways in which Dutch AD refers to sounds. In contrast, Igareda (2012) studied the specific role of music and lyrics in films and offers suggestions for a better integration of AD and music. Apart from this descriptive work there is limited experimental research on sound in AD. Fryer et al. (2013) studied how sounds affect the target audience's sense of presence. They confirmed that sound effects play a key role in stimulating presence in film, but that it depends on the sensory characteristics of the listener. For people with sight loss, words alone seem to provide a rich, imaginative experience. Sounds did not significantly increase their presence levels in the experiments conducted. Finally, the work done by Lopez and her colleagues takes a different approach. Rather than studying how sound can be treated in AD, these researchers look for ways to give sound a more prominent position in accessible products for people with sight loss. In earlier work, Lopez and Pauletto (2009) investigated the possibility of creating an alternative to an audio described film in the form of an artwork without any images that they termed an audio film, which is different from other aural narratives such as radio plays or audio dramas in that it is meant to be broadcast in cinemas and uses cinematographic techniques in an adapted form. They found that audio films can convey narrative information effectively, but that some aspects, such as unrealistic sounds, are more difficult to render. In later work, Lopez et al. (2020, 2021) studied whether AD could be enhanced by using more of the potential that sound design has to offer. Based on the principle of accessible filmmaking where accessibility is included from the earliest production stages of an audiovisual product and not as a mere afterthought, they explored whether enhanced AD could be used to reduce the need for traditional AD by implementing strategies such as (a) the use of sound effects to render narrative information on actions, characters, and settings, (b) the use of binaural audio to position sounds in different locations in space, and (c) the use of the first person. Audience feedback suggests that this is a very promising new avenue of research, but as in all the other contributions cited above, the authors mention the need for further research in this area.

This latter observation is far from new. As early as 1980, in the special issue of Yale French Studies devoted to cinema and sound, Altman pointed out that research into film was up to then predominantly image-bound. In his introduction to the special issue, Altman makes a very strong plea for a more systematic study of sound, particularly from a technological angle. While things have started to change, Altman's plea is still highly relevant, which can possibly be attributed to the fact that there still is a conceptual, theoretical, and methodological gap when it comes to studying the semiotics of sound (Kokkola & Ketola, 2015; Reviers, 2018; Van Leeuwen, 1999). Altman's observation was echoed by Chion, a film scholar. Already in 1985, Chion devoted a publication exclusively to the role of sound in film, entitled *Le son au cinema* [Sound in cinema]. In his work, Chion details a range of features of film sound, from the types of meanings it can express to its

particular semiotic qualities. Further, he illustrates how filmmakers can manipulate these qualities to enhance the meaning they contribute to the complete audiovisual text. Chion also stresses the importance of the relationship between sound and image for audiences as they extract meaning from a film. The aural and the visual modes in film are so closely interconnected that it is difficult to speak of one without mentioning the other: “on ne ‘voit’ pas la même chose quand on entend; on ‘n’entend’ pas la même chose quand on voit” [we don’t see the same thing when we hear; we don’t hear the same thing when we see]. (Chion, 1999, p. 3). Sound effects add semiotic value to filmic images. This helps to convey complex types of meanings that the individual modes alone cannot express as effectively. This multimodal interaction forms the core of the filmic medium and is what Chion calls the “audio-visual contract” in film. This strong audio-visual relation is, for obvious reasons, a challenging aspect for the practice and study of AD.

One of the research domains that studies the meaning-making potential of sound and that is quickly becoming a reference framework in audio-visual translation (AVT) and MA research is multimodality. This field of study starts from the premise that non-verbal modes can express the same types of meaning functions as language. And just like language, non-verbal modes have a set of resources with a semiotic potential from which authors can select the appropriate ones to convey their message (Taylor, 2020; Van Leeuwen, 2005).

The main focus of multimodality research with regard to sound has been on “(...) inventorizing the different material articulations and permutations a given semiotic resource allows, and describing its semiotic potential, describing the kinds of meanings it affords” (Van Leeuwen, 2005, p. 4). This includes detailed descriptions of the types of sounds that exist in audiovisual texts (music, sound effects, dialogue, silence, diegetic and non-diegetic sounds, on- and off-screen sounds, etc.), the properties of sounds and how they can be manipulated to carry different meanings (such as pitch, range, volume, nasality, etc.) and the effects sounds can have (such as perspective, cause and effect, emotional response, symbolic meaning, etc.) (See Reviers, 2018, for a detailed overview).

The approaches and terminology within multimodality research, however, are varied and heterogeneous. The field does not (yet) offer a systematic and straightforward analytical and terminological framework (Boria et al., 2019), which can be explained, among other things, by the complexity of the issues and the wide variety of multimodal artefacts. Such an undertaking becomes even more complicated when the wide variety of (intersemiotic) translations that are performed on such texts (e.g., AD) are studied. One method for multimodal analysis that has been applied to the study of MA is multimodal transcription, which was initiated by Baldry and Tibault (2006) and builds on social semiotics. It has been applied to AVT (Pérez-González, 2014; Taylor, 2003) and to the study of multimodal cohesion in AD (Remael & Reviers, 2018; Reviers, 2018).

Another gap that remains in the current work on sound relates to how sound contributes to the development of stories. The narratological approach to the study of audiovisual texts and their translations is particularly relevant for the study of AD in film and has become a widely recognized theoretical foundation within the field. It has been discussed by Vercauteren (2016) for instance and

constitutes the academic foundation of the ADLAB AD guidelines (Remael et al., 2014). Moreover, the narratological role of sounds in audiovisual texts is a research topic that has recently attracted interest beyond translation alone. It has inspired the development of a new area of study entitled audionarratology. Mildorf and Kinzel (2016) describe how the different aural components of a narrative text contribute to the creation of the storyworld (Herman, 2009). At the same time, they point out that these components “are still rarely taken into focus in narratological analysis” (Mildorf & Kinzel, 2016, p. 2) and that more research is needed to find out how they actually contribute to telling and interpreting stories.

The main focus of studies in audionarratology is on exploring how all kinds of sounds (music, language, prosody and sound effects) contribute to the creation of storyworlds through a narrative’s main constituents, namely actions, characters and spatio-temporal settings (Hühn et al., 2016; Vercauteren, 2016). However, literature on audionarratology remains scarce and constitutes an eclectic collection of approaches. Some authors focus on the role of music (e.g., Wulf, 1997, as cited in Mildorf & Kinzel, 2016) and its narrative contribution to affective meaning. Others, such as Huwiler (2016) or Lutostanski (2016), focus exclusively on aural text types, such as radio plays and audio dramas, where sounds function independently and create storyworlds in their own right. Such approaches are particularly valuable for AD, an acoustic genre as well, even though distinctly different, since AD is added to a pre-existing soundtrack that cannot be manipulated to fit the narrative development of the target text. Finally, some scholars focus on identifying the main narrative functions of sounds. Thom (1999), for instance, developed a list of narrative functions of sound (cf., next section). From this list of functions, it becomes instantly clear that sounds contribute to all levels of a story. Sounds can function independently and move the narrative forward and they can contribute to strengthening existing narrative building blocks expressed through other semiotic resources.

A final point that audionarratology scholars emphasise is the importance of active inferencing by audiences in reconstructing a narrative based on acoustic signs. This observation resonates with film scholars like Altman and Chion mentioned earlier, who emphasise the implicit links between sounds and images in films and how audiences interpret both simultaneously. As Mildorf and Kinzel (2016) say: “(...) the (aural) narrative only works when recipients try to make sense of what they hear, filling in the missing information by making assumptions and drawing inferences” (p. 11). The importance of this inferencing or what Van Leeuwen (2005) calls information linking is particularly relevant for the role of sound in AD, too. In this respect, Mildorf and Kinzel (2016) focus on the importance of cohesive ties when analysing sound from a narratological point of view, a sentiment that is echoed in multimodality research, where a lot of emphasis is put on the interaction between various semiotic resources in a text. A sound never functions on its own, and it is the interaction between resources that creates meaning. This is also called multimodal cohesion and it is something that has been underlined as a crucial consideration when it comes to AD (Reviers, 2018; Remael & Reviers 2018). Such cohesive ties can be made explicit in the source text and/or the AD, but in many cases, they remain implicit and depend entirely on the logical connections and information links audiences create based on the multimodal input.

Multimodal cohesion is the term used within text-analytical and social semiotics perspectives (Van Leeuwen, 2005; Reviers, 2018). When adopting audionarratological angle, such inferences are studied from the perspective of mental model theory (Johnson-Laird, 1983). When processing narratives, audiences create a mental image of the story they are presented with, or a mental model of “who did what to and with whom, when, where, and why, and in what fashion” (Herman, 2002, p. 9). This creative process combines input from the audiovisual text (bottom up) with the audience’s previous knowledge and experience (top down). The top-down process in particular helps audiences fill in narrative gaps and supplement information to make sense of the input they obtain from the audiovisual text. As a result, the process of story creation in the audience’s mind is to some extent personal. This top-down process is further developed within schemata theory. Schemata are cognitive structures that divide our generic knowledge into interrelated categories (humans, objects, settings, actions) (Alexander & Emmott, 2014; Sanford & Emmott, 2012). Audiences use such categories as a basis to interpret input from a text to support narrative reconstruction. These reconstruction processes are obviously relevant to AD to understand how audiences reconstruct a story based on an AD version of an audiovisual source text in combination with their specific background knowledge and experience (Remael et al., 2014; Vercauteren, 2016). So far, however, research has almost exclusively focused on how the visual meaning-making channels contribute to this story re-creation. The ways in which sounds contribute to the process have received far less attention (Mildorf & Kinzel, 2016; Vercauteren, 2022) and a call should be made for more research on the role of sound in AD from an explicitly narratological point of view.

The aim of the present article is to explore this role, using concepts from film studies, multimodality and audionarratology as a starting point. The specific questions we focus on are (a) what types of sounds are described, (b) how sounds are described, and (c) what the narrative position and function of the described sounds is. Given the current lack of proven solid analytical frameworks to study sound in AD, a final research question (d) relates to the methodology for sound analysis in AD and asks whether the theoretical framework and terminology described above are sufficient for a solid analysis of sound in AD and what the methodological needs for this genre are.

### **3. Methodology**

#### **3.1. Development of the Analytical Framework**

For our case study we decided to operate within a framework that is regularly used in the field of AD, namely narratology. A considerable part of both research and practice in AD is founded on the belief that one of the aims of mainstream films is to tell a story and that it is the task of the describer to ensure that people with sight loss can still understand it, despite not having access to all the meaning-making channels. The fact that “50 percent of the motion picture experience is aural” (sound designer Vincent LoBrutto, quoted in Buhler et al., 2010, p. xxii) makes analysing sounds and sound descriptions from a narratological angle all the more relevant. Knowing what types of sounds are



described and how, what position they occupy in the narrative and what narrative function they serve can offer a good starting point to understand how people with sight loss can recreate, process, and interpret stories and in this way help to confirm or modify existing AD guidelines.

### **3.1.1. Types of Sound in Film**

As Chion (2016) points out, in film “image is the constant focus of attention, but [...] sound at every moment brings about a number of effects, sensations and significations” (p. 150). This aural signification comes to us through the soundtrack, which is generally subdivided into three interacting, complementary components, namely, vocal sounds or dialogue, music and environmental sounds or sound effects (Bordwell et al., 2017; see also Barsam & Monahan, 2016; Buhler et al., 2010; Verstraeten, 2008). According to Buhler et al. (2010), these three components operate on two distinct, hierarchical narratological levels: “[d]ialogue occupies the sonic foreground and music and effects the background.” (p. 9). In other words, dialogue is the aural component that drives the narrative action forward, whereas music and sound effects provide the background to that action, functioning as an aural setting offering a sense of spatial and temporal (dis)continuity. For our analysis, we defined dialogue in a very broad sense. In addition to the actual verbal dialogue, characters regularly produce non-articulated utterances such as sobs, sighs, laughs, etc. that also add meaning to the narrative, which includes but is not limited to prosody and paralanguage. Given that it may not always be clear who produces these utterances, the describer may decide to explain them in the AD, hence our decision to broaden the definition. These three components will serve as the first element in our analysis to categorise the sounds we encounter.

### **3.1.2. Position of the Sounds in the Narrative**

For a correct understanding of the story, it is crucial that we know where the sounds – or rather their sources – are located in the film. According to Buhler et al. (2010), sounds are closely linked to the physical world in which they appear since people will normally try to anchor them to an object they can be associated with (p. 65). For people with sight loss, this anchoring process is presumably much more difficult because they do not have full access to the physical world and/or the film they are watching. In film, anchoring of sounds to a physical source operates on two distinct levels, namely “the level of the narration [...] and the level of the sound in the fantasy or screen world” (Buhler et al., 2010 p. 65). On the level of the narration, it is important to know whether sounds originate from a source within or outside the story world, i.e., whether they are diegetic or non-diegetic. On the level of the screen world, i.e., within the story world presented in the film, it is important to know if the source of the sound is presented on-screen or whether it is not visible and thus off-screen. Finally, a distinction can be made between sounds that occupy a figure position and those that occupy a ground position. This distinction is particularly relevant for AD since existing guidelines usually give describers the advice not to spoon-feed the target audience by describing clearly identifiable sounds (generally in a figure position) and to make identification of sounds that are less clear (generally in a

ground position) easier. With regard to their position in the narrative, we will categorise sounds according to these three features, i.e., whether they are diegetic or non-diegetic, whether their source is located on-screen or off-screen, and whether they occupy a figure or a ground position.

### **3.1.3. Function of the Sound in the Narrative**

Sounds contribute significantly to the audiovisual narrative of films and create different effects that can at any time serve one or more functions. The effects and functions of sound that can be created in texts through the choice of a specific type of sound and its position or by manipulating its properties, such as volume, pitch, and modality, are theoretically endless. Furthermore, the effect of a sound and therefore its function in the narrative also depends on the audience's interpretation while recreating the storyworld, as discussed earlier. As a result, it becomes extremely difficult to compile an exhaustive list or even provide a clear categorization of the effects and functions of sound in film for analytical purposes. Several scholars from film studies, social semiotics, multimodality, and narratology have, however, started to describe some of the various effects sounds can have as an indication of the function they can fulfil in the text. Chion (1985), for example, speaks of three ways in which sounds can be listened to: (a) causal listening – using sound to identify objects, settings or actions, (b) semantic listening, which relates to the meaning derived from spoken language, and (c) reduced listening – the affective, emotional, physical, and aesthetic qualities of a sound that are inferred from its overall texture. In contrast, social semiotician Van Leeuwen (2005) underlines the effects of the perspective that sounds create and how the figure or ground position of a sound can reveal the (changing) distance between characters, actions, and settings. He also emphasises the rhythmic properties of sound and their contribution to the creation of temporal settings. Finally, Fryer (2010) developed a categorization of six sound effects to guide the analysis of sounds in AD, including realistic sound effects that identify objects, actions or settings, symbolic effects, often referring to non-diegetic sounds that invoke connotative or abstract meaning, conventional effects, which refer to clearly identifiable sounds, like the sound of footsteps, and finally impressionistic sound effects that have an affective impact on audiences and create a mood or atmosphere. To conclude, some scholars have undertaken a similar categorization, but from a narratological angle, which is of particular value for the purpose of this article.

One such scholar is Verstraeten (2008), who claims that dialogue can serve to focalise sound, either in a realistic, distorted or even expressionistic way, to create sound cuts and/or an ominous or otherwise uncomfortable atmosphere. Music can serve to drive or twist the plot, to create a bridge between successive shots, to create a certain rhythm, to introduce a theme or character, to canalise or intensify emotions and it can be used for characterization. Finally, sound effects can be used to characterise settings and characters, to elicit emotions, and as a means of focalisation. Thom (1999) does not look at the three components separately, but says that dialogue, music, and sound effects can all have any of the following functions:



- (1) Suggest a mood, evoke a feeling
- (2) Set a pace
- (3) Indicate a geographical locale
- (4) Indicate a historical period
- (5) Clarify the plot
- (6) Define a character
- (7) Connect otherwise unconnected ideas, characters, places, images ,or moments
- (8) Heighten realism or diminish it
- (9) Heighten ambiguity or diminish it
- (10) Draw attention to a detail or away from it
- (11) Indicate changes in time
- (12) Smooth otherwise abrupt changes between shots or scenes
- (13) Emphasise a transition for dramatic effect
- (14) Describe an acoustic space
- (15) Startle or soothe
- (16) Exaggerate action or mediate it

(Thom, 1999, p. 9).

A comparison of the two contributions shows that many of the functions mentioned by Verstraeten (2008) and Thom (1999) overlap. In addition, the narratological functions mentioned by these authors correlate to some extent with the effects of sounds described by the film and multimodality scholars mentioned earlier, who point to similar semantic potential. For the present analysis, the classification by Thom (1999) seems more suitable than the other frameworks mentioned, since it takes an explicitly narratological approach and it is more concrete and elaborate when compared to Verstraeten's categories (2008). As a result, we will use Thom's (1999) categorization as a basis for our analysis.

#### **3.1.4. Type of Sound Description**

Finally, we also wanted to see what types of sounds are actually described. This raises the question of what counts as a sound description, a topic that has received little attention in AD research so far. One of the few authors who has addressed this issue is Reviers (2018). Based on a transcription and analysis of three clips from a corpus of Dutch audio described films within social semiotics and multimodality frameworks, she studied the ways in which audio describers refer to sounds from the soundtrack to disambiguate their role in the narrative. While this constituted a preliminary study based on a limited sample, she identified a non-exhaustive list of four levels of description:

(a) The descriptive unit mentions or reiterates the source of the sound directly through the use of nouns or verbs that refer to it (e.g., mentioning a door when it squeaks), (b) the descriptive unit evokes the source of the sound indirectly through a process of lexical relations (e.g., mentioning

setting the table to clarify the sounds of glasses and cutlery), (c) the descriptive unit (additionally) refers to the quality of the sound rather than the source to disambiguate it further and it uses nouns, verbs and occasionally adjectives to do so (e.g., mentioning someone rapping on the door instead of knocking), (d) the descriptive unit indirectly and implicitly supports the process of information linking, based on which audiences can infer the source of the sound (e.g., a police office is mentioned, which is the only clue to disambiguate the sounds coming from an interrogation). In addition, Reviers (2018) underlined that the degree of explicitness of a sound co-depends on other factors beyond the AD alone, such as timing and the level of synchronicity between the sound and the AD and the prominence of the sound in the soundscape, i.e., whether it occupies a figure or ground position.

### 3.1.5. Materials and Annotation

For our analysis we used the Flemish film *Loft* (Van Looy, 2008). This choice was partially motivated by the fact that the sounds of the opening scene of this film and their description had been the subject of an earlier study by Reviers (2018) and we wanted to build on and partially replicate this study. In addition, the annotation of the scenes in this article was realised in collaboration with two master's students. Since they had no prior experience with sound analysis, they started by re-annotating and re-analysing the same opening scene individually so that we could determine the accuracy of their work, compare it with the earlier analysis undertaken by Reviers (2018) and adjust the analytical approach and annotation strategy where necessary.

In addition to the opening scene (scene 1: 00:00:00 – 00:01:45), we analysed two longer scenes (scene 2: 01:21:18 – 01:25:57 and scene 3: 01:43:47 – 01:48:29), both of which were chosen because, contrary to the opening scene, they included dialogue sequences and less common sound effects (e.g., indicating a flashback). The annotation of each scene was completed by one of the two students.

As with the opening scene, their annotation and analysis were discussed within the team and whenever discrepancies were identified in the annotational approach, e.g., with regard to the function of the sound or the type of description, a consensus was sought based on the theoretical framework and the earlier work by Reviers (2018). For the analysis, a transcription table was created which included one row for every sound that was identified and columns to list the elements discussed in our methodological framework:

Table 1.

*Example of Transcription*

Time code	Characterization of the sound	Type of Sound <sup>1</sup>	AD (Yes/No)	Concrete Description	Type of AD (Reviers, 2018)	Position <sup>2</sup>	Function
01:45:30	Sound of a pen writing	Sound effect	Yes	Still wearing his gloves, he writes Sarah's farewell letter <sup>3</sup>	1 DG	On Fig	Heighten realism

Source: authors' own work.

During the annotation process the research team found that the available terminological and conceptual framework was certainly useful but required more critical reflection and further development. For instance, it was not always clear how to delineate a sound (type/position), the narratological categories were too broadly defined and it was not straightforward to attribute a sound to one or another category (e.g., evoke a feeling versus startle or soothe). In some cases, sounds seemed to fit into multiple categories (e.g., heighten realism, draw attention to a detail and clarify the plot). Finally, the way the AD referred to sounds was a matter of degree. It remains unclear in many cases whether the AD, for instance, refers to a sound implicitly through information linking or whether it does not refer to it at all. These limitations need to be considered in the following section, which presents the first exploratory analysis.

## 4. Analysis and Discussion

### 4.1. Types of Sound

Table 2 presents a breakdown of the different types of sounds we encountered, including an indication of how many were described and how many were not.

<sup>1</sup> In this column we either entered "dialogue", "music" or "sound effect".

<sup>2</sup> This column was further subdivided into three cells, one indicating whether the sound was diegetic (DG) or non-diegetic (NDG), whether its source was on-screen (On) or off-screen (Off) and whether the sound occupied a figure (F) or ground (G) position.

<sup>3</sup> Our translation. The film and its description were both made in Dutch as was the analysis in the dissertations.

Table 2.

*Types of Sounds and Number of Sound Descriptions*

Scene	Dialogue			Music			Sound Effects		
	Total	AD	No AD	Total	AD	No AD	Total	AD	No AD
Scene 1	0	0	0	7	0	7	28	7	21
Scene 2	31	7	24	22	0	22	87	36	51
Scene 3	24	6	18	13	0	13	69	27	42

Source: authors' own work.

We identified a total of 281 sounds, 55 (19.6%) of which were related to dialogues, 42 (14.9%) were music and 184 (65,5%) were sound effects. 83 (29.5%) of these sounds were described and 198 (70.5%) were not. A closer look at the described sounds shows that music was never included in the description, dialogue was described 13 times (15.7% of all sound descriptions) and sound effects were described 70 times (84.3% of all sound descriptions). While in scene 1, 100% of the sound descriptions were related to sound effects, scenes 2 and 3 show results that are more comparable to the overall picture. In scene 2, 16.3% of the descriptions related to dialogues and 83.7% to sound effects. In scene 3, 15.7% of the descriptions related to dialogues and 84.3% to sound effects.

A little under 30% of all sounds were described. Due to a lack of reference materials from previous research, we cannot make any statements as to whether this is a high or low percentage. However, some observations can be made. First of all, the finding that music was never described can possibly be ascribed to the fact that all sounds identified as music were non-diegetic. Since this music does serve various functions in the narrative, such as creating moods or emotions, it would be interesting to study in the future if these functions were integrated in the AD in other ways than an explicit reference, for example by means of specific words, stylistic choices, or voicing. Another observation is that very few instances of what we defined as dialogue (i.e., verbal but also non-articulated utterances) were described. The descriptive units that can be linked to dialogue all refer to non-articulated utterances. For example, when one of the characters, Luc, is sitting on the bed in the loft next to the dead body of a girl, Sara, who was murdered there, we can hear sobbing and to make it clear who is sobbing, the AD reads “Luc zit **huilend** op het bed naast Sara” [Luc is sitting on the bed next to Sara, **crying**]. A possible explanation for why no descriptive units referred to verbal dialogue is that the first scene did not include any dialogue at all and the other two scenes we analysed came much later in the film when the voices of the characters were already known, and speaker identification was no longer necessary. Finally, the vast majority of sounds that were included in the descriptive units (84.3%) related to sound effects (as opposed to the categories of music and dialogue; think of the sound of rain), which means that 38% of all sound effects encountered (70 out of 184) were described. From the following tables, in which we will focus only on the sounds that were actually described, it will become clearer what positions and functions these sound effects (in addition to the non-articulated utterances) occupied in the narrative.

## 4.2. Position of the Sound

Table 3 gives an overview of the position that the audio described sounds occupy in the narrative. For every scene we look at the number of diegetic versus non-diegetic sounds, on-screen versus off-screen sounds, and figure versus ground sounds.

Table 3.

*Position of the Audio Described Sounds in the Narrative*

	Diegetic or Non-Diegetic		On-screen or Off-screen		Figure or Ground	
	DG	NDG	ON	OFF	F	G
Scene 1	6	1	6	1	6	1
Scene 2	38	5	36	7	39	4
Scene 3	32	1	27	6	24	9
Total	76	7	69	14	69	14

Source: authors' own work.

The majority of the described sounds were *diegetic* sounds, the source of which was shown *on screen*. The sounds occupied a *figure* position. More specifically, 76 of the described 83 sounds (91.6%) were diegetic, versus only 7 (8.4%) non-diegetic. In 69 of the 83 sounds (83.1%) the source was visible on-screen while in 14 cases (16.9%) it was located off-screen. Finally, 69 of the 83 sounds (83.1%) occupied a figure position and 14 (16.9%) had a ground position. If we look at the percentages per scene, the same applies to all analysed material: In scene 1, 85.7% of the sounds described were diegetic versus 14.3% non-diegetic, 85.7% were on-screen versus 14.3% off-screen and 85.7% had a figure position versus 14.3% with a ground position. In scene 2, 88.4% of the described sounds were diegetic and 11.6% were non-diegetic, 83.7% were on-screen versus 14.3% off-screen and 90.7% had a figure position compared to 9.3 % that had a ground position. In scene 3, the results regarding the diegetic versus non-diegetic comparison are somewhat different from those in the other scenes and the overall result since 97% were diegetic and 3% non-diegetic. With 81.8% on-screen and 18.2% off-screen sounds, the results were comparable to those for scenes 1 and 2 and to the overall results. Finally, the percentages regarding figure (72.7%) versus ground (27.3%) position were again different from the overall results and the results from scenes 1 and 2.

A possible explanation for why the vast majority of the described sounds are *diegetic* sounds in a *figure* position with their source *on-screen* is that these sounds are arguably the most relevant for the narrative and therefore the most prominent in the source text, which means they are more easily recognised by the describer and hence included in the AD. However, this seems to go against existing guidelines, which say that the describer should avoid spoon-feeding the target audience and not describe what is already clear. Rather, the focus in these guidelines is on describing unclear sounds. One possible explanation for why these sounds are described most, regardless of the guidelines, may be that just like in the case of images, the most prominent sounds get described. A second explanation may be that, as was explained in the literature review, sounds and images cannot be

separated from each other. Rather, they create meaning in combination. In other words, when a visual action or object is described, the sound that goes with it is included in the description as well. However, further research analysing this aspect in AD as well as process research investigating the choices made by describers is needed to confirm or refute this explanation. One additional question that comes into play here relates to multimodal cohesion: If a sound that is clearly audible is described, the description explicitly creates coherence. If that sound were not described (as suggested in the guidelines), the audience would still be able to recognise it but there would be less (or no) cohesion, which could possibly make the audience question that source and/or impose a higher cognitive effort on their part. Again, this is something that requires further study.

#### **4.3. Narrative Function of the Sounds**

Next, we identified the narrative function of the described sounds to see if we could find any relation between sound description and narrative function. An overview of the narrative functions of the sounds that were described can be found in table 4.



Table 4.

*Narrative Function of the Audio Described Sounds*

Function	Scene 1	Scene 2	Scene 3	Total
Suggest a mood, evoke a feeling	1	5	1	7
Set a pace	0	0	0	0
Indicate a geographical locale	0	0	2	2
Indicate a historical period	0	3	0	3
Clarify the plot	5	4	10	19
Define a character	0	2	4	6
Connect otherwise unconnected ideas, characters, places, images or moments	0	0	0	0
Heighten realism or diminish it	0	5	3	8
Heighten ambiguity or diminish it	0	0	0	0
Draw attention to a detail or away from it	1	7	1	9
Indicate changes in time	0	0	0	0
Smooth otherwise abrupt changes between shots or scenes	0	0	0	0
Emphasise a transition for dramatic effect	0	0	0	0
Describe an acoustic space	0	0	0	0
Exaggerate action or mediate it	0	17	12	29

Source: authors' own work.

Table 4 shows a varied picture. The sounds included in the AD had various functions. The two functions encountered most were those clarifying the plot (19 instances or 22.9%) and exaggerating or mediating an action (29 or 34.9%). In other words, most of the described sounds supported the action of the narrative. 6 (7.2%) of the described sounds could be linked to the characters in the narrative, and 5 (6%) could be linked to the spatio-temporal setting. Seven (8.4%) of the sounds had what could be called an extradiegetic narrative function, suggesting a mood or evoking a feeling in the target audience. The categories “heighten realism or diminish it” and “draw attention to a detail or away from it” can be linked to all different narrative constituents; a more fine-grained analysis would be needed to see what their exact narrative referent is.

As mentioned in the methodology section, analysing narrative functions and effects of sounds is a very difficult exercise given the virtually endless possibilities that can be created and the partly individual nature of reception and interpretation. In addition, we are aware that table 4 does not completely represent all the functions of sounds in the narrative since we only included them in one

category while some probably could be fitted into two or more categories. It would be interesting to see replications of this exercise to find out if and how this part of the analysis can be further refined. However, it is clear that most of the sounds that were described (48 or 57.8%) had a function that supports the narrative action. Sounds linked to other narrative constituents were less often described as were sounds conveying metadiegetic emotional or affective meaning. This raises the interesting question regarding the extent to which the types of meaning that become less prominent or even get lost in the AD version are rendered through other AD features such as word choice, style or voicing, or maybe do not need to be described at all because they are clear without any AD (see also Fryer et al., 2013, who did not find any significant impact of sound effects on presence).

#### 4.4. Type of Sound Description

Finally, using the categories identified by Reviers (2018), we also looked at how the sounds that were included in descriptive units were actually described. The results are presented in table 5.

Table 5.

##### *Types of Sound Description*

	Explicit Source	–	Explicit Nature	–	Implicit Meronymy	–	Implicit Information Linking	–
Scene 1	3		1		3		0	
Scene 2	29		0		3		11	
Scene 3	22		1		4		6	
Total	54		2		10		17	

Source: authors' own work.

In most cases (54 or 65.1%) the sounds were described by explicitly referring to their source (as in "Sirens come closer"). In 2 descriptions (2.4%) a reference to the nature of the sound was made (as in "Rain gushes down on the hood of the car"). Ten of the 83 sound descriptions (12.0%) mentioned the sound implicitly by means of meronymy, i.e., by using words that are lexically related to the source of the sound (for instance the sound of breaking glass, preceded by "with flailing arms he keeps falling until he crashes down on the hood of the BMW"). In 17 cases (20.5%) the reference was implicit and had to be derived by means of information linking and inference.

The finding that most of the sounds were described by explicitly referring to their source can be attributed to the source often being visible on screen (83.1%, see Table 2). But again, since no comparable analyses exist, it is difficult to say if this is a general feature of sound description or not. In addition, the last of the four categories, i.e., an implicit reference to sound through information linking, turned out to be rather hard to use and often it was difficult to decide whether a sound was described implicitly or not described at all. For example, in scene two (at 01:23:14), a muffled bumping sound can be heard when two of the characters supporting a third one, Vincent, let go of

him so he loses his balance and falls on his back. The AD reads “Vincent is slap en half-bewusteloos” [“Vincent is limp and semi-unconscious”]. While this description indirectly explains the bumping sound (i.e., Vincent falling on his back because he is partly unconscious), some would say this constitutes a character description rather than a sound description. This again shows that images and sounds are intrinsically linked in the original and can therefore probably never be fully separated in AD either. But it also shows that the categories used to identify sound descriptions are not fully delineated yet and further research may be needed to refine them.

## 5. Conclusion

The aim of the present article was to develop a framework for a narratological analysis of sounds in audio described films. Given its small scale, it is important that its limitations are duly considered. First, the analysis only covered a very small sample of three scenes and focused solely on Flemish descriptions. Larger scale analyses in multiple AD languages are needed to corroborate or refute the results. In addition, the annotation and analysis processes revealed that the current aggregated conceptual and terminological framework based on insights from multimodality research, film studies and (audio)narratology lacks rigour and requires further development to allow for a fine-grained and nuanced inquiry. Finally, the text-analytical approach highlighted textual issues, but did not allow us to explore the describers’ decision-making process.

Nevertheless, this exploratory study points to interesting avenues for further research. The first finding is that music and dialogue are only rarely described, even though the role of music, for instance, has been underlined by researchers such as Igareda (2012). In the present study, sound effects are described the most, with a clear tendency to describe diegetic, on-screen sounds in figure positions. At first glance this result does not seem to match with what might be expected based on current guidelines and literature that advise describing only unclear sounds and not focus on what is evident. A third observed tendency is that most described sounds mainly contribute to conveying actions rather than characters or settings. All in all, there seems to be a rather fixed focus on specific types of meaning when it comes to describing sound. It is impossible, based on text analysis alone, to identify the reasons for this and our findings raise an extensive number of relevant new questions relating to how describers make their choices and how audiences receive the described film. How do AD audiences, for instance, draw on other types of meaning related to characters, settings, and emotions to recreate the filmic narrative? Do other semiotic channels, such as voicing, contribute to this process? To what extent can audiences disambiguate sounds that are not described and do such undescribed sounds in the target text hold the same semiotic potential as the original sounds and as sounds that are described? Do some sounds and meaning potentially get lost in translation and does this have a negative impact on audiences’ enjoyment? As Fryer’s et al. (2013) experiment indicated, sound effects did not have a significant impact on presence levels, for instance.

A final observation is that most described sounds are referred to explicitly. Nevertheless, a good number of sounds are disambiguated by the AD implicitly and most sounds are not described at all.

This observation points first and foremost to the importance of active inferencing and information linking by audiences. As mentioned earlier, the analytical framework requires further development to evaluate the level of explicitness in AD accurately and assess the impact on audiences. Does it mean AD audiences require a greater cognitive effort to recreate the story and does this have a negative impact on their enjoyment? Or can audiences rely on their mental models to recreate the story based on the sounds even without disambiguation by the AD? In brief, the analysis showed that when it comes to sounds there seems to be much more going on in the creation and reception of AD than research is currently addressing.

To conclude, this article hopes to have brought home the point that there is a profound need for a better understanding of how film sounds are processed by audiences in interaction with AD and what cognitive processes and efforts are required to do so successfully. Particularly for those sounds that were not included in the AD text, it is necessary to know how audiences can grasp their meaning. In the clips under study, for instance, several sounds like the breaking of glass, heavy rain, or the sound of a knife, were not described and audiences would have to infer or deduce the sources of such sounds. Priorities for further research are, first, the development of solid and fine-grained theoretical and terminological frameworks and, second, the development of experimental, cognitive research on the reception and processing of sound in AD.

## References

- Alexander, M., & Emmott, C. (2014). Schemata. In P. Hühn, J.C. Meister, J. Pier & W. Schmid (Eds.), *The living handbook of narratology*. Hamburg University. <http://www.lhn.uni-hamburg.de/article/schemata>
- Altman, R. (1980). Introduction. *Yale French Studies*, 60, 3–15.
- Baldry, A. P., & Thibault P. J. (2006). *Multimodal transcription and text analysis*. Equinox.
- Barsam, R., & Monahan, D. (2016). *Looking at movies. An introduction to film* (11th ed.). W.W. Norton & Company.
- Bordwell, D., Thompson, K., & Smith, J. (2017). *Film art. An introduction* (6th ed.). McGraw-Hill Education.
- Boria, M., Carreres, A., Noriega-Sanchez, M., & Tomalin, M. (2019). *Translation and multimodality*. Routledge.
- Buhler, J., Neumeyer, D., & Deemer, R. (2010). *Hearing the movies. Music and sound in film history*. Oxford University Press.
- Chion, M. (1985). *Le son au cinema [Sound in the Cinema]*. Edition de l'Etoile.
- Chion, M. (1999). *L'audio-vision: son et image au cinema [Audio-vision: sound and image in the cinema]*. Editions Nathan
- Chion, M. (2016). *Sound. An acouological treatise* (J. Steintrager, Trans). Duke University Press (Original work published 2010).
- Fryer, L. (2010). Audio description as audio drama. A practitioner's point of view. *Perspectives: Studies in Translatology*, 18(3), 205–213. <https://doi.org/10.1080/0907676X.2010.485681>
- Fryer, L., Pring, L., & Freeman, J. (2013). Audio drama and the imagination. The influence of sound effects on presence in people with and without sight. *The Journal of Media Psychology*, 25(2), 65–71. <https://doi.org/10.1027/1864-1105/a000084>

- Greening, J., Petré, L., & Rai, S. (2010). A comparative study of audio description guidelines prevalent in different countries. RNIB.
- Herman, D. (2002). *Story logic: Problems and possibilities of narrative*. University of Nebraska Press.
- Herman, D. (2009). *Basic elements of narrative*. Wiley-Blackwell.
- Huwiler, E. (2016). A narratology of audio art. Telling stories by sound. In J. Mildorf & T. Kinzel (Eds.), *Audionarratology. Interfaces of sound and narrative* (pp. 99–115). De Gruyter.
- Hühn, P., Meister, J. C., Pier, J., & Schmid, W. (2010). *The living handbook of narratology*. Hamburg University. <http://www.lhn.uni-hamburg.de>
- Igareda, P. (2012). Lyrics against images: Music and audio description. *Monti: Monografías de traducción e interpretación. Multidisciplinary in Audiovisual Translation*, 4, 233–254. <https://doi.org/10.6035/MonTI.2012.4.10>
- Independent Television Commission. (2000). *ITC guidance on standards for audio description*.
- Johnson-Laird, P.N. (1983). *Mental models. Towards a cognitive science of language, inference and consciousness*. Cambridge University Press.
- Kokkola, S., & Ketola, A. (2015). Thinking outside the “Methods Box”: New avenues for research in multimodal translation. *Vakki publications*, 4, 219–228.
- Lopez, M., Kearney, G., & Hofstädter, K. (2020). Seeing films through sound. Sound design, spatial audio, and accessibility for visually impaired audiences. *British Journal of Visual Impairment*. <https://doi.org/10.1177/0264619620935935>
- Lopez, M., Kearney, G., & Hofstädter (2021). Enhancing audio description. Inclusive cinematic experiences through sound design. *Journal of Audiovisual Translation*, 4(1), 157–182. <https://doi.org/10.47476/jat.v4i1.2021.154>
- Lopez, M., & Pauletto, S. (2009). The design of an audio film. Portraying story, action and interaction through sound. *JMM: The Journal of Music and Meaning*, 8, Winter 2009.
- Lutostanski, B. (2016). A narratology of radio drama. Voice, perspective, space. In J. Mildorf & T. Kinzel (Eds.), *Audionarratology. Interfaces of sound and narrative* (pp. 117–132). De Gruyter.
- Mildorf, J., & Kinzel, T. (2016). *Audionarratology. Prolegomena to a research paradigm exploring sound and narrative*. In J. Mildorf & T. Kinzel (Eds.), *Audionarratology. Interfaces of sound and narrative* (pp. 1–26). De Gruyter.
- Pérez-González, L. (2014). *Audiovisual translation: Theories, methods, issues*. Routledge
- Remael, A. (2012). For the use of sound. Film sound analysis for audio description. Some key issues. *Monti: Monografías de traducción e interpretación*, 4, 255–276. <https://doi.org/10.6035/MonTI.2012.4.11>
- Remael, A., Vercauteren, G., & Reviers, N. (Eds.) (2014). *Pictures painted in words: ADLAB audio description guidelines*. Edizioni Università di Trieste.
- Remael, A., & Reviers, N. (2018). Multimodality and audiovisual translation. Cohesion in accessible films. In: L. Pérez-González (Ed.), *Routledge handbook of audiovisual translation* (pp. 260–280). Routledge.
- Reviers, N. (2018). Tracking multimodal cohesion in audio description. *Linguistica Antverpiensia New Series. Themes in translation studies*, 14, 22–35. <https://doi.org/10.52034/lanstts.v17i0.477>
- Sanford, A. J., & Emmott, C. (2012). *Mind, brain and narrative*. Cambridge UP.
- Szarkowska, A., & Orero, P. (2014). The importance of sound for audio description. In A. Maszerowska, A. Matamala, & P. Orero (Eds.), *Audio description. New perspectives illustrated*. John Benjamins BV.
- Taylor, C. (2003). Multimodal transcription in the analysis, translation and subtitling of Italian films. *The Translator*, 9, 191–205. <https://doi.org/10.1080/13556509.2003.10799153>

- Taylor, C. (2020). Multimodality and intersemiotic translation. In L. Bogucki & M. Deckert (Eds.), *The Palgrave handbook of audiovisual translation and media accessibility*, (pp. 83–99), Palgrave Studies in Translation and Interpreting.
- Thom, R. (1999). Designing a movie for sound. FilmSound.org. Learning space dedicated to the art and analyses of film sound design.  
[https://www.filmsound.org/articles/designing\\_for\\_sound.htm](https://www.filmsound.org/articles/designing_for_sound.htm)
- Van Leeuwen, T. (1999). *Speech, music, sound*. McMillan.
- Van Leeuwen, T. (2005). *Introducing social semiotics*. Routledge.
- Van Looy, E. (Director). (2008). *Loft* [Film]. Woestijnvis.
- Vercauteren, G. (2016). *A narratological approach to content selection in audio description* [Unpublished doctoral dissertation]. University of Antwerp.
- Vercauteren, G. (2022). Narratology and/in audio description. In E. Perego, & C. Taylor (Eds.) *Handbook of audio description*. Routledge
- Verstraeten, P. (2008). *Handboek filmnarratologie* [Handbook of Film Narratology]. Vantilt.
- Wulf, C. (1997). Ohr. In C. Wulf (Ed.) *Vom Menschen: Handbuch Historische Anthropologie* [On People; Handbook on Historical Anthropology] (pp. 459–464). Beltz.