Narratology and/in audio description

Gert Vercauteren
Published online on: 07 Apr 2022

How to cite: Gert Vercauteren. 07 Apr 2022, Narratology and/in audio description from: The Routledge Handbook of Audio Description Routledge
Accessed on: 22 Nov 2023
1. Introduction and definitions

Audio description (AD) has been defined in various ways. The essential idea that can be found in all these definitions is that it makes audiovisual products accessible to people with sight loss by translating the visual and aural elements that these people do not have access to into a verbal commentary that helps them understand and enjoy the original product. As such, AD can be considered as a form of intersemiotic (Jakobson, 1959) or intermodal (Kaindl (2013), Geerinck and Vercauteren (2020)) translation. If we take this translation theoretical perspective as a point of departure, we can say that the task of audio describers essentially consists of two main phases: first, they must carry out a systematic and detailed analysis of the source text to determine its exact meaning. Second, they must decide how to render that meaning in a suitable target text. Both these phases contain very specific difficulties. On the one hand, the source text, be it a film, a play or an opera, is multimodal in nature, creating meaning “from the combination of and interaction between different semiotic channels: images, dialogues, sounds and text-on-screen” (Vercauteren, 2012: 209). In other words, audio describers need very good insights into these different channels and the ways in which they interact in order to get a full understanding of the source text. On the other hand, one of the inherent limitations of AD is that it cannot interfere with the semiotic channels the target audience does have access to: the dialogues and relevant sound effects. This means that audio describers will usually not be able to include all the information they want in their target text and they will have to set priorities. To do so, they need to know what elements in the source text are most relevant and what elements are secondary. In short, what they need is a solid theoretical framework that can serve as a foundation for both source text analysis and target text creation.

In that respect, narratology has always sparked the interest of AD scholars. Already in her early account of AD research, Braun (2007) identified narratology as one of the fields of study that “could be one of the pillars for researching AD and developing future audio describers’ understanding of some of its difficulties” (p. 20). This is by no means surprising given that many of the audiovisual products that are made accessible through audio description involve telling a story. As such, narratology, defined by Meister (2014) as a “variety of theories, concepts and analytical tools . . . that play a central role in the exploration and modelling of our
ability to produce and process narratives in a multitude of forms” (par. 2) can offer a solid theoretical foundation for the creation of audio descriptions. On a general level, narratology can be beneficial to AD in three respects.

2. Narratology as a universal analytical framework

First of all, narratology allows audio describers to analyse any narrative source text in a systematic way. Particularly in the early stages of the discipline, narratologists tried to identify and define universal principles of narrative (Meister, 2014). In other words, they were looking for features that transcend any specific medium and can therefore be applied to all kinds of narrative. Although this approach later received its fair share of criticism – as will be explained later – early formal or structuralist narratology continues to offer a useful toolbox for anyone analysing a narrative text (Sommer, 2004). Two main characteristics seem particularly relevant for AD, namely that all narratives use the same general constituents (i.e., characters, actions or events and the spatio-temporal settings in which these take place) and that they are all constructed through a three-tier process.3 The first tier in the process is what Bal (1997) calls a fabula, an abstract construct without any concrete representation that she defines as “a series of logically and chronologically related events that are caused or experienced by actors” (Bal, 1997: 5). One important remark that has to be made with regard to this definition is that it mentions only two out of the three main narrative constituents, namely events (or actions) and actors.4 The third one, namely the spatio-temporal settings in which these events are set, is not included, a fact that Bal (1997) explains by pointing out that these settings can normally be inferred by the audience and, when that is not the case, the audience will simply provide a setting itself. In the case of the audiovisual products that constitute the source text of an audio description, however, Bal’s (1997) claim seems more problematic. As Chatman (1978) points out, verbal stories such as novels could take place “ ‘nowhere in particular’ . . . [but] movies have difficulty in evoking this kind of nonplace” (p. 106). Indeed, in film virtually every shot contains visual information about the time and place in which the action is set. It is therefore imperative that this third constituent of the story is duly taken into account in the source text analysis as well.

In the following phases of the construction process, the fabula will undergo two fundamental transformations. A first transformation is achieved in the second tier, which Bal (1997) calls the story, defined as “a fabula that is presented in a certain manner” (p. 5). The verb presented is somewhat misleading here, since the story is still an abstract construct without any concrete (re)presentation. The transformations taking place on the story level are of an organisational nature and can be grouped into three broad categories: a) the organisation of the actions, essentially to achieve specific narrative effects and elicit various emotions in the audience (Bordwell, 1985; Vandaele, 2018); b) the physical and mental characterisation of the agents, predominantly to make them unique and to motivate their actions and reactions; and c) the materialisation of the spatio-temporal setting against which the story will unfold.

3. Narratology as a medium-independent framework

The second transformation the fabula undergoes happens in the third tier, called the narrative text or “a text in which an agent relates . . . a story in a specific medium, such as language, imagery, sound, buildings or a combination thereof” (Bal, 1997: 5). In other words, it is on this level that authors decide in what medium they will tell the story and what specific
techniques they will use to present that story in a concrete form. Again, narratology offers valuable insights for audio describers on this level. Indeed, as Sommer (2004) points out:

narratology has proved to be a reliable basis for interdisciplinary and transgeneric studies that focus on the forms and functions of narrative both in those texts traditionally classified as fictional narratives (novels, short stories, epic poems) and in other narrative art forms such as drama, film, hypertexts and painting.

(p. 6, my emphasis)

All the media mentioned previously have their own specific ways of presenting narratives. Even when a novel and a film use the same characters and settings as main constituents and organise them in the same way, for instance, presenting parts of the narrative chronologically and others by means of flashbacks, the specific way in which the narrative will be presented to the audience will be radically different, since both media use their own representational techniques. Again, this is highly relevant for source text analysis in audio description, particularly since watching films is deceivingly easy. Or as Monaco (2009) puts it: “we know very well that we must learn to read before we can attempt to enjoy or understand literature, but we tend to believe, mistakenly, that anyone can read a film” (p. 174). It is true that we do not need to be film experts in order to enjoy a film, but audio describers, who have to make that film accessible to people with sight loss, cannot simply “absorb” the images. They must understand how they are composed and what principles and techniques are used to create specific meanings.

In summary, the films that are the narrative source texts in AD are the result of an elaborate, multi-faceted set of transformations that will ultimately determine the unique nature of the film the audience will be presented with. But this creative process does more than just create specific effects. It also offers a roadmap for the audience to process the narrative text. Authors start from the chronological fabula, then decide how they will organise the chain of actions that take place and motivates them through the characters’ traits and determine in what space and time the actions will be set. Finally they decide how all this will be concretised in the narrative text. The audience will then work in the opposite direction: based on the concrete cues they are presented within the narrative text, they will interpret the presentational and organisational choices made by the author and work their way back to the initial fabula. As such, the audio description must provide all the information the audience needs to do so. Here again, narratology can offer valuable insights into how to decide what information it is that the audience uses and needs to process stories.

4. Narratology as an author-oriented and audience-oriented framework

Indeed, narratology studies both story creation and story processing/interpretation. In other words, it studies narratives both from the author’s and from the audience’s perspective. One of the fundamental objections against the structuralist approach mentioned earlier was that it did not take into account any cognitive aspects (Sommer, 2004). Structuralism seemed to imply that the meaning of a narrative resided exclusively in the text itself; that it was the sum of all the narrative constituents and the way in which they were organised and presented. With the cognitive turn in the early 1990s, it was acknowledged that narrative meaning is the result of a complex interaction between the text and its audience: rather than just passively receiving a narrative, audiences actively engage with the text, bringing their own knowledge and expectations to bear on the text to create constantly developing and evolving storyworlds (Herman,
2002). For audio describers, knowing what kinds of knowledge audience members use to process and interpret narratives, knowing what cues from the text and mental principles they use to create those storyworlds is crucial. It gives them the necessary insights to decide what information is most relevant and should get priority in their AD.

5. Basic narratological insights and current research on narratology in AD

Since narratology is such a vast field with ramifications into many other disciplines, it will not be possible to explore in this chapter all the ways in which it can help audio describers. In the following sections, we will therefore limit ourselves to an overview of current contributions to the field of AD that use narratology as a theoretical framework, either to gain more insight in the source text or to develop principles for the creation of the target text.

6. Source text analysis

Various authors have turned to narratology to create frameworks and principles for source text analysis in AD and all three main narrative constituents have received attention in this respect. First, there are the actions or events. As Porter Abbot (2008) remarks, “without an event or an action, you may have a ‘description’, an ‘exposition’, an ‘argument’ . . . but you won’t have a narrative” (p. 12). These actions serve different functions that must be taken into account when analysing the source text.

In her contribution to the ADLAB guidelines, Mazur (2015) focuses on two of these functions. First, she points out that they “are an essential part of a film narrative, moving the story forward” (p. 19). They are what Pitkänen (2003) calls “plot-advancing information” (p. 24), in other words, they are the narrative constituent that makes the story develop and as such they should be one of the main points of focus in both the analysis of the source text and the audio description resulting from it. Second, Mazur (2015) emphasises that actions and reactions are one of the ways in which we get to know the characters. They constitute what Margolin (2007) calls the characters’ communicative or behavioural dimension: the way characters behave provides insights into their inner personality; their cognitive and emotional world.

In his article on the usefulness of narratological insights for audiovisual translation and media accessibility, Vandaele (2018) points to the fact that actions do not only serve intra-diegetic functions like the ones that have been presented. When analysing their source text, audio describers must be aware that the actions, or rather the way in which they are presented (or not), can elicit specific emotions in the audience. Based on Bordwell’s (1985) narratological approach to film, Vandaele (2018) links the way in which actions are communicated to three main narrative emotions, namely suspense, curiosity and surprise. In order to create suspense – an emotion that is usually aimed at in thrillers – actions are presented chronologically and/or the presentation of certain future actions is delayed. When filmmakers want to elicit curiosity, they can decide to ostensively delay the presentation of past actions; they can present the effect of a certain action (a dead body) but not the cause (the murder – and the murderer). The third emotion, surprise, is generated when filmmakers delay the presentation of certain actions in a non-ostensive way; without the audience noticing that certain actions are not presented. Audio describers must be aware of these different extradiegetic functions and make sure that their target audience can recognise and experience them too.

Finally, Vercauteren (2012) turned to narratology to design a strategy for the description of temporal relations between actions or events in subsequent scenes. Within the broader
framework of content selection in AD, his contribution presents a schematic model that helps audio describers analyse whether or not actions and the relations between them have to be described. The basic premise in this analysis is that – in theory – the target audience should not know more than any sighted viewer, which means that actions and their temporal relations should only be described when they are also clear to the sighted viewer. If this is not the case, priority can be given to other information that is clear. However – and this goes for all narrative information in the AD – research has shown that the processing of aural information requires more mental effort than that of visual information (Fresno, 2014). Moreover, explicitation is one of the universals of translation (Laviosa-Braithwaite, 1998) and target texts tend to be more explicit than the corresponding source texts. In other words, even if actions and their relations are not clear in the filmic source text, describers can decide to render them explicit in the AD if they think it is beneficial to the processing of the narrative as a whole.

Narratologists have long acknowledged that the second main narrative constituent, the characters, are more difficult to study and define. Indeed, even when narrative characters can sometimes be prototypical or even stereotypical, it is very hard to reduce them to pre-defined sets of features and traits that allow for a systematic analysis. Or as Herman and Vervaeck (2005: 70) observed: “[i]t is not at the level of character analysis that structuralist narratology has made its most significant contribution”. Still, the study of narrative character received its fair share of attention in structuralist narratology and has been applied to source text analysis in AD too. Vercauteren (2016b) combined three accounts to present a framework to assist audio describers in analysing narrative characters and determining priorities in terms of character information to include in their description. A first account is that of Margolin (2007), who acknowledges that every character in any narrative work has its own unique identity that can be discovered by analysing that character’s properties, which can be grouped into three dimensions (Margolin, 2007: 72–73): a) physical properties such as age, sex, posture, and so on – any exterior characteristic that allows the audience to identify and recognise a character based on his/her external appearance; b) communicative and behavioural properties, in other words, all the verbal utterances and physical (inter)actions that define a character and c) mental properties, which can further be subdivided into perceptual, emotive, cognitive and volitional properties — everything a character sees or hears, feels, thinks or wants, or the characteristics that define a character’s inner life. By analysing these three dimensions, audio describers already get a fairly accurate idea of how a character is perceived by the audience and as such they can determine what elements can be included in the description. A second account that can further refine that analysis and help describers prioritise that information is the one by Rimmon-Kenan (1983), who analyses characters on the basis of three vectors, namely a) their complexity or the number of character traits they have, ranging from one to an infinite number of traits; b) their development, in other words, whether they remain the same throughout the entire narrative or undergo some kind of change and c) the access to their inner life, which can be very limited or very wide. Rimmon-Kenan (1983) sees these three vectors as sliding scales and characters can be situated anywhere on these scales. As such, this account can help audio describers decide how much information about a character to include in the description. Characters with a very limited number of traits, that hardly develop throughout the narrative and offer little insight into their inner life, will probably need less description than highly complex characters that develop a lot and provide full access to what they see, hear or feel. Finally and linked to the development vector from the previous account, Bal (1997) studies how characters appear throughout the narrative. The starting point is that a character can be new (when it is introduced for the first time) or known (when it appears again later in the narrative). If a character is new, we usually know nothing about it and all the information
we get is presented for the first time. If a character is known, the information the audience receives when it appears again can be a repetition of what is already known, or an addition of new physical, behavioural or mental traits. This additional information can confirm what is already known or it can change the perception of that character (i.e., when the character does something the audience does not expect based on what they know about that character). Finally, throughout the narrative, characters are linked to other characters and to settings and these relations can remain the same or change as well. A similar approach to character analysis for AD was taken by Mazur (2015), who, in addition to the parameters discussed earlier, points to the distinction between main characters, supporting characters and background characters or extras to determine how much attention they should receive in the AD as a whole. In another contribution, Mazur (2014) focuses on one particular feature of narrative characters, namely their gestures and facial expressions as a reflection of their mental dimension. More specifically, she presents a classification of these gestures and facial expressions as discourse-supporting, discourse-filling or discourse-conflicting as a way to determine how important they are in terms of description. A final contribution worth mentioning here is that by Benecke (2014), who provides guidance with respect to the description of a character feature that has not been discussed previously, namely its proper name. Various authors have shown that proper names fulfil specific functions in language systems in general (Manini, 1996) and narratives in particular (Margolin, 2002; Rimmon-Kenan, 1983). They anchor the characters as unique participants in the narrative and can contribute to their characterisation, providing information about their geographical, ethnic and social background. In audio description, the question of when to name a character by its proper name is still a thorny issue. Some guidelines suggest naming characters only when they are named in the film, which echoes the discussion that the target audience of the AD should not know more than the other audience members. Others suggest a character’s name can be given earlier, for example if it is a well-known historical figure or if its identity does not have to remain a secret. Since proper names are usually shorter than characterising descriptions, the latter approach seems to offer the advantage that describers will have more time to include other relevant narrative information in their AD. Benecke (2014), however, develops a model that allows describers to fixate characters in other ways within an evolving AD model, so that they can be named when their name is also mentioned in the original text (Benecke in this volume).

The third main narrative constituent, the spatio-temporal setting in which the narrative takes place, has received less attention in AD research so far, just as it has in narratology. Dennerlein (2009) explains this by the fact that narratology was initially concerned with the construction of narratives as temporal and causal sequences of events, while Herman (2002) points out that “if space was discussed at all it was used negatively to mark off setting from story... description from narration proper” (p. 265). However, various authors have shown that the spatio-temporal settings in which the story is set are often much more than a mere “background for the characters and their actions” (Ryan, 2012, § 29). They regularly serve a symbolic meaning and are crucial for the development of the story. Based on examples taken from the film Inglourious Basterds (Tarantino, 2009), Vercauteren and Remael (2015) present a framework to analyse narrative settings to determine relevant spatio-temporal information and decide what to include in the description. In addition to distinguishing between the temporal and the spatial dimension of the setting, they build on work by Pitkänen (2003) to point out that a setting can be global or local, that – as pointed out earlier – it can be a background to the action but can also serve a symbolic function and that it can be real or imagined. In addition, audio describers must take into account that the setting or parts of it can be presented explicitly or implicitly. A second account that is included in their framework is the one developed by Bal...
(1997) and is very similar to the one used for analysing narrative characters: again, it first has to be determined whether a setting is new or known. To determine the importance of a new setting and hence the possible attention it can get in AD, the describer should establish whether it is just a background (less AD) or serves a symbolic meaning (more AD). Once the setting is known, subsequent appearances in the narrative can be repetitions, in which case little to no AD is needed (just a general reference to the setting). If new information is presented this can again confirm what we already know, or the setting can change. This again will determine what and how much will be described. Finally, settings are often linked to characters and/or their actions, an important element to include in the analysis because it means that descriptions, if needed for the sake of brevity, can then refer to settings by mentioning a character, rather than the setting itself.

7. Target text creation

After the source text analysis has been completed and it has been determined which elements are eligible for description, describers can turn their attention to the creation of the target text. Two fundamental questions that have to be answered in this step are a) what elements do I ultimately include in my description? and b) how do I formulate my description? Both these questions have been approached from a narratological perspective in AD research.

The first question was already referred to in the introduction to this chapter. Given that descriptions cannot interfere with dialogues or sound effects, describers will usually not have enough time to include all narrative elements from the source text in their AD and they will have to set priorities. Fresno (2014, 2016), Fresno et al. (2014, 2016) and Vercauteren (2016a), among others, have turned to cognitive narratology to provide insights into how information can be prioritised in AD. The fundamental idea underlying all these contributions is that audiences process narratives by creating mental models (Johnson-Laird, 1983) or, as Herman (2002) calls them, “storyworlds”; mental representations of “who did what to and with whom, when, where, why and in what fashion” (Herman, 2002: 9). Drawing on work by Emmott (1997) and Schneider (2001), Vercauteren (2016a) explains that this is a dynamic process in which audiences use both information from the text (bottom-up processing) and personal knowledge (top-down processing), making mental model creation to some extent an individual endeavour: people who are familiar with certain references will be able to create different and/or richer mental models than people who are not. For instance, if a person reading the sentence “In the distance they spotted the characteristic contours of the Gherkin” knows that the “Gherkin” refers to the Swiss Re Building in London’s financial district, she may integrate much more information in the mental model and make a lot more inferences and assumptions allowing for a more thorough processing of the narrative, than if she does not. However, it is safe to say that mental model creation begins with bottom-up processing of information from the text, more specifically, information about the three main narrative constituents discussed earlier: actions, characters and settings. The models themselves can be conceived of as three-level structures (Emmott, 1997; Vercauteren, 2016a), as shown in Table 5.1.

The first and fundamental level that forms the basis of the mental model includes the actions that are presented. Since these are the backbone of any narrative as explained earlier, they are the information the audience will first accommodate in their model. On the second level of the model, all the actions presented will be framed in a general context; in other words, the audience will look, on a very general level, at who performs and/or undergoes the actions, when and where. In other words, they will include global information on characters and settings that allow them to situate the actions in the narrative world. This global information
Narratology and/in audio description

is no more than a label, for instance, the name of a character, a place or a time. On the third level, these general labels are enriched with specific information such as physical details of the characters, particulars about the place in which the action is set or the time at which it takes place. Throughout the narrative this three-level structure will be updated and expanded (cf. the discussion about characters and settings in section 2.2). Moreover, the audience will try to link all the successive actions (and scenes) by means of a process called contextual monitoring (Emmott, 1997). The basic premise underlying this monitoring is that the audience assumes things remain the same unless signalled otherwise. As such, audience members will update their mental model when characters enter or leave a contextual frame (frame modification (Emmott, 1997)), when the action moves to a different frame (frame switch (Emmott, 1997)) and when the action returns to a previous frame (frame recall (Emmott, 1997)). As Vercauteren (2016a) explains, the same principles for narrative processing by the audience can be used to select relevant content and prioritise information in audio description: given that descriptions (in general) contain a verb, the actions also form the backbone of the AD. They then must be framed within a general context by providing – if this information is known to the audience – information that allows the audience to know what characters are present and in what setting the action takes place. This is the basic information that is needed to process the narrative, but if time allows, more specific information about entity representations can be added to the description. In order to allow the audience to follow the unfolding of the narrative, the describer has to look for modifications, switches or recalls in the contextual frames and provide the information the audience needs to be able to recognise them.

Based on the principles of mental model creation, Fresno (2016) and Fresno et al. (2016) specifically looked at what information about characters can be included in audio description and how to prioritise this information. Specifically, they looked at the physical features of characters and experimentally tested which of these features were best remembered and hence most useful to include in the description. In another contribution focussing on the same topic of character description, Fresno et al. (2014) turned their attention to the second fundamental question in AD text creation, namely how to formulate the description. More specifically they tested whether long sentences combining all relevant character features or short sentences in which the different features were split up between different ADs resulted in better comprehension and richer mental models. They found that using shorter sentences containing less information yielded better results than using longer ones. In their 2016 paper, the same authors addressed a far more contentious issue in AD, namely the question whether audio descriptions should objectively reflect what is seen on screen or can include interpretative or subjective elements. The origin of this debate can be found in early AD guidelines (e.g., Benecke and
Dosch (2004) or Snyder (2010)), which advise audio describers to objectively describe what they see on screen so that the description does not “impose the describer’s own feelings but rather provokes the listener’s” (Rai et al., 2010: 61). Various authors have questioned this stark dichotomy and one-sided approach to AD. Kruger (2010) points out that this view does not reflect the far more complex and nuanced reality of narrative fiction: narratives don’t tell stories in an objective way but always include more or less personal (and hence subjective) choices made by the author or filmmaker and by the audience. Therefore, he claims, ADs should incorporate a narrative function too, a function that reflects the original nature of the source text. On the other hand, he acknowledges that not all source texts have the same degree of narrativity and so he suggests that the access provided by audio description should be seen as a “descriptive-narrative continuum” (Kruger, 2010: 233) ranging from AD that is explicitly descriptive, for example in documentaries, to audio narration that “moves away from a strict fidelity to what can be seen on-screen in favour of a coherent narrative” (Kruger, 2010: 233).

In other words, audio narration is not so much an independent audio track that tells what is shown on screen, but forms one, integrated narrative with the other elements from the source text to which the target audience does have access. Palmer and Salway (2015) make a similar claim, starting from a philosophical principle they call the thought-action continuum (Palmer, 2004). As the name suggests, according to this principle, descriptions can reflect a character’s actions (i.e., what is shown on screen) or they can reflect the thoughts or intentions underlying these actions. So a description like “she frowns” is nearer to the action end of the continuum whereas “she is puzzled” would be situated nearer to the thought end. In addition to these fundamental contributions to this question, other authors have started looking at more practical implications of these fundamentally different AD styles. For example, Fresno et al. (2016) studied the effects of different audio descriptions of physical traits of narrative characters on the audience’s memory. They created a visual description that reflected what was seen and a semantic description that reflected the underlying meaning of what was seen and they found that the latter yielded better results. They explained this finding hypothesising that participants who were presented with the visual descriptions experienced a higher processing load because they had “to go through the merely visual descriptions and extract their semantic meaning” (Fresno et al., 2016: 160). Walczak and Fryer (2017) took a similar approach to test the effect of different AD styles on the presence of the target audience. They created standard descriptions adhering to rules of objectivity and creative descriptions that included features such as “camera work and subjective descriptions of the characters, their actions and scenes crucial to the plot” (Walczak & Fryer, 2017: 6, my emphasis) – in other words, features that explicitly integrated narrative elements in the AD. They found that these creative descriptions resulted in higher levels of presence. Finally, Geerinck and Vercauteren (2020) took the various contributions on AD style discussed earlier as a starting point to analyse how the behavioural and mental dimension of narrative characters was described in different Flemish and Dutch TV series. They found that objective (action-oriented, descriptive or visual) descriptions and more interpreted (thought-oriented, narrative or semantic) descriptions were used equally often to represent the characters’ behavioural dimension; to represent the actions they performed and underwent. On the other hand, the vast majority of the descriptions of the characters’ mental dimension – their emotions, feelings and so on – included an explanatory, interpreted element, which could indicate that describers were led by the narrative to create their descriptions.

The final observation of the previous paragraph indicates that we still know relatively little about the interaction between narrative and audio description and that the usefulness of narratology for AD research must be explored much more. In the following section some possible avenues for future research will be suggested.
8. Future directions

In their introduction to the edited volume *Audionarratology. Interfaces of Sound and Narrative* (Mildorf & Kinzel, 2016b), published in the *Narratologia* series, Mildorf and Kinzel analyse the current landscape of narratological research in the age of multimodal audiovisual media and come to the conclusion that it is predominantly focusing on the visual channel. If the verbal channel is studied at all, it is almost exclusively on the written word and not the spoken word. Despite the fact that earlier studies have confirmed the importance of vocal qualities such as rhythm, intonation, tone of voice, and such, “the sonic qualities of fictional texts and the perceived or imagined features of narrators’ and characters’ voices are still rarely taken into focus in narratological analysis” (Mildorf & Kinzel, 2016a: 2). The authors go on to demonstrate that sounds other than the spoken word did not get the attention they deserve either. Similar observations can be made for audio description. Early definitions describing AD as a service “providing a narration of what is seen, i.e., giving an oral description of visual elements such as action costumes or props found in cultural events” (Matamala, 2005: 9) or as “the addition of a descriptive narrative to accompany the key visual elements of theatre, television, cinema and other visual media” (Greening & Rolph, 2007: 127) exclusively focusses on the visual channel in the audiovisual source text and most of the research to date also has a marked visual bias. This is not to say that no research on sound in AD has been done. Remael (2012), for example, studied the role of sound in the filmic text and the challenges it creates for AD, particularly with the aim of providing insights into how multimodal coherence can be achieved in AD. Reviers (2018) took this research a step further by analysing how AD units and sound effects from the original film interact. Finally, Walczak and Fryer (2018) tested how different voicing styles of AD, namely synthetic versus human, impacted on the presence and emotional experience of the target audience. They found that the human AD resulted in higher levels of presence than the synthetic one in drama, but that there was no significant difference in documentary. Similar results were found for preference: the participants clearly preferred the human AD in drama, but there was no clear-cut preference when it came to the voicing of AD in documentary. Although the authors did not hypothesise on this, one possible explanation of these findings may be the different level of narrativity of these two genres. Indeed, much more research into the various links between narrative, sound and AD is needed. Future studies could build on the work by Walczak and Fryer (2018) to explore how other sounds than dialogue influence the audience’s emotional experience. Mildorf and Kinzel (2016a) point out that sounds “can trigger in us an emotional response that we may still remember years later” (p. 7). So, just like voices, they may increase the audience’s presence or immersion and as such enhance the narrative experience. Likewise, sounds also have a clear cognitive dimension (Mildorf & Kinzel, 2016a: 11). They contribute to the creation of storyworlds (D. Herman, 2002) and have the potential of activating whole mental schemata (e.g., the setting of a restaurant or other location, the celebration at a party, etc.). In the field of AD, better insights into the narrative functions and possibilities of sounds are beneficial for both source text analysis and target text creation. For source text analysis, we need to know how sounds contribute to the creation of the story and its different constituents: voices can tell us a lot about characters and their emotions, the position of a sound can tell us more about the immediacy of settings, and so on. As such they can help the audio describer determine what is narratively (most) relevant. The potential of good insights into the narrative qualities of sound is probably even bigger for target text creation: if we understand how sounds conjure up mental frames, how they narrativise the storyworld, how they elicit emotions in the audience, we gain a powerful tool that helps us decide what information to include in the AD
and what information to leave out, because the audience will grasp it, or its function will be fulfilled through the soundtrack.

As this last sentence shows, this calls for more (experimental) reception research. Some of the contributions discussed earlier, such as by Fresno et al. (2016) or by Walczak and Fryer (2018), already started exploring the effects different AD styles have on the target audience, but there clearly is more room for further research there. There are already a few indications that a more narrative description style has an impact on the audience’s presence (Walczak & Fryer, 2017) and that it may also influence the audience’s cognitive load when processing audio descriptions (Fresno et al., 2014), but these are just initial explorations that need to be investigated more thoroughly. To date no experimental research has been done to confirm Fresno et al’s (2014) claim. Also further studies can analyse how sounds, either by themselves or integrated in a coherent AD, impact on the audience’s emotions and/or cognitive load. And although some research indicates that audiences prefer a more subjective narrative description style (Chmiel & Mazur, 2012, 2016), other findings seem to indicate this preference may be related to genre (Walczak & Fryer, 2017): in short, this link between AD style and level of narrativity of the source text needs further attention as well. All these research avenues can contribute to creating a more solid and wider basis for narratology as a theoretical framework underlying the study and practice of AD.

9. Conclusion

It was the aim of the present chapter to explore whether and how narratology can provide useful insights to further both the study and practice of AD. From the introduction it became apparent that it can be beneficial for AD in various respects: narratology studies the universal principles underlying story construction and as such offers a comprehensive tool for systematic analysis of narratives. In addition, narratology is not limited to the study of narratives in one particular medium but studies the principles underlying story creation in all kinds of media, which is favourable for a field like AD, in which the source text and target text are characterised by different dominant semiotic systems (visual vs. verbal). Finally, narratology not only looks at how stories are created but also at how they are processed and interpreted by their audience, which helps audio describers and AD researchers understand what information the target audience needs to grasp the story that is told.

In the following sections, a translation theoretical perspective was adopted to analyse, based on existing research in AD, how narratology can offer a solid foundation for the two main phases in the translation process, source text analysis and target text creation. In terms of source text analysis, all the main constituents of narratives — actions, characters and settings — have received attention, and principles have been developed that allow audio describers to analyse them systematically in their source text. As far as target text creation is concerned, AD research adopting a narratological approach is more fragmented and limited. The questions of how to decide what to include in a description and how to formulate such description have been addressed, but no conclusive answers have been provided yet.

This was also emphasised in the final section, which looked at future steps that can be taken to create further interdisciplinary links between narratology and AD. Experimental research and reception studies can be undertaken to study audience preferences and to investigate the various cognitive and other mental effects of different AD styles on audience members. Furthermore, additional research is certainly needed to explore how sound contributes to the
creation and reception of audio described products. In short, narratology undoubtedly offers a solid framework for AD and some of its principles have already been studied and applied. However, there are many more narratological insights and propositions that can enrich our understanding of how audio descriptions can support the story that is being told in the source text and how they can enhance the end-user experience.

Notes

1 This reference to people with sight loss is by no means meant as reductionist. Although they are the primary target audience of audio description, we are aware of recent developments in the field of media accessibility that advocate a universalist approach to inclusion (Greco, 2018, 2019) or approach the service from the perspective of the different semiotic channels that can be made accessible rather than from that of the target audience (Remael et al., 2019).

2 In the remainder of the text we will consider the source text to be a film or TV series, but the ideas and principles put forward are also applicable to other audiovisual products and static forms of art.

3 See Herman and Vervaeck (2005) chapter 2 for a complete overview. From their discussion it becomes clear that all narratologists use their own terminology. In the present paper, we will adopt the terminological concepts used by Bal (1997).

4 Bal deliberately avoids using the term characters in this definition. In this embryonic stage of the story, the actors are abstract constructs that do not possess physical or psychological characteristics yet, and as such can encompass an extensive variety of both human and non-human entities that can cause and/or undergo events.

5 See also Vandaele (2018) for an application of Bordwell’s (1985) theory to audiovisual translation.

6 In terms of target text creation, she explores the usefulness of traditional translation strategies for the audio description of gestures and facial expressions and discusses one of the main dividing issues in the field of AD, namely the choice between objective and subjective descriptions (cf. also Kruger (2010), Palmer and Salway (Palmer & Salway) or Fresno et al. (2014)). Since these elements are not strictly related to narratology, they will not be discussed further in the present chapter.

7 Possible explanations are the fact that artists “constantly experiment with and generate new effects through sound. As a result it becomes impossible to compile an exhaustive list of the possible effects or functions of sound . . . ” (Reviers, 2018: 25) or the fact that it is difficult to fix and transcribe sounds making it hard to develop a clear methodology to study them (Mildorf & Kinzel, 2016a).

8 See the analysis of the setting in the opening scene of Inglourious Basterds (Tarantino, 2009) by Vercauteren and Remael (2015), which shows that sounds can generate immediate switches from global to local settings.

10. Further reading


11. References


12. Filmography

