LUND UNIVERSITY

Audio Description of Art

The Role of Mental Imagery and Embodiment

Holsanova, Jana

Published in:
Lund University Cognitive Studies

2021

Link to publication

Total number of authors:
1

Audio Description of Art: The Role of Mental Imagery and Embodiment

Jana Holsanova

LUND
UNIVERSITY

Abstract
How can we make visual art accessible for audiences with visual impairment and blindness? Can audio description help blind and visually impaired (BVI) audiences to understand and experience art? Which ingredients must AD contain to evoke and stimulate the creation of vivid internal mental images? Can AD engage the target audiences and contribute to their aesthetic experience and enjoyment? What other means are there to enhance inclusion and immersion? The paper focuses on how audio description can evoke and stimulate the creation of vivid internal images via verbal descriptions of art and thereby contribute to an embodied aesthetic experience for the BVI audiences. It introduces research on image perception, image description, and mental imagery, relevant for audio description; summarise guidelines and recommendations for audio description of art; and offers authentic examples of AD of art tailored for the target audiences. Finally, the paper underlines the importance of reception studies and discusses the issue of inclusion and multi-sensory experiences.

Audio Description of Art: The Role of Mental Imagery and Embodiment[1]

Jana Holsanova

How can we make visual art accessible for audiences with visual impairment and blindness? Can audio description (AD) help blind and visually impaired (BVI) audiences to understand and experience art? Which ingredients must AD contain to evoke and stimulate the creation of vivid internal mental images? Can AD engage BVI audiences and contribute to their aesthetic experience and enjoyment? What other means are there to enhance inclusion and immersion?

This paper focuses on how audio description can evoke and stimulate the creation of vivid internal images via verbal descriptions of art and thereby contribute to an embodied aesthetic experience for the BVI audiences. After a short introduction on audio guides and audio description, we present results from empirical studies on image perception, image description, and mental imagery that are relevant for audio description. Next, we summarise guidelines and recommendations for audio description of art concerning creation of vivid internal images and embodiment and illustrate with authentic examples from AD. We underline the importance of feedback from BVI audiences and the need for a systematic study of end users' experiences, needs, and preferences. Finally, we discuss the issue of inclusion and multi-sensory experiences.

Introduction

Several art museums worldwide offer stand-alone descriptions of artworks, pre-recorded audio and multimedia guides, and physical or virtual guided tours. These guides and descriptions are produced by museum curators and art pedagogues for the general public. They contain standard information about the artwork (artist, nationality, title, date, dimensions), a general overview of the subject/motive, form, and colour; information about the technique, medium, and style of the artwork, and a description of its visual appearance. The aim of these verbal descriptions is to educate; to offer cultural experience, enjoyment, and relaxation; to attract more visitors to the museums; and to increase interest in art. These kinds of artwork descriptions are usually approximately three to five minutes long and can be found on museum websites, in mobile applications, and in media players at the museums. Example 1 illustrates a general verbal description of a work of art. It is the beginning of a five-minute-long visual description of Henri Matisse, *Dance (I)*, produced by MoMA (the Museum of Modern Art), and published on the museum's website.

Example 1. MoMA
Narrator 1: *7–2 Dance (First Version).* 'Painted in 1909 by the French artist Henri Matisse, 1869 to 1954. Oil on Canvas, eight feet six inches high by twelve feet nine inches wide. 260 by 390 centimetres.'
Narrator 2: 'This large painting is of five women holding hands in a circular dance, on a field of green, before a bold blue sky. The sinuous, sensuous curves of the approximately life-size dancers in their circular swirl, convey a vivid sense of rhythm and movement. The painting is dominated by large expanses of just four colours. A hilly shape, in what might be described as emerald green,

suggests a lawn. It creates a wavy horizon where it meets the deep blue sky, about one-third of the way up the canvas. The five women wear no trace of clothing, and their skin is pale flesh, tinted with pink. The last colour is a flat black that outlines the dancers and represents their shoulder-length hair. On close examination, one can see that the black of their hair is sometimes tinted with brown or green, and the black that outlines the figures ranges from very dark, to almost grey….'

While this description contains some ingredients that could stimulate the imagination of blind and visually impaired audiences and help them to access, understand, and enjoy this painting, it is not intended specifically for BVI audiences and thus is not adapted to their particular needs and preferences. In fact, only a few museums provide information and services for the target group of visually impaired visitors in order to encourage their inclusion by museums (Schoene 2021). Such services would, in addition to audio description, also include texts and signs in Braille, tactile maps for orientation, and an opportunity to touch the artworks or models. There are only a few examples of art museums and exhibitions designed and organised by an organisation for the blind. One of these is Museo Tiflológico in Madrid, inaugurated by the National Organisation of the Spanish Blind (ONCE) in 1992. It is a museum 'born by decision of its users and designed by these tailored needs' that focuses on education, experience, and inclusion.

Audio description

One important way to enhance the accessibility of works of art and inclusion for BVI audiences is through audio description. According to Perego (2019:333), 'Audio description (AD) is the acoustic verbal description of the visual elements of any cultural (static and dynamic) product for the benefit of people with visual impairment.' AD helps BVI audiences to access, better understand, and enjoy the details of the visual experience of film, theatre, and other cultural events. In the field of translation studies, AD has been characterised as an intermodal audiovisual translation since the content of images is transferred into words (Jakobson 1959, Reviers 2017).

There are many challenges for AD of artworks, the first of these being how to describe the visual appearance of an artwork for audiences that have never been able to see and to stimulate the creation of inner mental images that lead to understanding and enjoyment. The next challenge concerns the fact that the interpretation and aesthetic experience of an artwork is not only based on what we see, but also on what we know, associate, feel, and experience. This is, however, tricky, since the audio describer is not supposed to express personal associations (cf. Guidelines and recommendations). A further challenge is how to mediate the communicative purpose of the artwork and how to accordingly choose relevant aspects of the work to describe. According to Reviers (2017), audio description of film must respect the communicative purpose of the film intended by the filmmaker, adequately translate the message, and at the same time meet the communicative needs of the BVI recipients and adapt the AD for them (Reviers 2017). However, for many artworks, we do not even know what the original intentions of the artist were, and we are left with the communication of the artwork itself. Finally, there is a difference in the role of AD for reception of a work of art in comparison to video or film. In the reception of an audio described film (consisting of dynamic images, sounds, music, and dialogue), parts of the work are received and processed by BVI audiences from the original in combination with AD. However, in the case of an artwork that consists of a static image or sculpture, the reception is not strengthened by any other auditive channels and it must rely solely on the AD. Consequently, AD plays a much more significant role for the reception and it is very important for both the creation of vivid internal images and for aesthetic experiences.

Before we focus on the guidelines and recommendations for what audio description of art should contain in order to stimulate imagery, let us first have a look at research on how viewers perceive, conceptualise, and verbally describe complex images, and on what role mental imagery plays in this process.

Image perception and image description: two windows into the mind

In what follows, we will summarise results from studies on image perception and image description. These studies are relevant for the research on audio description in two respects: First, the dynamic processes of image perception and image description uncover the creativity and complexity of meaning-making and the interpretation of images. Second, the way sighted viewers perceive complex images, for instance, what aspects they focus on in their visual discovery and in their verbal descriptions, can serve as an inspiration for AD practices in general and for the AD of art in particular.

Current theories on visual perception stress the cognitive basis of art perception. We 'think' art as much as we 'see' art, according to Solso (1994). Our way of perceiving objects in a scene can be triggered by our interests, associations, expertise, or context. Holsanova (2001, 2008), using eye tracking methodology and verbal protocols, conducted a series of studies on how sighted viewers perceive, conceptualise, and verbally describe complex images. The aim of these studies was to uncover the underlying cognitive processes. Since it is not possible to directly access the content of the mind, Holsanova (2001) developed a dynamic, sequential method, in which she combined eye movement data and spoken description data. Eye movement data shows the path of image discovery and allows us to follow which image elements have been given attention, when, in what order, and for how long. This reflects human thought processes and offers us a window into the mind. Data on spoken language descriptions allows us to come closer to how the image has actually been perceived, conceptualised, and experienced by the viewers. Thus, spoken descriptions offer us another window into the mind. Both of these kinds of data have been used in order to compare the content of the visual and verbal foci of attention, and to gain insights about the underlying cognitive processes.
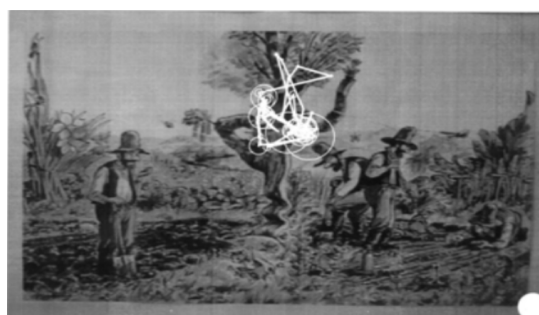
there are three birds in the tree

Figure 1. From Holsanova (2001). Eye movement patterns for one participant's looking at a children's book illustration by Sven Nordqvist (1990: *Kackel I grönsakslandet* [published in English as *A Ruckus in the Garden*, 1991]) and describing the central part of the image: 'In the middle of the field is a tree with three birds doing different things. One bird is sitting on its eggs in a nest, the other bird is singing, at the same time as the third, female, bird is beating a rug or something'.

Thanks to the combination of verbal and visual data, the following tendencies could be observed: viewers created clusters of image elements with spatial proximity ('three birds') but also clusters of image elements that were distributed across the scene, based on categorical or taxonomical similarity ('flying insects'). In addition, the viewers established mental groupings based on abstract concepts and associations 'it looks like early summer') and returned to previously inspected image elements and viewed them with another concept in mind (cf. Figure 2).

Figure 2: From Holsanova (2001). Examples of viewing patterns accompanying an image description

there are three birds in the tree
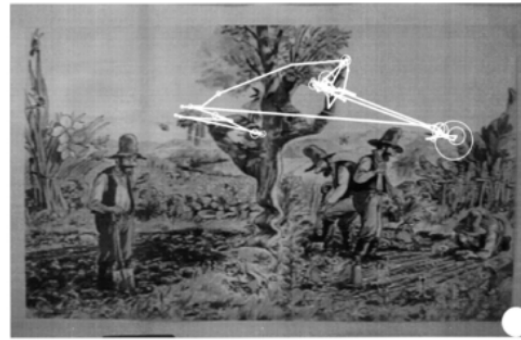
there are flying insects

Pettson is digging

actually, there are two different fields
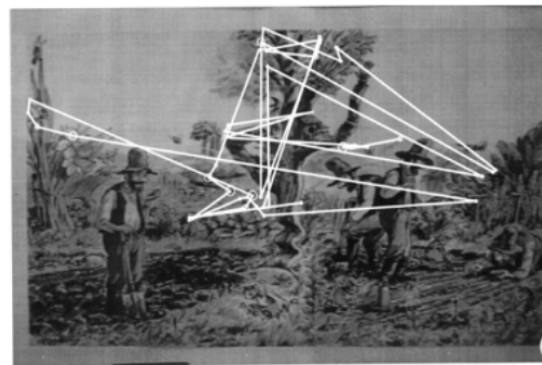
it looks like an early summer

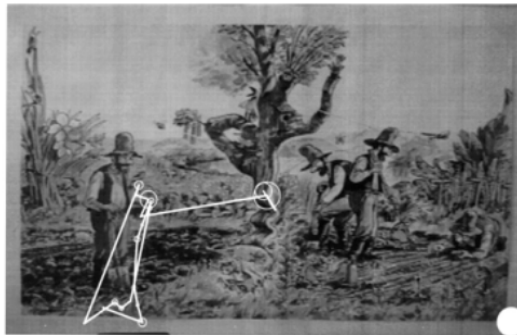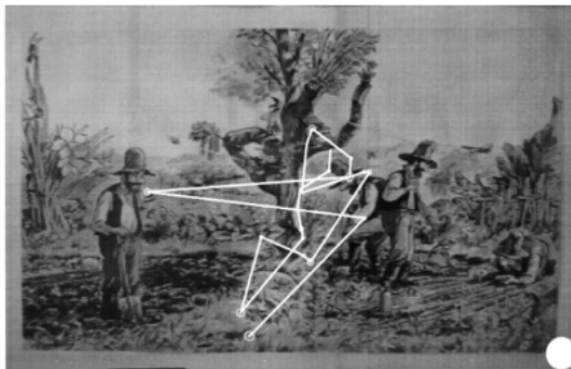Participants often used metaphors and similes for comparison ('a dragonfly looking like a double winged aeroplane'). The majority of viewers proceeded according to the saliency principle – they viewed and described visually salient elements first – and according to the animacy principle – they viewed and described human beings and other animate image elements prior to inanimate ones.

Concerning the time-course of image perception, viewers were able to get an initial holistic impression or the 'gist' of the scene very early in the viewing process (Henderson 2007). On the basis of this gist, a general overview of the scene has been formulated early in the description process. Some regularities in the visual exploration have been found, similar to those detected by Buswell (1935). The visual exploration started with an initial 'global' phase consisting of a general survey where the eyes move with a series of relatively short pauses over the main portions of the image. This phase was then followed by a subsequent 'local' phase of image exploration consisting of a series of long fixations, concentrated on small areas of the image, evidencing detailed examination of those sections. These phases were also reflected in the verbal descriptions. However, even differences among individuals were found. For instance, there was a clear distinction between two styles of image descriptions: a static descriptive style, focusing on spatial

6

arrangement, composition, and visual details, and a narrative style, focusing on temporal aspects and dynamic events (Holsanova 2001, 2008). The explanation for this is as follows: Meaning-making is an interactive meeting between the recipient, the image, and the situational context (Holsanova 2014b). All of these three aspects mediate the process of meaning-making. First, the visual appearance of an image serves as a starting point for the process of meaning-making through an interplay of various means of expression. Second, the recipients play an active role and co-create its meaning. Their personal characteristics mediate perception and interpretation of the image. Differences among individuals arise thanks to the variety of the recipients' backgrounds, interests, previous knowledge, expectations, domain and genre knowledge, expertise, emotions, and attitudes. Third, the context in which the image is displayed, perceived, and interpreted plays an important role in meaning-making (Holsanova 2021).

Mental imagery in the sighted and the blind

Another area of research relevant for audio description of art is mental imagery. This refers to the human ability to see things with our mind's eye and is apparent in a wide range of everyday situations. According to Finke (1989:2), mental imagery is 'the mental invention or recreation of an experience that in at least some respects resembles the experience of actually perceiving an object or an event'. We 'see' images in our minds when we recall episodes from our past, when we plan things to do in the future, solve problems, or read an exciting novel. Historically this phenomenon has been very hard to study, but with present day eye tracking technology a new window for understanding the mind has opened. Mental imagery relies to a large degree on the same processes that are active during perception when we interact with the external world (Laeng et al. 2014; Richardson et al. 2009). In other words, the same activation occurs when we take in information through direct sensory input as when we create internal mental images in our minds. Current research in cognitive science has demonstrated that mental imagery is accompanied by spontaneous eye movements, and that these eye movements closely reflect the content and spatial layout of the imagined scene (Johansson et al. 2006).

Our research team has conducted several studies on mental imagery and measured the effects of internal images using eye-tracking methodology (Holsanova, Hedberg, & Nilsson 1999, Johansson, Holsanova & Holmqvist 2006, 2013). In the first scenario, viewers inspected a complex picture. Afterwards, they recalled the picture by orally describing it while looking at a blank screen. In the other scenario, viewers looked at a blank screen while listening to a verbal description of a scene and recalled the scene description by retelling it orally. Results from this study revealed that participants moved their eyes to appropriate spatial locations while describing the image from memory, while listening to the spoken scene description (that was never seen in the first place), and while retelling it from memory. Figures 3 and 4 illustrate the results from Johansson et al. (2006).
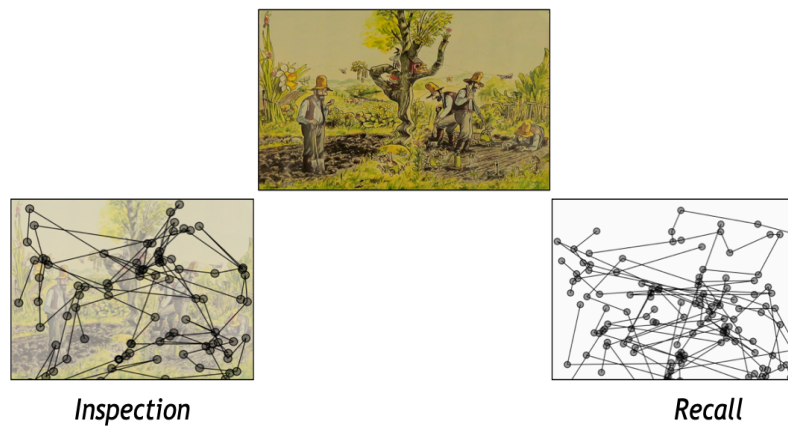
Figure 3. Inspection and recall of a complex picture. From Johansson et al. (2006).
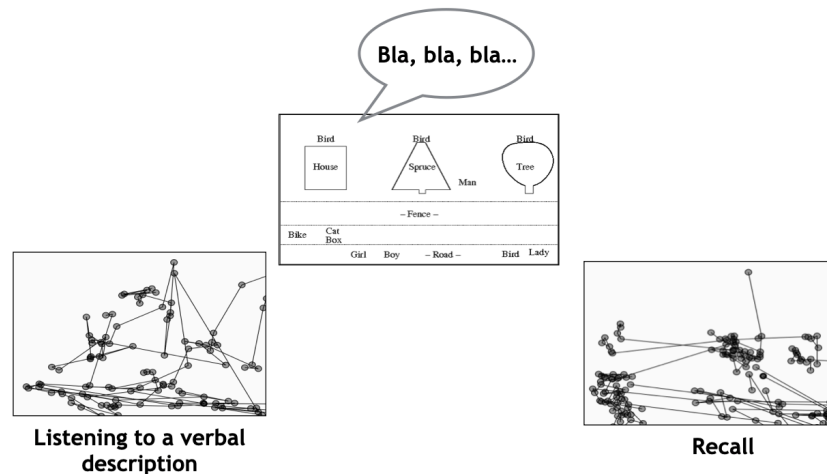

Figure 4. Listening to the spoken scene description and retelling it from memory. From Johansson et al. (2006).

The eye movement patterns closely reflected the content and spatial layout of the imagined scene. In their minds' eye, the viewers 'saw' the scene as if it were in front of them and 'painted' it with their eyes on the blank screen. The results also indicate that the effect was equally strong during recall, irrespective of whether the original elicitation was spoken or visual. Here is an obvious link to audio description: By using verbal descriptions we can evoke and stimulate the creation of vivid internal images for visually impaired and blind audiences.

But how does this work for people with visual impairment and for people who are congenitally blind? Can they form mental images? Contemporary research suggests that even people who are congenitally blind experience some kind of mental images, but that these differ from those of sighted people in several respects (Cattaneo and Vecchi 2011; see also Johansson 2016). When vision is impaired, the other senses become more important. Sound, touch, smell, and taste are a natural part of our sensory world, and can be used to complement or may even be a substitute for vision. For example, touch is used to discern details, shapes, and textures (Noordzij, et al. 2007), while movements, proprioception, hearing, and the sense of smell can be used to assess how different things are placed in relation to one's own body (Eimer 2004). To sum up, individuals with congenital blindness use other senses than sight to create rich mental models of their environments. Their mental imagery is often represented in a very spatial manner, and instead of being visual it is more dependent on embodiment and on haptic and motor imagery (e.g., Cattaneo & Vecchi 2011; Noordzij et al. 2007). Audio descriptions must therefore be adapted to the BVI audiences and their needs and preferences.

Guidelines and recommendations for the AD of art

How should a verbal description of an artwork be tailored for BVI audiences? In the following, we will have a look at what the guidelines say about the AD of art and focus on the following questions: How can AD help blind and visually impaired audiences to understand and experience art? Which ingredients must AD contain to evoke and stimulate the creation of vivid internal images? How can AD contribute to the aesthetic experiences of BVI audiences?

The challenge for audio describers is to produce creative, informative, vivid, and precise descriptions, condensed in short texts (Perego 2019). Audio description should respect the style and means of expression of the artist and at the same time meet the needs and preferences of the BVI recipients and adapt to them. It should enhance mental imagery and give the BVI audiences the means to engage with a work of art and immerse themselves in the mood and feelings that the artwork evokes. Thus, there are many challenges for audio describers, concerning both the language and the content of AD.

Clear and precise language

Several guidelines point out that AD of art should use clear and precise language (*Art Beyond Sight, ABS* 1996) and take into account the heterogeneous group of end users, not only in terms of blindness (total vs. partial) and educational background, but also in terms of familiarity with the special languages and contents of the arts (Perego 2019). Ideally, museum AD should be short yet informative, to the point and factual yet evocative, highly pertinent and intriguing, focused on content, but it should also avoid complex sentence structures as well as 'flowery' prose and an 'exuberant' vocabulary (Giansante 2015:9). This sets very high expectations which seem to be difficult to achieve. Perego (2019) analysed a selection of eighteen stand-alone ADs of works of art exhibited at the British Museum in London. The ADs were produced by VocalEyes and the scripts were created collaboratively by describers and curators, with the input from visually impaired testers. In the result of her corpus study, Perego found that the language of AD, although spoken, had the character of written expository texts, with long words, abstract verbs, complex syntax, technical terms, and the use of expert language. The museum ADs guaranteed vivid, imaginative, and diverse language but were more complex than advised, merging the features of narrative, descriptive, and informative texts.

Vivid details in order to evoke inner images

According to the guidelines for verbal description by *Art Beyond Sight* (ABS 1996), verbal description of art should contain vivid details and diverse language. It should present enough information so that listeners can form an image in their minds, and come to their own opinions and conclusions about a work of art. Perego gives some examples: 'The need to describe vividly requires the use of adjectives which are necessary to produce semantically rich noun groups (the great black bird's wings; its pink, mottled, un-feathered neck and head) and contribute to the visual intensity and meticulousness of the description' (Perego 2014, pp. 28–30). She notes that the production of a visually intense text is difficult, especially when there are constraints of time or space.

To illustrate vivid description in AD of art adopted for BVI audiences, we include two transcribed extracts from a student assignment at the Syntolksutbildningen (Audio Describer Training) at Fällingsbro Folkhögskola, Sweden. The assignment was to audio describe two artworks by Anders Zorn, *Midsommardans* (1897) and *Kaikroddaren* (1886), exhibited at the National Museum in Stockholm, and to record the audio descriptions. A time limit for each description was set at three minutes.

Eli Tistelö, a teacher in the Audio Describer Training program, instructed the students[2]. The recording was to include both basic facts about the work of art and about the artist (title,

---

[2] The full list of instructions in Swedish can be found in the Appendix.

artist, style/period, background/context, format, material) along with the audio description of the work. Eli instructed the students to keep the AD of the artwork separate from the facts about the artist and the painting in order to make it easier for the target audience to build an inner image with the help of the AD[3]. She commented further on the various aspects of AD: 'Give a summarising heading first, before details. Use a rich vocabulary to reflect mood and feeling. Describe posture and facial expressions. Try to stay close to the image, check what is actually visible. If you get a certain feeling, try to figure out what in the image gives you this feeling and describe it. Remember to leave associations, fantasies, and free thought space to the target audiences. Find verbs other than 'we can see', use, for example, 'there is'. When recording your AD, try to vary your vocal delivery by using a different tempo, stress, and intonation'. Students were provided with a number of questions as guidance for their formulations of AD:

- What does the image represent? (People, things, environments, events)
- How is the image composed? Is there a background / foreground? Is there a clear direction? Is there a centre?
- Is there light and darkness in the picture? Where do the light / shadows come from?
- Is the image in colour or black and white? And what is the nature of the colour? (e.g., strong, greyish, luminous, pale, …)
- How is the picture painted? (e.g., with coarse brush strokes, depicted exactly as in a photo, greatly simplified, blurred, …)
- Are there symbols in the picture?
- Does the image evoke any emotion in you? Feel free to try to include feeling in the description and think about what in the image produced the feeling. (e.g., A room with dim lighting. Flowers in bright, happy colours. A solitary umbrella.)

Example 2:
Midsommardans [Midsummer Dance], Zorn 1897, student JPS (extract 01:03–2:43)
'The subject is a log dance in Dalarna. The dance is in full swing and the dancers are swirling around on the grass in front of a brown log cabin. In front there are men playing the violin. In the background you can see a maypole of the type that brings to mind Dalarna. And you see the gables of a red cottage with white corners peeking out. The people in the painting wear folk costumes. The women have black ankle-length skirts, white blouses that cover their entire arms. They wear white kerchiefs with discreet floral patterns, and they have red bodices that extend like braces across the back and over the shoulders and pick up in the front where the bodice covers the bust. Men wear brown breeches. They have white shirts with brown vests and black short jackets on top of them. And on their heads they wear black hats. These are reminiscent of a flatter round Stetson model. It is light outside. Not a cloud in the sky. But there is a faint haze over the picture. It is as though there is rain waiting over the horizon. The light shines in from the right and is reflected on the gables of the red cottage. This is an image that displays a collective energy. Here I feel that people shift between maintaining tradition and correct behaviour at the same time as they indulge in a wild game.'

Example 3:
Kaikroddare [Turkish Boatman in the Constantinople Harbour], Zorn 1886, student AK (extract 00:55–3:00)
'The painting depicts rowers out on the Bosphorus. A caique is a kind of long, narrow, light, kayak-shaped rowing boat. In the foreground is a male rower captured in the picture as if the artist himself

---

[3] The extracts in Example 2 and 3 contain only the AD part of the presentation.

was sitting in the boat. And the rower has outstretched muscular arms and he is just about to pull the oars towards his body. He is wearing a long white garment, with short sleeves, slightly rolled up. On his head he has a white turban. He has a dark beard and moustache. And his narrow face is turned slightly to the side towards a spectator. Diagonally behind him another caique rower approaches in oncoming traffic, and he has two ladies as passengers who can be glimpsed under a red parasol. And even further away you can see another boat on its way towards us. To the right in the picture you can discern a lively harbour with ships with high masts and smoke from chimneys and, closest to us, a larger ship. The scene is shrouded in an early hazy evening light that falls in from the left corner of the painting and barely breaks through the cloud cover. Pink and pale yellow from the sun are reflected in the water's surface where you can also see the shadow of the oncoming caique and a faint reflection of the two ladies under the red parasol. There is an overall contrast between the heavy grey sky and the evening sun, between the grey lively harbour far away in the background and the shimmering water surface near us where the caique boats with their passengers glide forward with the help of the movements of the rower's oars. Zorn lets us mostly guess the motif with the help of different fields of colour. The caique rower's face is not detailed, but the grim, worn features are prominent through lines and shadows. In the sky you can see elements of clear brushstrokes and how the thin watercolour paint has flowed and dried.'

Aesthetic experience, embodiment, and immersion

According to Neves (2016), listeners should be given the means to engage with, and immerse themselves in, an art exhibit. Immersion implies deep mental involvement in something. One important element is to mention mood and feeling, another is to describe facial expressions, glances, gestures, and body postures depicted in a painting or a sculpture. A description of facial expressions, glances, gestures, and body postures can – in combination with a description of feelings and emotions – create a better context not only for understanding and imagination but also for empathy and engagement. When listening to AD, the target audiences can then recognise the described bodily expressions, emotions, and feelings from their own previous physical experiences and become engaged through embodiment. The ABS guidelines (1996) also recommend referring to other senses as analogues for vision when describing surfaces, for instance, by describing the tactile quality of a material.

Example 4 illustrates some of these factors. It is a one minute long audio description of a sculpture from an exhibition at Sundsvall Museum[4]. After identifying the main subject, the audio describer mentions the material of the work of art and its tactile quality and includes a vivid description of the child, its clothes, boots, body posture, and facial expression.

Example 4: *In Oversized Shoes* by Maria Eriksson. Sundsvall Museum. Audio description by Ingela Hofsten.
'A sculpture depicting a child who appears to be two years old judging by its body and height. It is made of black ceramics that feels rough to the fingers. The child is wearing something that looks like linen and underpants and a hat with a small brim. The clothes shimmer in metallic green. On both feet the child has golden boots, which have ended up on the wrong feet. The child looks up at the viewer with a grumpy face, the corner of his/her mouth pulled down, and with arms crossed. It's easy to imagine that the child just said "I don't want to!"'.

According to theories of embodied cognition, our bodily experiences and physical interactions with the outside world both influence our thinking and help us think. Current research shows that our bodily experiences affect how we interpret, evaluate, and understand visual and linguistic information. They facilitate language comprehension both when we are listening to spoken

---

[4] Sundsvall Museum, Region Västernorrland, presented thirty-one works of art from twenty-six artists at the exhibition 'Art Injection III – Outlook' from 10 October, 2020 through 3 January, 2021. Ten works of art have been audio described by Ingela Hofsten for about one minute per art work. The AD shown in Example 4 has been translated from Swedish by the author. The Swedish original AD can be found in the Appendix.

language descriptions (Bergen et al. 2007) and when we read text (Zwaan & Taylor 2006; Hauk, Pulvermuller et al. 2004). When we conceptualise a scene or listen to a spoken scene description, we conduct a mental simulation of bodily experiences of actions on objects and interactions with other persons. For instance, hearing the word 'dog' automatically triggers the sensorimotor, affect, and mental states associated with our experiences of dogs (e.g., what dogs look like, how they move, feel, etc.). The idea is that embodied processes are encoded into the knowledge system during the initial experience and later repeated via simulation mechanisms in response to the original stimulus (e.g., seeing a dog) or associated stimuli (e.g., hearing the word *dog*) (Schendan 2012).

ABS guidelines even go a step further. Apart from listening to descriptions of gestures, body postures, and facial expressions, ABS (1996) encourages the end users to imitate the depicted figure's pose and gestures. 'Since everyone is aware of his or her own body, this activity provides a concrete way of understanding difficult poses depicted in the painting.' This encouragement to create understanding through re-enactment is in line with current research in neuroscience and social psychology. There is increasing evidence to support the strong links between language and action. The research on imitation, empathy, and mirror neurons has shown that we have a common representational format for action and perception that facilitates imitation. And imitation and mimicry facilitate empathy (Iacoboni 2009).

Frames of reference, analogies, and the meeting of minds

Since visitors have different frames of reference based on different backgrounds and experiences, some descriptions could be difficult to understand. For instance, the description 'the cottage is completely dilapidated' could be considered concise and effective from the audio describers' point of view, but the question is whether the users can associate and imagine all the details included in these concepts? Also, abstract concepts like 'horizon' or 'shadow' can be difficult to describe and to grasp. It is therefore very important to discuss our understanding and associations in order to achieve a 'meeting of the minds' (Holsanova 2016). ABS guidelines recommend in this respect describing intangible concepts with analogies (ABS 1996). In the following example, the visual phenomena of a shadow is explained for congenitally blind persons by using analogies.

Example 5: ABS and Paula Gerson, 'The Building Blocks of Art'
'Use of light and shadow in a painting can be explained by referring to the feeling one has when sitting in front of a window on a sunny day. The parts of the face and body that feel the warmth are said to be in the light. Those parts not being warmed by the sun are said to be in shade or shadow. To understand the concept of a cast shadow, imagine yourself standing in the kind of shower where the water comes out in a fairly narrow spray. As you stand in front of this spray, the front of your body gets wet, but not your back. If the water were a light source, the front of your body would be highlighted, and the back would be in shadow. Additionally, because the front of your body blocks the water, there would be a spot on the shower floor behind you where water does not fall. If the water were a light source, your body would block the flow of light, and the light would not reach the area of the shower floor behind you. The dark area behind you is called a cast shadow.'

Evaluative aspects, associations, and interpretation

Interpretation and aesthetic experience of the artwork is not only based on what we see but also on what we know, associate, feel, and experience. It is, however, tricky since the audio describers are not supposed to express personal associations and feelings. ABS guidelines (1996) recommend the use of objective references rather than those that might affect a blind person's point of view. An 'objective' AD is, however, problematic considering the range of individual and contextual factors that influence our way of describing a scene. ABS further requires that the BVI audiences must have the freedom to interpret the work of art for themselves. In her analysis of museum AD, Perego

(2019) found that 'evaluative/emotive adjectives are generally used in moderation in order to limit appraisal, to avoid conveying an explicit or implicit degree of judgment or bias, or positive or negative connotation, which would restrict the interpretative freedom of the people with visual impairment'. In this respect, it is obvious that audio describers need feedback from the blind and visually impaired users in order to adapt and optimise AD for the target audiences and to find out what preferences, perceptions, and understanding they have. It is therefore important to review verbal descriptions with visually impaired advisors to find an effective language, the sought-after clarity and length of the descriptions, and an appropriate pace. Also, conversations about an artwork between sighted and blind audiences can be very useful in order to discover not only what can be seen but also feelings, memories, associations, interpretations, and to achieve a 'meeting of minds'.

Multi-sensorial access to art

In previous sections, we focused on auditory solutions to improve accessibility of visual art in museums in the form of audio description and audio guides. In addition, tactile solutions have been developed and used in art museums, e.g., text and signs in Braille, tactile maps for orientation, and the possibility of touching the artworks or models. For instance, the National Museum in Stockholm introduced tactile images and auditive experiences already in the 1990s (Rangner Jacobsson 2021). Moderna museet in Stockholm is exploring tactile means by introducing experiential practices and workshops for visually impaired visitors, offering a toolbox and materials to create artworks, and touchable models and details from a painting (Hillberg & Wernström-Pitcher 2021). Researchers in inclusive technology Matthew Butler, Leona Holloway, and Kim Marriott from Monash University, Australia, are currently working with gallery staff, artists, and end users to create accessible versions of artworks. The researchers have developed a collection of accessible works and explored a range of technologies, including 3D printing, 3D modelling and scanning, laser cutting, and touch screen technology. They are currently evaluating whether these accessible versions of visual art are meaningful and valuable to blind and low vision visitors.

However, despite the use of single modality solutions, such as audio description or tactile graphics, blind and visually impaired people still face challenges to experiencing and understanding visual art independently. Therefore, some museums use a combination of 3D prints or relief images and audio description to enable both tactile and auditive experience and to stimulate the senses hearing, touch, and sometimes even smell. Multi-sensorial access increases the experience – as has been underlined in the ABS guidelines and elsewhere. Therefore, an art exhibition should incorporate sound in creative ways and visitors should be allowed to touch the artworks. This would give the BVI visitors an immediate, personal experience and the possibility of exploring an artwork at their own pace. An alternative is touchable materials, such as reproductions, samples of materials (canvas, clay), tools (paintbrushes, chisels, and hammers), and replicas of the objects depicted in an artwork. Yet another option is to offer tactile illustrations of artworks. Relief images in conjunction with AD provide additional information for the visitors. Finally, for a better aesthetic experience and enjoyment, voice quality and vocal delivery is important. It adds an additional dimension to the atmosphere, mood, and feeling.

Cavazos Quero et al. (2021a) from Sungkyunkwan University in South Korea developed a prototype of an interactive multimodal guide that uses audio and tactile modalities to improve the autonomous access to information and experiences of visual artworks at art museums. The prototype is composed of a touch-sensitive 2.5D artwork relief model that can be freely explored by touch. Blind and visually impaired users can access localised verbal descriptions and audio by performing touch gestures on the surface while listening to themed background music being played simultaneously. Results from a usability survey indicate that this particular multimodal approach is easy to use and improves confidence and independence when exploring visual artworks. The researchers also propose a multi-sensory colour code system that uses sound and scent to represent colours (Cavazos Quero et al. 2021b).

Summary and conclusions

This article aimed at answering the question of how we can make visual art accessible for audiences with visual impairment and blindness (BVI). We claim that audio description (AD) helps target audiences access, better understand, and enjoy the details of a visual art experience.

Research on the dynamic processes of image perception and image description uncovers the creativity and complexity of meaning-making and the interpretation of images. Meaning-making is an interactive meeting between the recipient, the image, and the situational context (Holsanova 2014b). First, the visual appearance of the image serves as a starting point for the process of meaning-making. Second, the recipients mediate, using their personal characteristics, the perception and interpretation of the work of art. Third, the context in which the image is displayed, perceived, and interpreted plays an important role for meaning-making (Holsanova 2021). The way sighted viewers perceive complex images can serve as an inspiration for AD practices. Research on mental imagery shows that by using vivid verbal descriptions we can evoke and stimulate the creation of vivid internal images for visually impaired and blind audiences. Mental imagery is very important for understanding and enjoyment.

The task of the audio describer is to describe the visual appearance of an artwork, to stimulate the creation of inner mental images that lead to understanding and enjoyment in the target audiences, and that contribute to their aesthetic experience of the artwork. At the same time, audio describers are not supposed to express evaluations, personal associations, and feelings. Instead, they must leave the space of interpreting the work of art to the BVI audiences.

According to the guidelines, AD should be short yet informative and use clear and precise language providing vivid details in order to evoke inner images. It is recommended that a summarising heading be given first, before details, and that a rich vocabulary be used to reflect mood and feelings. Target audiences should be given the means to engage with and immerse themselves in an art exhibit. AD of facial expressions, glances, gestures, and body postures can – together with a description of feelings and emotions – create a better context not only for understanding and the imagination but also for empathy and engagement. Embodiment plays an important role in this respect. Since visitors have different frames of reference based on different backgrounds and experiences, some descriptions could be difficult to understand. The guidelines suggest describing abstract concepts with the help of analogies.

Apart from use of single modality solutions such as audio description, a multi-sensorial access is recommended since it enhances the experience. Currently, several prototypes are being developed to enable both tactile and auditive experiences and to stimulate the senses of hearing, touch, and sometimes even smell.

Finally, there is an urgent need for systematic research on the reception of AD. From a reception perspective we need to know the following: How do BVI recipients perceive, understand, and experience the AD of art? How do they imagine the content of the described artwork and do they feel involved? Which preferences they have concerning the length of the AD so that the descriptions do not overtax their cognitive capacities, disrupt their attention, or spoil their enjoyment (Perego 2019)? A reception perspective on AD is the focus of a three-year interdisciplinary research project 'How Blind Audiences Receive and Experience Audio Descriptions of Visual Events' being conducted at Lund University in Sweden and financed by FORTE (the Swedish Research Council for Health, Working Life and Welfare). Regularly held workshops with a reception perspective on AD are also part of the activities of the cooperative initiative 'Audio Description for Accessible Communication' (ADACOM) at Lund University, led by the author of this article.

References

Buswell, Guy Thomas. 1935. *How People Look at Pictures: A Study of the Psychology of Perception in Art.* Chicago: University of Chicago Press.

Cattaneo, Z. & Vecchi, T. 2011. *Blind Vision. The Neuroscience of Visual Impairment.* Cambridge, MA: MIT Press.

Cavazos Quero, Luis, Jorge David Iranzo Bartolome, and Jundong Cho. 2021a. 'Accessible Visual Artworks for Blind and Visually Impaired People: Comparing a Multimodal Approach with Tactile Graphics'. *Electronics* 10(3)(January):297. DOI: 10.3390/electronics10030297.

Cavazos Quero, Luis, Jorge David Iranzo Bartolome, and Jundong Cho. 2021b. 'Multi-Sensory Color Code Based on Sound and Scent for Visual Art Appreciation'. *Electronics* 10(14)(July):1696. DOI: 10.3390/electronics10141696.

Finke, R. A. 1989. *Principles of Mental Imagery*. Cambridge, MA: MIT Press.

Giansante, L. 2015. 'Writing Verbal Descriptions for Audio Guides.' *Art Beyond Sight: Museum Education Institute*. Retrieved from http://www.artbeyondsight.org/mei/verbal-description-training/ writing-verbal-description-for-audio-guides/

Hättich, Achim & Martina Schweizer. 2020. 'I Hear What You See: Effects of Audio Description Used in a Cinema on Immersion and Enjoyment in Blind and Visually Impaired People'. *British Journal of Visual Impairment* 38(3): 284–98.

Henderson, John M. 2007. 'Regarding Scenes'. *Current Directions in Psychological Science* 16(4): 219–22.

Hillberg, Malin & Anita Wernström-Pitcher. 2021. 'Moderna Museet — Guided tours for visual impaired'. Presentation at the symposium 'Exploring visual art auditively. Art, language, and sounds in museums and culture'. Inter Arts Center Malmö, May 2021.

Holsanova, J. 2021. 'The Cognitive Perspective on Audio Description: Production and Reception Processes'. In *Handbook of Audio Description,* edited by C. Taylor and E. Perego. London: Taylor & Francis.

Holsanova, J. 2014. 'In the Eye of the Beholder: Visual Communication from a Recipient Perspective'. In *Visual Communication*. [Handbooks of Communication Science, HoCS], edited by David Machin, Chapter 14, 331–55. Berlin: De Gruyter.

Holsanova, Jana. 2020. 'Att beskriva det som syns men inte hörs. Om syntolkning'. *Humanetten*, 44: 125–46, DOI: https://doi.org/10.15626/hn.20204406.

Holsanova, J. 2019. *Bildbeskrivning för tillgänglighet*. Myndigheten för tillgängliga medier, rapport nr. 6.

Holsanova, J. 2016. 'Kognitiva och kommunikativa aspekter av syntolkning'. In *Syntolkning – forskning och praktik,* edited by J. Holsanova, M. Andrén, and C. Wadensjö. Lund: Lund University Cognitive Studies 166/ Myndigheten för tillgängliga medier, rapport nr. 4: 17–27.

Holsanova, J. 2011. 'How we focus attention in picture viewing, picture description, and during mental imagery'. In *Bilder, Sehen, Denken*, edited by K. Sachs-Hombach and R. Totzke. Herbert von Halem Verlag: Köln, 291–313.

Holsanova, J. 2008. *Discourse, Vision, and Cognition*. Amsterdam and Philadelphia: John Benjamins.

Holsanova, J. 2001. *Picture Viewing and Picture Description: Two Windows on the Mind*. PhD diss. Lund University Cognitive Studies 83.

Holsanova, J., B. Hedberg, and N. Nilsson. 1999. Visual and Verbal Focus Patterns when Describing Pictures'. In *Current Oculomotor Research: Physiological and Psychological Aspects,* edited by Wolfgang Becker, Heiner Deubel, and Thomas Mergner. New York: Plenum.

Holsanova, J., R. Johansson, and V. Lyberg-Åhlander, 2020. 'How the Blind Audiences Receive and Experience Audio Descriptions of Visual Events – A Project Presentation', In *Book of Extended Abstracts. 3rd Swiss Conference on Barrier-free Communication*, 39-41.

Iacoboni, M. 2009. 'Imitation, Empathy, and Mirror Neurons'. *Annu. Rev. Psychol*. 60:653-70. doi: 10.1146/annurev.psych.60.110707.163604.

Johansson, R. 2016. 'Mentala bilder hos seende och blinda'. (Mental images in sighted and blind). In *Syntolkning – forskning och praktik,* edited by J. Holsanova, C. Wadensjö, and M. Andrén. Lund University Cognitive studies 166/MTM:s rapportserie nr 4, 29-38.

Johansson, Roger, Jana Holsanova, and Kenneth Holmqvist. 2006. 'Pictures and Spoken Descriptions Elicit Similar Eye Movements During Mental Imagery, both in Light and in Complete Darkness'. *Cognitive Science* 30(6): 1053–79.

Johansson, R., J. Holsanova, and K. Holmqvis. 2013. 'Using Eye Movements and Spoken Discourse as Windows to Inner Space. In *Conceptual Spaces and the Construal of Spatial Meaning. Empirical evidence from human communication,* edited by C. Paradis, J. Hudson, and M. Magnusson, 9–28. Oxford: Oxford University Press.

Laeng, B., I.M. Bloem, S. D'Ascenzo and L. Tommasi 2014. 'Scrutinizing Visual Images: The Role of Gaze in Mental Imagery and Memory'. *Cognition* 131, 263–83.

Neves, J. 2016. 'Enriched Descriptive Guides: A Case for Collaborative Meaning-Making in Museums'. *Cultus* 9(2), 137–53.

Noordzij, M.L., S. Zuidhoek, and A. Postma. 2007. 'The Influence of Visual Experience on Visual and Spatial Imagery'. *Perception* 36, 101–12.

Nordqvist, Sven. 1990. *Kackel i grönsakslandet*. Bromma: Opal.

Perego, Elisa. 2019. 'Into the Language of Museum Audio Descriptions: A Corpus-Based Study, *Perspectives* 27:3, 333-49. DOI:10.1080/0907676X.2018.1544648.

Rangner Jacobsson, Jeanette. 2021. 'Access to visual art through tactility and audio description. Nationalmuseum 1994-2019'. Presentation at the symposium 'Exploring visual art auditively. Art, language, and sounds in museums and culture'. Inter Arts Center Malmö, May 2021.

Richardson, D. C., G.T.M. Altmann, M. J. Spivey, and M.A. Hoover. 2009. 'Much Ado about Eye Movements to Nothing: A Response to Ferreira et al.: Taking a New Look at Looking at Nothing'. *Trends in Cognitive Science* 13(6): 235–36.

Remael, A., N. Reviers, and G. Vercautere, eds. 2015. 'Pictures Painted in Words: ADLAB Audio Description Guidelines'. Trieste: EUT.

Salzhauer Axel, Elisabeth, Virginia Hooper, Teresa Kardoulias, Sarah Stephenson Keyes, and Francesca Rosenberg. 1996. 'Making Visual Art Accessible to People Who Are Blind and Visually Impaired'. New York: Art Education for the Blind, Inc.

Schendan, H.E. 2012. 'Semantic Memory'. *Encyclopedia of Human Behavior, 2nd ed.*: 350–58.

Schoene, Janneke. 2021. 'Transmediating meaning? Artistic audio guides and aesthetic experience'. Presentation at the symposium 'Exploring visual art auditively. Art, language, and sounds in museums and culture' Inter Arts Center Malmö, May 2021.

Solso, R. L. 1994. *Cognition and Visual Arts*. Cambridge, MA:. MIT Press. A Bradford Book.

Acknowledgement

Appendix: Examples and instructions in Swedish

1. Sundsvall Museum, Region Västernorrland: I för stora skor av Maria Eriksson. Syntolkning: Ingela Hofsten

I för stora skor av Maria Eriksson. *'En skulptur som föreställer ett barn som ser ut att vara i tvåårsåldern till kropp och längd. Den är gjord i svart keramik som känns skrovlig mot fingrarna. Barnet har på sig något som ser ut att vara linne och underbyxor och en hatt med litet brätte. Kläderna skimrar i metalliskt grönt. På fötterna har hen guldfärgade stövlar, som hamnat på fel fot. Hen tittar upp mot betraktaren med ett truligt och trumpet ansiktsuttryck, neddragna mungipor och armarna i kors. Det är lätt att tänka sig att hen just har sagt 'Jag vill inte!'.*

https://www.taltidningenvasternorrland.se/2020/v49/syntolkning-av-region-vsternorrlands-utstllning-konstinjektion-iiiutblick/
https://www.rvn.se/sv/Utveckling/Kultur-och-bibliotek/konst-vasternorrland/konst-i-landstingets-lokaler/syntolkad-konst/

2. Syntolksutbildningen, Fällingsbro folkhögskola, Region Örebro
Lärare: Eli Tistelö; Kursansvarig: Lotta Lagerman

Instruktion
Nu är det dags att syntolka konst! Det är ju olika vad vi har för förhållande till konst. Kanske är det ett nytt och okänt område för dig, kanske är det en hemma-zon? Kolla på dina anteckningar från min genomgång av konst-syntolkning och gör dig redo för att beskriva en målning. Veckans uppgift är att göra en inspelad syntolkning av en bild. I år är det 100 år sedan Anders Zorn dog. Med anledning av det visar Nationalmuseum ca 150 verk av Zorn och Zorn blir också vårt tema i år. Var och en får arbeta med ett verk av konstnären. Alla verken finns på Nationalmuseum i Stockholm. Den bild som fallit på din lott får du i chatten.

Övningen handlar om att
- söka fakta och använda dem för att skapa en tillgänglig berättelse
- beskriva en bild med ord
- skriva en text
- välja och välja bort och förhålla sig till en given tid
- göra en inläsning

Börja med att söka på konstnären. Läs om honom och försök hitta om det finns något skrivet om den bild du ska jobba med. Samla lite fakta och kanske några värdefulla ord. I andras beskrivningar av verket kan du ibland hitta bra ledtrådar till hur du själv vill beskriva det. Sedan ska du skriva en kort text och läsa in den. Inläsningen får **max vara 3 min** lång och jag vill att du ska försöka att läsa in **utan** att det låter som att du läser.

Ta tid på dig själv när du läser. Stryk eller lägg till tills du hittat rätt längd på texten. När vi syntolkar konst ingår det oftast att också kort berätta vem som gjort verket, vad det heter samt kanske ytterligare några korta fakta. Det gäller att välja omsorgsfullt och att välja bort mycket. Det gäller också att försöka formulera sig så att lyssnare med olika bakgrund kan förstå. Undvik bild-tekniska termer som lyssnaren kanske inte är bekant med. Gör alltså två delar i din beskrivning: Ett avsnitt om Zorn och **om** bilden och ett med syntolkning av bilden. I själva syntolkningen av bilden är det ofta bra att börja med en sammanfattning snarare än att nämna för många detaljer/komponenter. Försök att först hitta en 'enkel' 'rubrik t ex: 'Bilden föreställer 10 barn som leker på en äng.'

I den här övningen vill jag att du använder listan med frågor. Eftersom bilderna är olika kommer inte alla frågor att passa på alla bilder, så var öppen och välj de frågor som just du har nytta av. Väv gärna ihop svaren på frågorna, så att du får flyt när du pratar. (Ex: Salvador Dalí föddes i början av 1900-talet i Spanien och målade surrealistiskt…). Kanske vill du också välja att säga något som inte finns i frågelistan, men väg dina ord på guldvåg och kom ihåg att *allt* går inte att ta med. Sträva efter att <u>inte blanda</u> syntolkning av verket med fakta om konstnären och bilden, då blir det lättare att bygga sig en inre bild med hjälp av din syntolkning.

Här är frågorna:
1. Grundfakta om verket
Vad heter målningen?
Vad heter konstnären?
Vad heter du som pratar?
Vad är det för format på målningen?
I vilket material är den gjord? (ex: olja på duk)
Vilken tid kommer den ifrån?
Vill du berätta något mer? (om inramning? Ingår det i en serie?)

2. Kort om konstnären
Varifrån kommer konstnären?
Vilken stil/konstinriktning tillhör verket/konstnären? (ex surrealismen)
Säg gärna några förklarande ord om stilen.
Vill du berätta något mer om konstnärens bakgrund eller sammanhang?

3. Syntolkning av verket
Är det ett föreställande eller icke-föreställande verk?
Om det är föreställande. *Vad* föreställer bilden? (Människor, saker, miljöer, skeenden)
Hur är bilden komponerad? (Finns det bakgrund/förgrund? Finns det en tydlig riktning? Finns det ett centrum?)
Finns det ljus och mörker i bilden? Varifrån kommer ljuset/skuggorna?
Är bilden i färg eller svartvit? Och hur är färgen? (Kulörstark, gråaktig, självlysande, blek…)
Hur är bilden målad? (Med grova penseldrag, exakt avbildande som ett foto, starkt förenklat, suddig…)
Finns det symboler i bilden?
Väcker bilden någon känsla hos dig? Försök gärna ta in känslan i beskrivningen och fundera över *vad* i bilden som gav känslan. (Ex: Ett rum med <u>dämpad</u> belysning. Blommor i <u>klara, glada</u> färger. Ett <u>ensamt</u> paraply.)

Nedan finns två transkriberade avsnitt från två inspelade syntolkningar av Anders Zorns verk Midsommardans (1897)och Kaikroddaren (1886). Syntolkningen gjordes av två kursdeltagare i Syntolksutbildningen (2021), JPS och AK.

01:03–2:43 Midsommardans, Zorn 1897, syntolkat av kursdeltagare JPS, Syntolksutbildningen
'… Motivet är en logdans i Dalarna. Dansen pågår för fullt och de dansande virvlar runt på gräset framför en brun timmerstuga. Där framför sitter män och spelar fiol. I bakgrunden kan man se en midsommarstång av den typ som för tankarna till Dalarna. Och man ser gavlarna på en röd stuga med vita knutar som där tittar fram. Människorna i målningen de bär folkdräkt. Kvinnorna har svarta ankellånga kjolar, vita blusar som täcker hela armarna, de bär vita hucklen med diskreta blommönster, de har röda livstycken som sträcker sig som hängslen över ryggen fram över axlarna och tar vid på framsidan där livstycket täcker bysten. Männen, de bär bruna knäbyxor. De har vita

skjortor med brun väst och överst en svart kort jacka. Och på huvudet så bär de svarta hattar. De påminner om en plattare rund Stetson modell. Det är ljust ute. Inte ett moln på himlen. Men det ligger ett svagt dis över bilden. De är som att ett väntande regn ligger brotom horisonten. Ljuset slår in från höger och det speglar sig på den röda stuggaveln. Det är en bild som har en samlad energi. Här upplever jag att människorna spelar mellan att upprätthålla tradition och korrekt uppförande samtidigt som de ger sig hän åt den vilda leken'.

00:55–3:00 Kaikroddare, Zorn 1886, syntolkat av kursdeltagare AK, Syntolksutbildningen

'... Målningen föreställer roddare ute på Bosporen. Och en kaik är nån sorts långsmal lätt roddbåt, lite kajak-formad. I förgrunden finns en manlig roddare fångad på bild som om konstnären själv satt i båten. Och kaikroddaren har sträckta muskulösa armar och han ska precis till att dra årorna mot kroppen. Han är klädd i en vit särk, med korta, lite uppkavlade ärmar, på huvudet så har han en vit turban. Han har mörkt skägg och mustasch. Och det här smala ansiktet är vänt lite med sidan mot åskådaren. Snett bakom honom så kommer en annan kaikroddare i mötande trafik och han har två damer som passagerare som skymtar under ett rött parasoll. Och ännu längre bort kan man skönja ännu en båt på väg åt vårt håll. Till höger i bilden så kan man ana ett livlig hamn med skepp med höga master och rök från skorstenar och närmast oss ett större fartyg. Scenen den är insvept i ett tidigt disigt kvällsljus som faller in från målningens vänstra hörn och knappt bryter igenom ett molntäcke. Rosa och blek gult från solen reflekteras i vattenytan där man också kan se den mötande kajkens skugga och en svag spegelbild av de här två damerna under det röda parasollet. Det finns en övergripande kontrast mellan den gråtunga himlen och kvällssolen, mellan den gråa livliga hamnen långt där borta i bakgrunden och den skimrande vattenytan nära oss där kajkerna med sina passagerare glider fram, med hjälp av roddarnas årtag. Zorn låter oss mest ana motivet med hjälp av olika färgfält. Kajkroddarens ansikte är inte detaljerad men de här bistra tärda dragen är framträdande genom streck och skuggor. På himlen ser man inslag av tydliga penseldrag och hur den tunna akvarellfärgen runnit och torkat'.