

# How the blind audience receive and experience audio descriptions of visual events - a project presentation

Dr. Jana Holsanova, Associate Professor  
Dr. Roger Johansson, Associate Professor  
Dr. Viveka Lyberg Åhlander, Associate Professor

Lund university  
Sweden



Holsanova, ARSAD 2019



# How the blind audiences receive and experience audio descriptions of visual events

- researchers from Lund University, Sweden
- awarded a grant from FORTE (Swedish Research Council for Health, Working Life and Welfare)
- three years project (2019-2021)
- interdisciplinary research & various methods: cognitive science, linguistics, psychology, logopedics
- include persons with visual impairment and blindness in the research
- co-operation with AD training (Fellingsbro), Swedish authorities and interest organisations



Photo of Roger Johansson, Jana Holsanova and Viveka Lyberg Åhlander

# Research initiatives in Sweden

The first Swedish research initiative started nine years ago in the form of workshops (organized by Jana Holsanova and Cecilia Wadensjö) 2010-2014

At these workshops, researchers, AD educators, audio describers, visually impaired recipients of AD, and representatives from government agencies and disability organizations discussed critical topics surrounding the production and reception of AD.

These research initiatives resulted in the book "*Audio description - Research and Practices*" (Holsanova et al., 2016).

Members of the present research group have previously conducted relevant basic research on:

- how information from verbal descriptions is mentally visualized in time and space (Holsanova et al. 1999, Johansson et al. 2006, 2011, 2013)
- how the interaction between attention, cognition and event segmentation when visual scenes are verbally described (Holsanova 1999, 2001, 2008, 2011, 2014a,b) and
- how the relationship between voice quality and verbal comprehension affects comprehension (Lyberg-Åhlander et al., 2015, Haake et al., 2014)
- cognitive aspects of AD, reception of AD, preferences of the end users (Holsanova 2015, Holsanova et al. 2015, 2016, Strukelj 2016, Holsanova, Johansson, Blomberg & Gårdenfors, in prep.)

# Cognitive approach to AD

AD makes the contents of visual scenes accessible for blind and visually impaired audiences and offers a richer and more detailed understanding and enjoyment

AD evokes inner mental images and enhances meaning-making for the non-sighted audience (Holsanova 2015, 2016)

There is currently a great need for establishing a scientific foundation for AD (e.g., Kruger & Orero, 2010) and a surge for more research on:

- how particular event descriptions are actually received (Holsanova 2015, Holsanova et al. 2016, Johansson, 2016)
- how both blind and sighted audiences perceive, imagine, segment and remember properties of the unfolding events (Johansson, 2016)
- how the voice quality, speech rate and pausing of the audio describer affects understanding and enjoyment (cf. Lyberg-Åhlander, 2015, Haake et al., 2014)

# Mental imagery

There is substantial evidence that people without sight indeed can use mental imagery similarly as sighted people when they internally simulate a course of events (Johansson, 2016; Cattaneo & Vecchi, 2011).

There are also fundamental differences in the underlying processes (Johansson, 2016; Cattaneo & Vecchi, 2011).

Blind audiences appear to depend more on haptic and motor imagery when they imagine and/or recall spatiotemporal information (e.g. Cattaneo & Vecchi, 2011; Noordzij et al., 2007).

Such processes are sequentially structured and cognitively more demanding than when sighted people engage in corresponding activities and severely limit how much concurrent information one can process (Postma et al., 2006).

While the sighted also generate mental images sequentially (Johansson et al., 2006), the visual modality typically allows for simultaneous perception of many objects, which enables them to visualize much larger “information chunks” in each step.

This is crucial to consider in communication between the sighted and the blind and fundamental for an audio describer when selecting what to describe as well as how to verbalize it.

# Event cognition

Event cognition is the study of how we perceive, conceive, remember and communicate about such events and *event segmentation* is our ability to conceive the boundaries of when an event starts and ends (e.g. Radvansky & Zacks, 2014).

All stories have temporal and spatial dimensions connected to its characters and the actions they are involved in. When watching movies, we thus segment information into meaningful events according to changes in the spatiotemporal reference frame (e.g. Zwaan & Radvansky, 1998).

Reception of film narratives is a highly complex and multimodal process, where the audience need to keep track of who did what, with whom, where, when and why for every event in the story, including the spatiotemporal circumstances, and constantly update this information in accordance with how the narrative unfolds (Vercauteren & Remael 2014, ADLAB 2014).

Event segmentation abilities have been critically linked to the updating of working memory and to the integration of novel information with associated long-term memories (e.g., Radvansky & Zacks, 2014; Zwaan & Radvansky, 1998).

If one segments well, this improves comprehension and saves valuable cognitive resources (Boltz 1992).

Event segmentation capabilities are fundamental for comprehending and remembering narratives for both the sighted and the blind.

To date, virtually nothing is known about how, or if at all, the sighted and the blind differ in how they perceive and remember events.

# Aural properties of spoken descriptions

Research shows that the quality of the speaker's voice affects the listening effort as well as listener's attitudes and comprehension of the spoken message (e.g. Lyberg-Åhlander, et al., 2015; Rogerson & Dodd 2005).

Listeners process the message more slowly and frequently miss content bearing words when listening to a dysphonic (hoarse) voice. This often has direct effects on their comprehension, but it also increases cognitive load and hence, the listening effort and motivation to listen to a spoken narrative.

Temporal aspects of spoken messages, such as speech rate, fluency and pause distribution are known to be highly important for the understanding of a spoken message (Eklund, 2004), and are inherently linked to how verbal information units are segmented when described events are conceived and segmented in spoken narratives (Eklund, 2004; Huff et al., 2018).

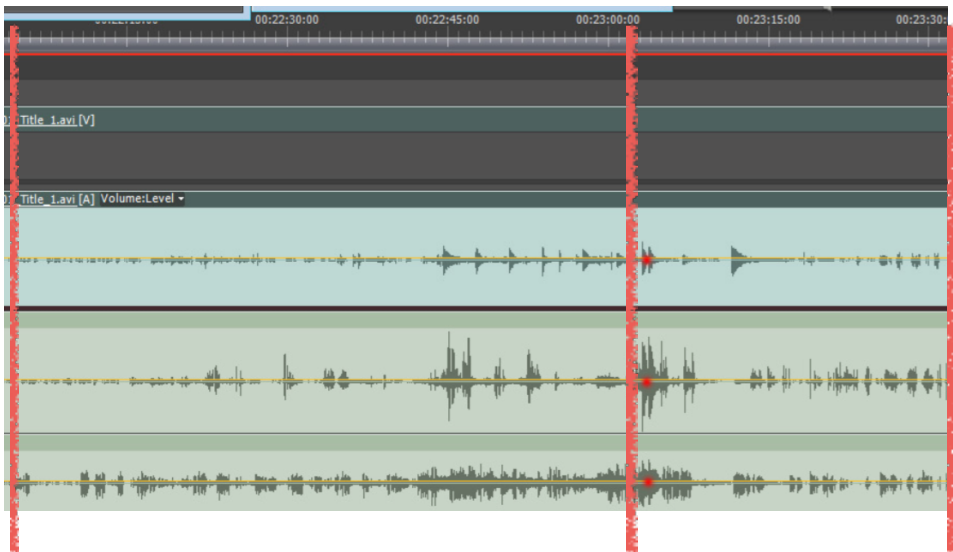
Not only speed but also intonation, style and vocal delivery may affect reception (Cabeza-Cáceres 2013, Walczak 2017).

Temporal features of speech, the role of voice quality, and how such factors interact with event segmentation and information processing, are crucial to consider when attempting to understand the principles that underlie successful communication through AD.

# Focus of the new project

## Cognitive aspects of AD

- (1) processing of visuospatial and temporal information
- (2) event segmentation
- (3) evaluating AD reception





# General aims

## Theoretical:

- to investigate how the blind audience understands critical aspects of verbal event descriptions and to **identify key principles** that underlie successful communication between the sighted and the blind.


## Applied:

- to use the project's outcomes to **support the training of audio describers** and AD practices, and ultimately to **facilitate the understanding and accessibility of visual information** for the visually impaired, audio describers and AD practices, and ultimately facilitate the understanding and accessibility of visual information for the visually impaired.

# Specific aims

- Systematically investigate how the blind understand and experience **visuospatial** and **temporal** properties of spoken event description, and to specify how **voice quality** interacts with those factors.
- Systematically investigate how the blind **conceive** and **segment** spoken event descriptions, and to specify how **speech rate** and **pausing** interact with those factors.
- Apply outcomes from point 1 and 2 to authentic AD and evaluate how this affects the **processing**, **quality** and **enjoyment** of the blind end users.

# Study 1

- Investigates how the blind experience and understand visuospatial and temporal information critical for successful AD.
- 40 blind participants and 40 sighted participants (as a control ip).
-  Verbal statements describing visuospatial relations or temporal change with various degrees of ambiguity.
- After each statement they will be asked to evaluate (1-6):
  - (1) how well they could imagine the content of the statement;
  - (2) their overall comprehension of the statement;
  - (3) how pleasant it was to listen to the statement.
- Voice quality: half of the statements will be presented with a simulated *dysphonic voice* and the other half will be presented with a *typical voice* (randomly distributed and counterbalanced over each condition).

# Study 2

- Investigates how the blind conceive and segment spatiotemporal event descriptions critical for successful AD.
- 40 blind participants and 40 sighted participants (as a control group).
- movie clips/verbal descriptions: 10-12 minutes long
- *Indicate event boundaries* in verbal descriptions of events from authentic movies and contrast this with the sighted control group, who will perform the same task but instead watch the corresponding movie clips.
- The verbal descriptions will comprise both *spatial and temporal event boundaries* that will be expressed in two discourse types: verbally *explicit* and verbally *implicit*.
- Manipulation of speech rate and pausing

# Spatial change



*In the house* sits a guy, around 20 years old, and is cleaning his shotgun. He lubricates it with oil, polishes it. He has a white tank top, mustache, dark wavy hair. His mother is standing in the living room, folding sheets.



This young guy is now *outside in the landscape*. The sun is shining, he has a bag on his back, a brown vest, white shirt. The empty landscape spreads out kilometer after kilometer, just a few shrubs here and there.

# Temporal change



We have left 1945 and now we are in 1993.

# Study 3

- The third study will evaluate authentic AD of a complete movie for 2 groups of AD receivers:
  - 1) with AD from a professional audio describer (N = 20);
  - 2) with AD from the same professional audio describer (N = 20) when he/she has been trained to use specific strategies based on findings from study 1 and 2.
- The reception of the two alternative versions of AD will be measured through a novel combination of quantitative and qualitative methods developed by members of the research group (Holsanova et al., 2015).
- Study 3 will implement results from study 1 and 2 in authentic AD of a complete movie and evaluate to what degree the present project actually can inform and improve AD practices.
- To develop the *strategies* that we will train the audio describer to use, we will first organize a *workshop* dedicated to this purpose. The workshop will consist of project members, AD educators, researchers, professional audio describers and end users of AD.
- Together we will discuss the outcomes from study 1 and 2 and use this knowledge to formulate appropriate *guidelines* that we can use for the training. As an added value, we intend to integrate those guidelines in future AD education.

# Audio Description Network

Jana Holsanova, Cognitive Science, LU  
Roger Johansson, Psychology, LU  
Peter Gärdenfors, Cognitive Science, LU  
Mats Andréén, linguistics, semiotics, LIU  
Johan Blomberg and Frida Blomberg, Linguistics, LU  
Viveka Lyberg-Åhlander, Logopedics, LU  
Nils Holmberg, Media & Communication, LU  
Erik Hedling, Ann-Kristin Wallengren, Anders Marklund, Film studies, LU  
Cecilia Wadensjö, Translation and Interpretation, SU  
Lotta Lagerman, AD training, Örebro/Fellingsbro  
Charles Forceville, Media studies, Amsterdam  
Petr Kaderka, Linguistics, Czech Academy of Sciences, Prag  
Swedish Braille Authority (Punktskriftsnämnden)  
Swedish Agency for Accessible Media (MTM)  
National Agency for Special Needs, Education and Schools (SPSM)  
Young Visually Impaired (US)  
Swedish Federation of Visually Impaired (SRF)



**LUND**  
UNIVERSITY

Holsanova, ARSAD 2019





## Selected publications

**Holsanova, J.**, Blomberg, J., Johansson, R., Blomberg, F. & Gärdenfors (in prep): Event structure in audio description of film.

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

**Holsanova, J.** & Wadensjö, C. (2016): Syntolkning - forskning, tankar och visioner. I: Holsanova, J. Andrén, M. & Wadensjö, C. (eds.): *Syntolkning – forskning och praktik*. Lund University Cognitive Studies 166/ Myndigheten för tillgängliga medier, rapport nr. 4, s. 13–15.

**Holsanova, J.**, Wadensjö, C. & Andrén, M. (eds.) (2016): *Syntolkning – forskning och praktik*. Lund University Cognitive Studies 166/ Myndigheten för tillgängliga medier, rapport nr. 4, s. 13–15.

**Holsanova, J.** (2015). Cognitive approach to audio description. In: Matamala, A. & Orero, P. (eds.): *Researching audio description: New approaches*. London: Palgrave Macmillan, pp. 49–73.

**Holsanova, J.**, Hildén, A., Salmson, M., & Kesen Tundell, V. (2015): *Syntolkning och uppläst text. En studie om hur användarna vill ha det med riktlinjer för audiovisuella medier*. Stockholm: Tundell & Salmson.

**Holsanova, J.** (2015): *Syntolkning – forskning och utbildning*. Myndigheten för yrkeshögskolan.

**Holsanova, J.** (2014b): In the eye of the beholder: Visual communication from a recipient perspective. In: David Machin (Ed.): *Visual Communication*. Handbooks of Communication Science [HoCS]. Chapter 14, pp. 331–355. De Gruyter.

**Holsanova, J.** (2014a): Reception of multimodality: Applying eye tracking methodology in multimodal research. In: Carey Jewitt (Ed.), *Routledge Handbook of Multimodal Analysis*. Second edition. Chapter 20, pp. 285–296. London: Routledge.

**Holsanova, J.** (ed) (2012): *Methodologies for multimodal research*, Visual communication 2012, Special issue, Vol. 11(3), Sage.

Boeriis, M. & **Holsanova, J.** (2012): Tracking visual segmentation. Connecting semiotic and cognitive perspectives. *Visual communication*, 11 (3), 259–281, Sage.

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

**Holsanova, J.** (2008): *Discourse, vision, and cognition*. John Benjamins Publishing Company: Amsterdam/Philadelphia.

**Holsanova, J.** (2001): *Picture Viewing and Picture Description. Two windows on the mind*. Doctoral dissertation. Lund University Cognitive Studies 83.

**Holsanova, J.** (1999): *På tal om bilder. Om fokusering av uppmärksamhet i och strukturering av talad beskrivande diskurs*. Licentiatavhandling. Lund University Cognitive Studies 78.