LUND UNIVERSITY

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

Holsanova, Jana; Johansson, Roger; Lyberg Åhlander, Viveka

2020

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Holsanova, J., Johansson, R., & Lyberg Åhlander, V. (2020). *How the blind audience receive and experience audio descriptions of visual events – a project presentation*. 39-41. Paper presented at Swiss Conference on Barrier-free Communication, Zurich, Switzerland.

Total number of authors:
3

Creative Commons License:
CC BY-ND

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Download date: 18. Oct. 2026

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

Jana Holsanova
Lund University / Cognitive science
Box 188, 22100 Lund, Sweden
jana.holsanova@lucs.lu.se

Roger Johansson
Lund University / Psychology
Box 188, 22100 Lund, Sweden
roger.johansson@psy.lu.se

Viveka Lyberg-Åhlander
Lund University/ Logopedics
Box 117, 22100 Lund, Sweden
viveka.lyberg_ahlander@med.lu.se

Abstract

We present a three-year multidisciplinary research project that started in 2019 and is funded by FORTE (Swedish Research Council for Health, Working Life and Welfare). The theoretical aim of the project is to gain a better understanding of the principles that underlie successful communication between the sighted and the blind via audio description (AD). The aim of a series of experiments is to identify similarities and differences in how the sighted and the blind understand, segment and experience visual, spatial as well as temporal properties of an event. The applied goal is to increase the quality of AD and to support the training of audio describers and AD practices, and ultimately facilitate the understanding and accessibility of visual information for the visually impaired.

1 Introduction

A lot of information in everyday life is visually processed through sight. For instance, to conceive visual events when watching a movie is something that is often taken for granted. However, for audiences with severe visual impairment this is not possible. But through supporting audio description, information accessibility can substantially improve. Audio description (AD) offers a richer and more detailed understanding and enjoyment for blind and visually impaired audiences. Spoken descriptions of visual scenes evoke inner mental images and enhance the process of meaning-making for target audiences (Holsanova 2015, 2016a).

There is currently a great need for establishing a scientific foundation for AD (Kruger & Orero 2010) and a surge for cognitively oriented research on audio description from a recipient perspective (Holsanova 2014, 2016a). How do end users perceive and understand verbal descriptions of visual events (Holsanova 2016a,b; Johansson 2016)? What preferences do they have (Holsanova et al. 2015)? How do blind and sighted audiences imagine, segment and remember properties of the unfolding events (Johansson, 2016)? How does voice quality, speech rate and pausing of the audio describer affect understanding and enjoyment (Fryer 2010; Lyberg-Åhlander, 2015)?

Thus, the focus of the current multidisciplinary project is on processing of visuospatial and temporal information, on segmentation of events in a film narrative, on evaluation of AD reception and on the role of voice quality.

2 Aims of the project

The theoretical aim of the project is 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 via audio description (AD). The applied goal is 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.

In particular, we aim to: a) Systematically investigate how the blind understand and experi-

© 2019 The authors. This article is licensed under a Creative Commons 3.0 license, no derivative works, attribution, CC-BY-ND.

ence visuospatial and temporal properties of spoken event description, and to specify how voice quality interacts with those factors. b) Systematically investigate how the blind conceive and segment spoken event descriptions, and to specify how speech rate and pausing interact with those factors. c) Apply outcomes from steps a) and b) to authentic AD and evaluate how this affects the processing, quality and enjoyment of the blind end users. The project also includes workshops with researchers, practitioners of AD, and blind users of AD. The studies are embedded in the theoretical framework of event cognition, mental imagery and empirical studies on aural properties of spoken descriptions.

3 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 (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 (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).

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 (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.

4 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.

5 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 (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, and are inherently linked to how verbal information units are segmented when described events are conceived and segmented in spoken narratives (Eklund 2004). Not only speed but also intonation, style and vocal delivery may affect reception (Cabeza-Cáceres 2013; Walczak & Fryer 2017; Fernández-Torné & Matamala 2015).

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.

6 Relevance and impact

The ability to access and understand visual media is of great importance in today's multimedia society. Audio description offers a richer and more detailed understanding and enjoyment for the visually impaired audiences and is crucial for the inclusion of the visually impaired in the society.

This implies higher demands on AD quality. To achieve this, basic research on how a visually impaired audience process event descriptions in relation to a sighted population is fundamental, as well as basic research on what, how and when such information should be communicated in AD. However, research about how to optimize this communication is scarce in the current literature. The present project will provide invaluable knowledge with direct implications for the education and training of audio describers, and it is of great strategic importance for accessible information in the society, at the working place and in everyday life.

References

Boltz, M. 1992. Temporal accent structure and the remembering of filmed narratives. *Journal of Experimental Psychology: Human Perception and Performance*, 18(1), 90.

Cabeza-Cáceres, C. 2013. *Audiodescripció i recepción. Efecte de la velocitat de narració, l'entonació i l'explicitació en la comprensió fílmica*. PhD. http://www.tdx.cat/handle/10803/113556

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

Eklund, R. 2004. *Disfluency in Swedish human–human and human–machine travel booking dialogues* (Doctoral dissertation, Linköping University Electronic Press).

Fernández-Torné, Anna & Matamala, Anna, 2015. Text-to-speech vs. human voiced audio descriptions: a reception study in films dubbed into Catalan. *The Journal of Specialised Translation* Issue 24 – July 2015, 61–88.

Fryer, L. 2010. Audio description as audio drama - a practitioner's point of view. *Perspectives: Studies in Translatology* Vol. 18, No. 3, September 2010, 205-213.

Holsanova, J. 2016a. Cognitive approach to audio description. In: Matamala, A. & Orero, P. (eds.): *Researching audio description: New approaches*, 49–73. London: Palgrave Macmillan.

Holsanova, J. 2016b. Kognitiva och kommunikativa aspekter av syntolkning. In: Holsanova, J. Andrén, M. & Wadensjö, C. (eds.): *Syntolkning – forskning och praktik*. (Audio description – research and practices). Lund University Cognitive Studies 166, 17–27.

Holsanova, J., Hildén, A., Salmson, M., & Kesen Tundell, V. 2015. *Audio description and audio subtitles.* A study of user preferences with guidelines for audiovisual media. Stockholm: Tundell & Salmson.

Holsanova, J. 2014. In the mind of the beholder: Visual communication from a recipient perspective. In Machin, D. (Ed.) *Visual communication*, Mouton - De Gruyter, 331–354.

Johansson, R. 2016. Mentala bilder hos seende och blinda. In J. Holsanova, C. Wadensjö, & M. Andrén (Eds.). *Syntolkning - forskning och praktik*. Lund University Cognitive studies 166, 29-38.

Johansson, R., Holsanova, J. & Holmqvist, K. 2006. Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness. *Cognitive Science*, 30(6), 1053-1079.

Kruger, J.-L. & Orero. P. 2010. Audio Description, Audio Narration – A New Era in AVT. Special Issue of *Perspectives: Studies in Translatology,* Vol 18: 3.

Lyberg-Åhlander, V., Brännström, K. J., & Sahlen, B. 2015. On the interaction of speakers' voice quality, ambient noise and task complexity with children's listening comprehension and cognition. *Frontiers in Psychology*, 6, 871.

Noordzij, M.L., Zuidhoek, S., Postma, A. 2007. The influence of visual experience on visual and spatial imagery. *Perception*, 36, 101–112.

Postma, A., Zuidhoek, S., Kappers, A. M., & Noordzij, M. L. 2006. Haptic spatial orientation processing and working memory. *Cognitive Processing*, 7(1), 181.

Radvansky, G. A., & Zacks, J. M. 2014. *Event cognition*. Oxford University Press.

Rogerson, J., & Dodd, B. 2005. Is there an effect of dysphonic teachers' voices on children's processing of spoken language?. *Journal of Voice*, 19(1), 47-60.

Vercauteren, G., & Remael, A. 2014. Spatio-temporal Settings. In: *Audio Description: New Perspectives Illustrated*.Amsterdam/Philadelphia, John Benjamins, 61-80.

Zacks, J. M., Speer, N. K., & Reynolds, J. R. 2009. Segmentation in reading and film comprehension. *Journal of Experimental Psychology*: General, 138(2), 307.

Walczak, A. & Fryer, L. 2017. Vocal delivery of audio description by genre: measuring users' presence. *Perspectives Studies in Translatology* 26(1):1-15.