Mind the Boundaries: Neurocognitive and AI-Driven Insights into Event Perception in Audio-Described Films

Roger Johansson, Maja Rudling, Johan Mårtensson, Jana Holsanova

ARSAD 2025

https://syntolkning.ht.lu.se/

How the blind audience receive and experience audio descriptions of visual events

What?

- Identify the **perceptual** and **neurocognitive** factors that contribute to effective communication between sighted and blind individuals during **audio descriptions (AD)** of visual events.

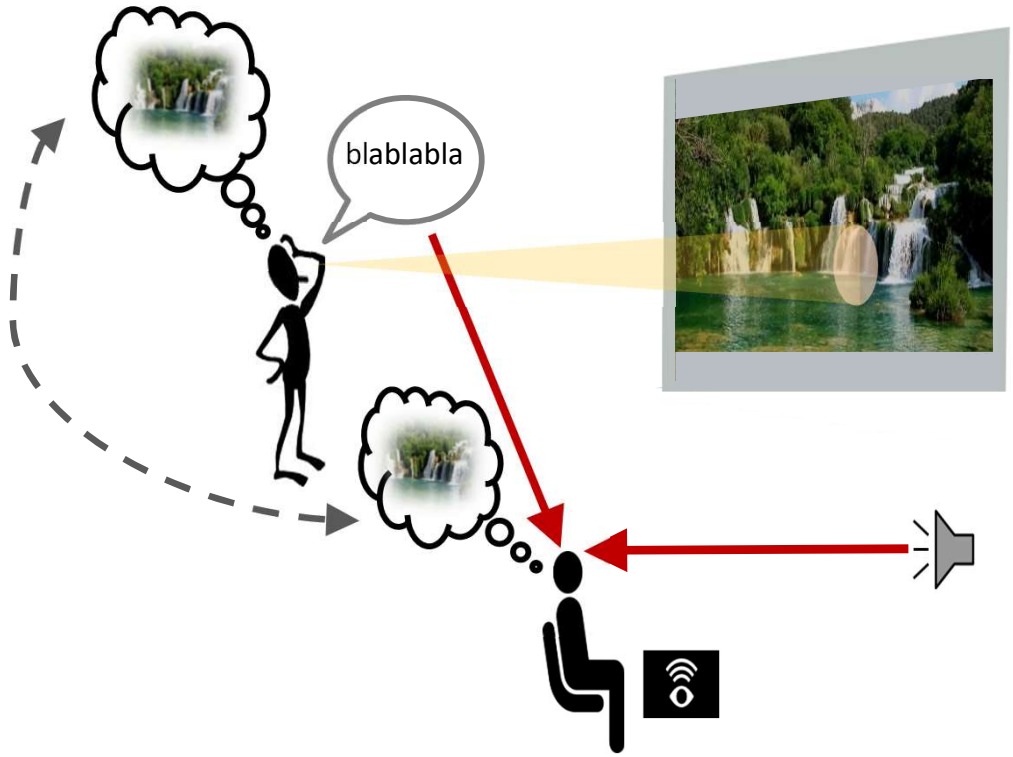
How?

- Conduct experimental studies using techniques and methods from **cognitive neuroscience** and **experimental psychology**.

Why?

- Enhance **understanding** of how fundamental **perceptual** and **neurocognitive** processes influence communication between sighted and blind individuals.

- Apply this knowledge to increase the **quality of AD** and **AD practices**, and ultimately facilitate the understanding and **accessibility of visual information** for the visually impaired.



Event boundaries & the human mind

- Although our experiences unfold as a continuous stream, we do not perceive them that way. Instead, we organize them into **distinct events**, with meaningful changes in **space** and **time**, serving as natural **event boundaries** (e.g., Radvansky & Zacks, 2014; Zacks et al., 2001).
- People tend to segment experiences in similar ways, suggesting that **event segmentation** is guided by **shared cognitive mechanisms** (e.g., Baldassano et al., 2018).
- Information encountered at event boundaries is **better remembered** and **improves comprehension** (e.g., Huff et al., 2014).

Event boundaries & audio description

Event boundary

”Hon ser rakt på sin far”

”She looks straight at her father”

”Uppgivet lutar Gerlof sig tillbaka”

”Resigned, Gerlof leans back”

”**Nästa dag.** Julia ställer en papplåda i den röda bilens öppna bagageutrymme ”

”The next day. Julia places a cardboard box into the open trunk of the red car”

Holsanova, J., Blomberg, J., Blomberg, F., Gärdenfors, P., & Johansson, R. (2023, *Journal of Audiovisual Translation*)

Study 1 - Event boundary perception in audio described films

- **Scientific aim:** Systematically examine how sighted and non-sighted individuals experience and understand film narratives as a chain of events with and without audio description.

- **Specific focus:** Investigate how different AD approaches influence the perception of critical event boundaries.

- **Applied goal**: Understanding how AD influences event segmentation is crucial for improving accessibility and ensuring a more equivalent narrative experience for visually impaired audiences.

Study 1 - Event boundary perception in audio described films

- 44 participants (28 female, 2 non-binary).
- Mean age of 37.5 years (SD = 12.8)
- 27 sighted and 17 non-sighted individuals (congenitally blind or lost sight early in life)
- The first 43 minutes of the **Swedish film *Skumtimmen*** (*Echoes from the dead*, Alfredsson, 2013)
- Pre-coded into **92 spatiotemporal event boundaries** in a pervious study (Holsanova et al., 2023)
- Sighted individuals: exposed to the original film
- Non-sighted individuals: exposed to the film with AD

Two AD versions (by a female professional audio describer):

1. Explicit version **2. Implicit version**

Indicating & registering event boundaries

"NOW!"

"Hon ser rakt på sin far"

"She looks straight at her father"

"Uppgivet lutar Gerlof sig tillbaka"

"Resigned, Gerlof leans back"

"**Nästa dag.** Julia ställer en papplåda i den röda bilens öppna bagageutrymme "

"**The next day.** Julia places a cardboard box into the open trunk of the red car"

Explicit event boundaries in AD

"NOW!"

"Hon ser rakt på sin far"

"She looks straight at her father"

"Uppgivet lutar Gerlof sig tillbaka"

"Resigned, Gerlof leans back"

"**Nästa dag.** Julia ställer en papplåda i den röda bilens öppna bagageutrymme "

"**The next day.** Julia places a cardboard box into the open trunk of the red car"

Implicit event boundaries in AD

"NOW!"

"Hon ser rakt på sin far"

"She looks straight at her father"

"Uppgivet lutar Gerlof sig tillbaka"

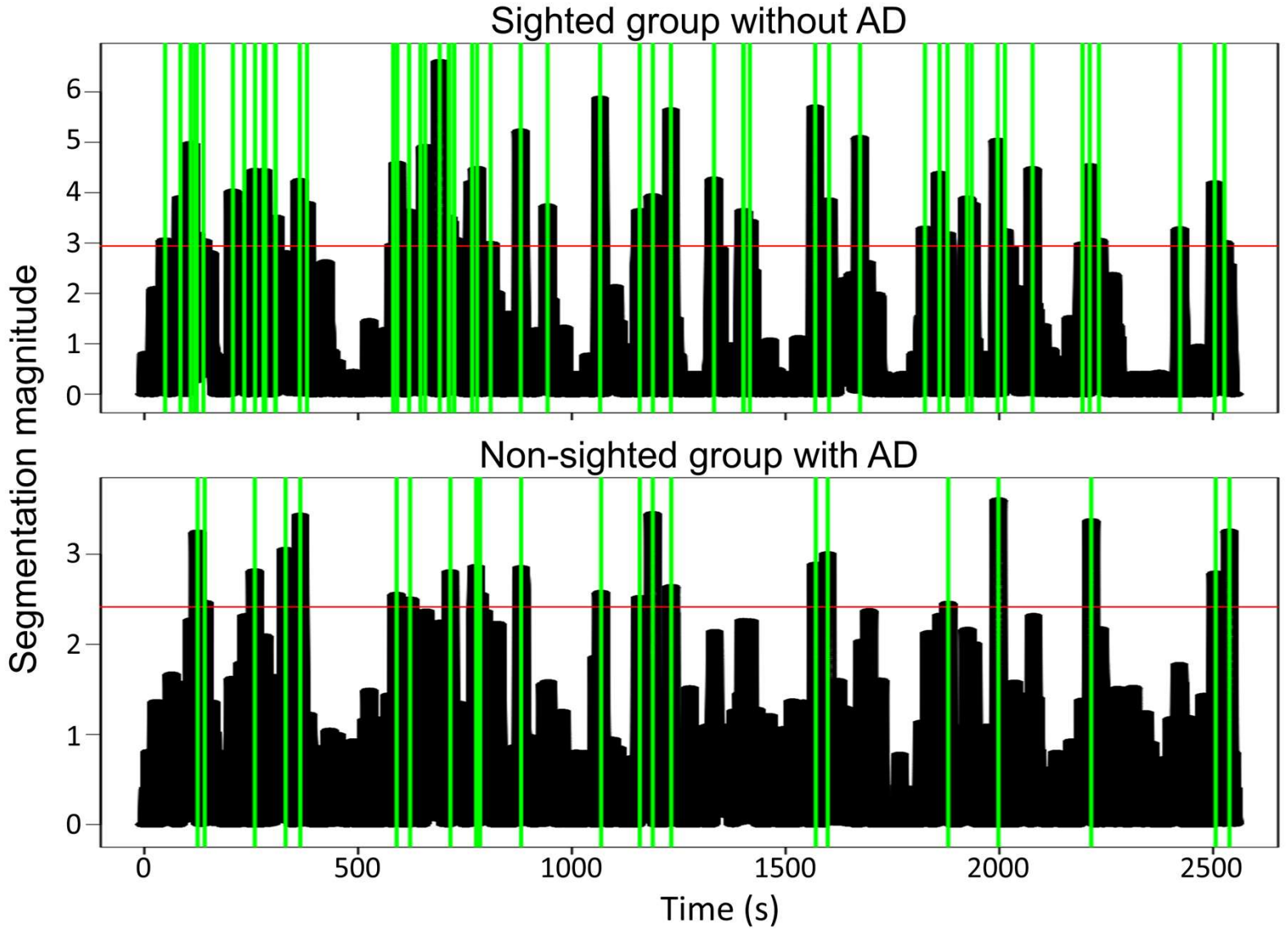
"Resigned, Gerlof leans back"

"Julia ställer en papplåda i den röda bilens öppna bagageutrymme "

"Julia places a cardboard box into the open trunk of the red car"

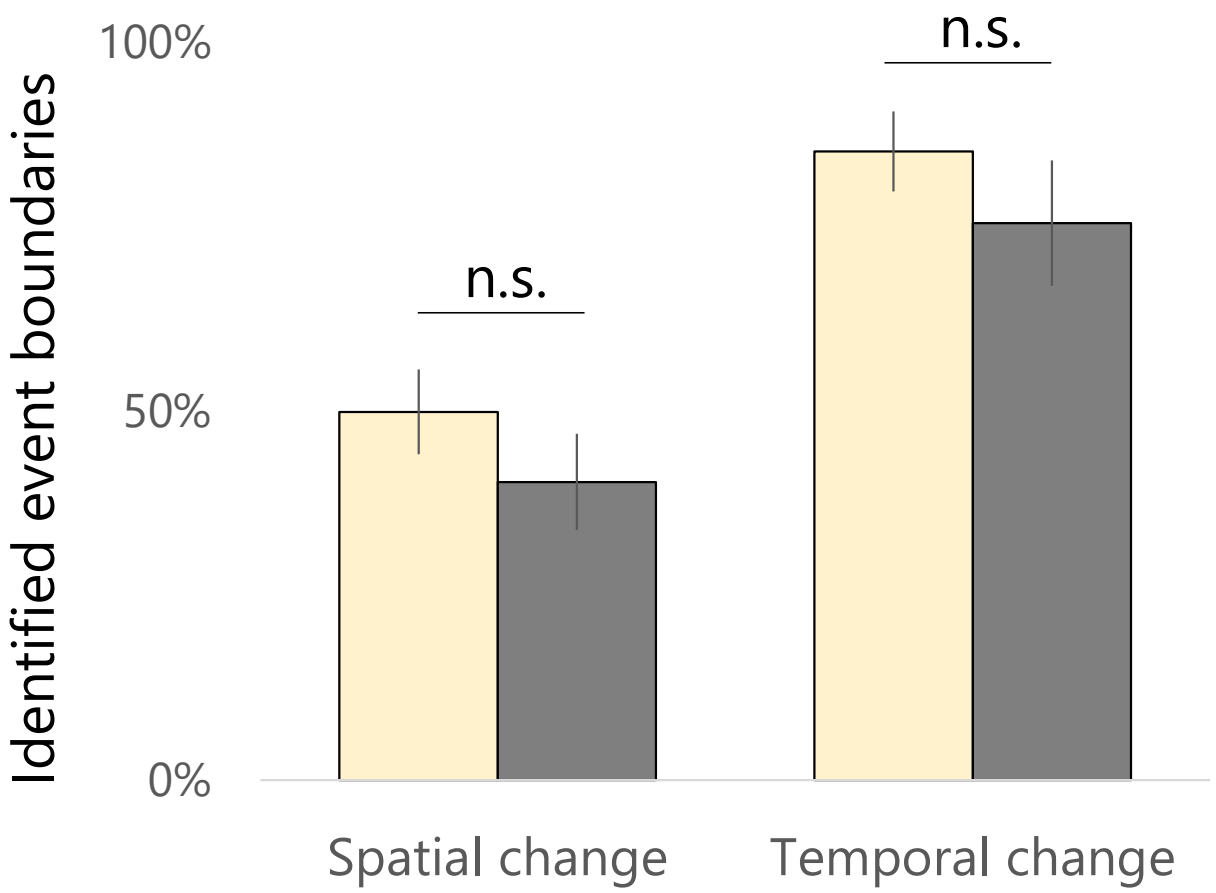
Results: Identified Event Boundaries

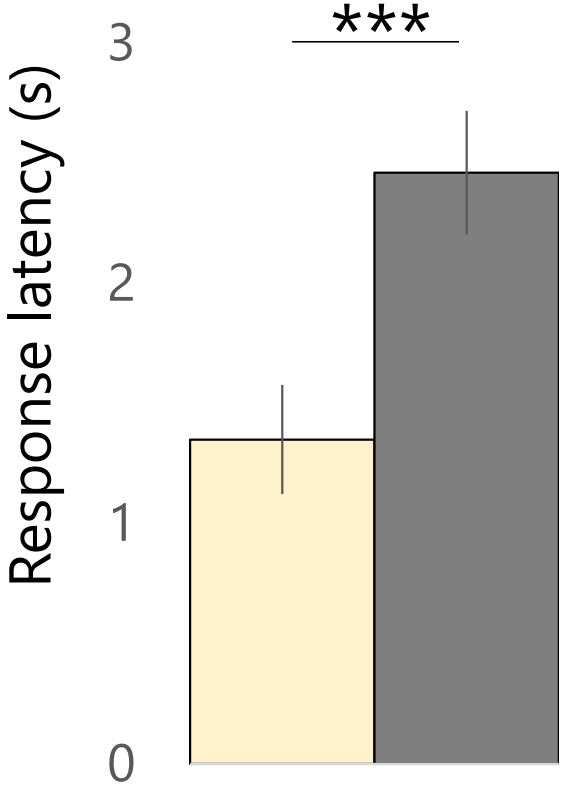
Johansson, R., Rastegar, T., Lyberg-Åhlander, V., & Holsanova, J. (2024, *Applied Cognitive Psychology*)



Results: Identified Event Boundaries

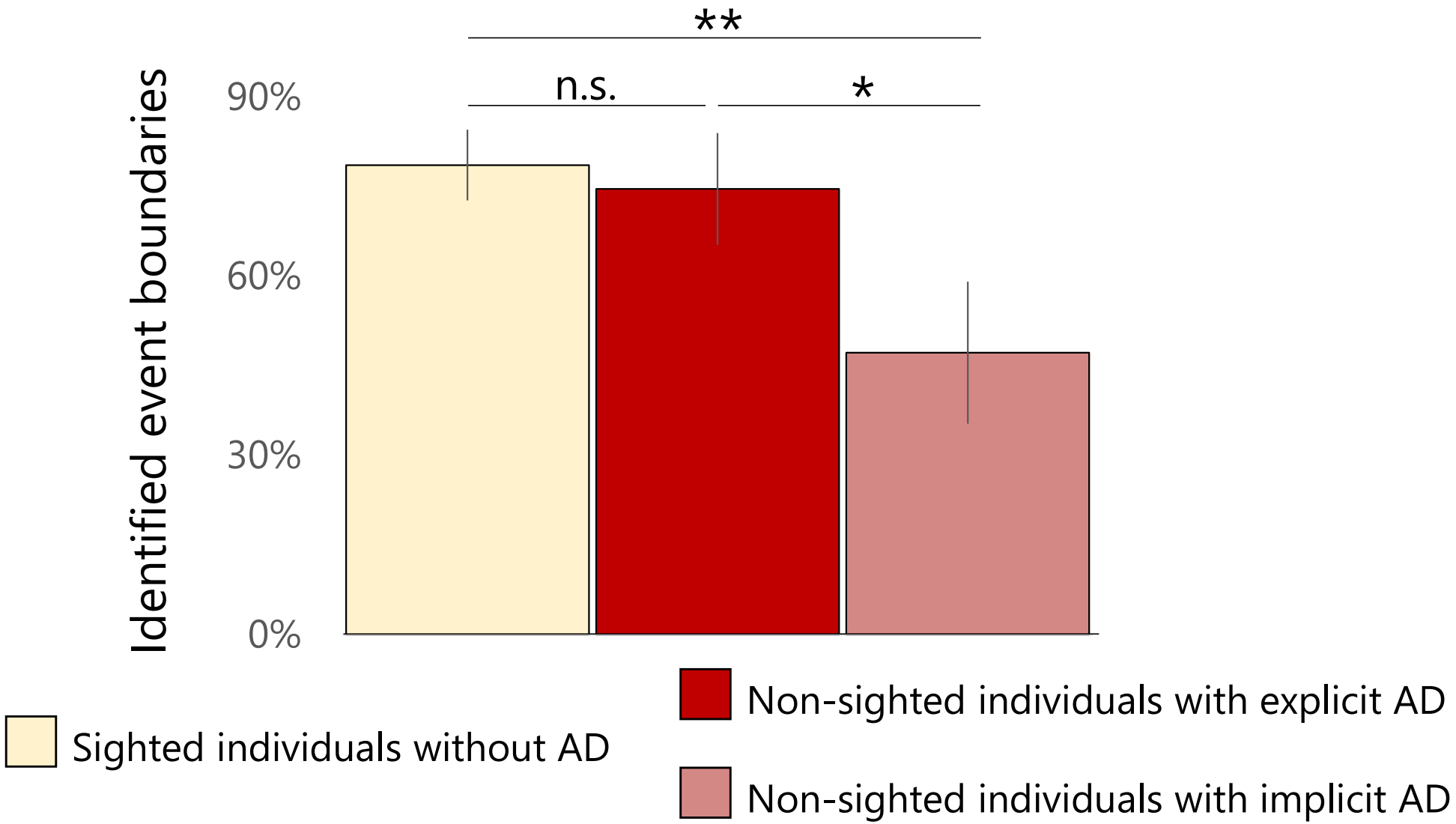
Johansson, R., Rastegar, T., Lyberg-Åhlander, V., & Holsanova, J. (2024, *Applied Cognitive Psychology*)





Results: Explicit versus Implicit Event Boundaries

Johansson, R., Rastegar, T., Lyberg-Åhlander, V., & Holsanova, J. (2024, *Applied Cognitive Psychology*)

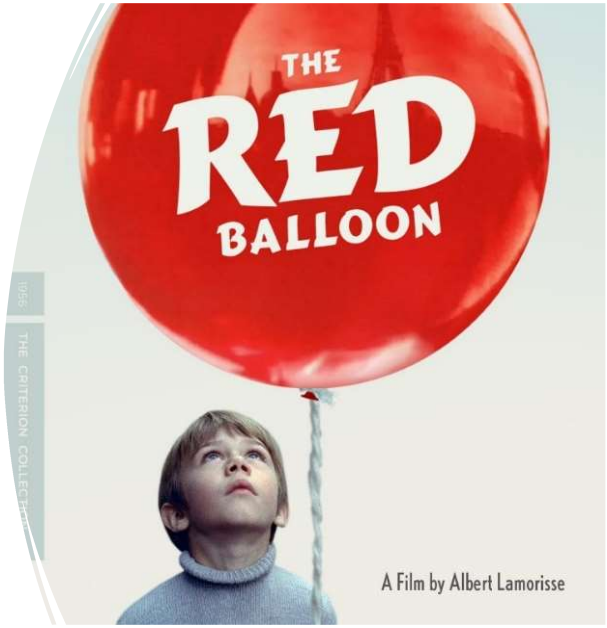


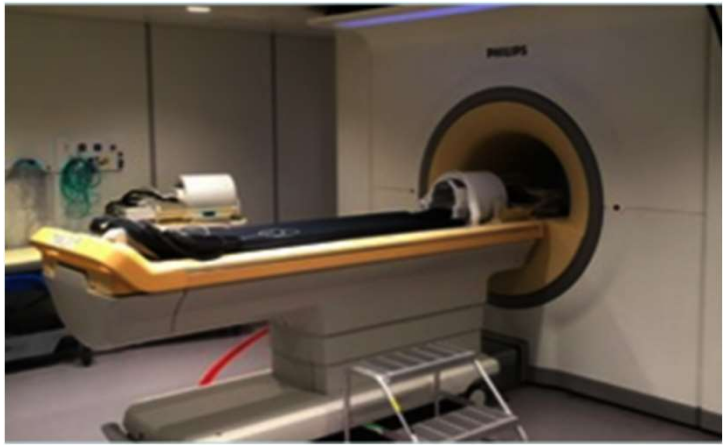
Summary & Conclusions

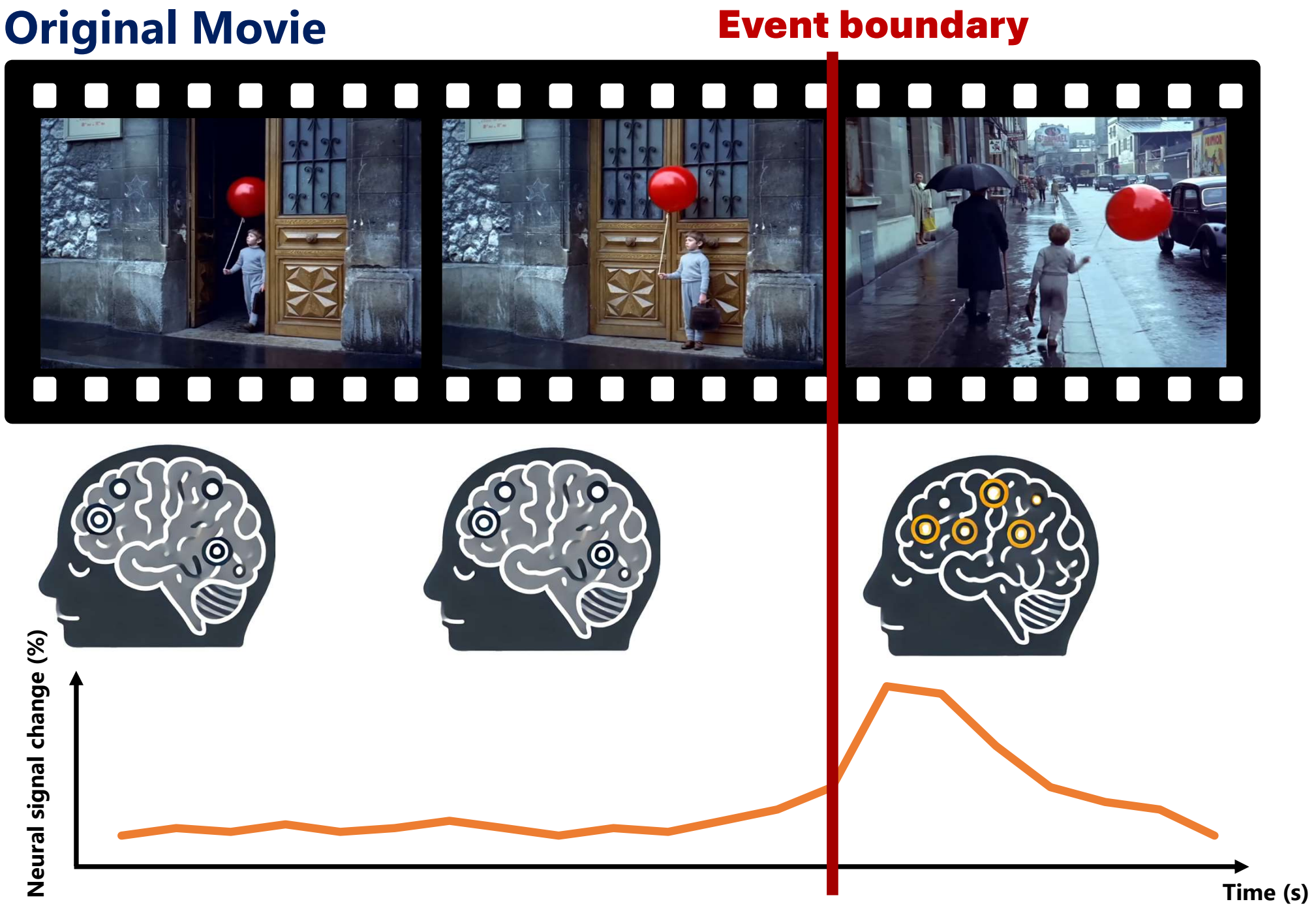
- Non-sighted participants perceived event boundaries similarly to sighted participants, indicating that **AD effectively conveys event structure**.
- **Implicit AD**, where event boundaries were not explicitly conveyed, **reduced the perception** of these transitions, likely weakening appropriate event segmentation of the unfolding narrative.
- **Explicit AD enhanced the perception** of event boundaries, supporting a clearer understanding of the narrative event structure and aligning it more closely with how sighted individuals process the film.
- These findings highlight the need for clear, explicit AD to improve comprehension, cinematic experience, and achieve **narrative equivalence**, with implications for **AD practices** and audio describer training.

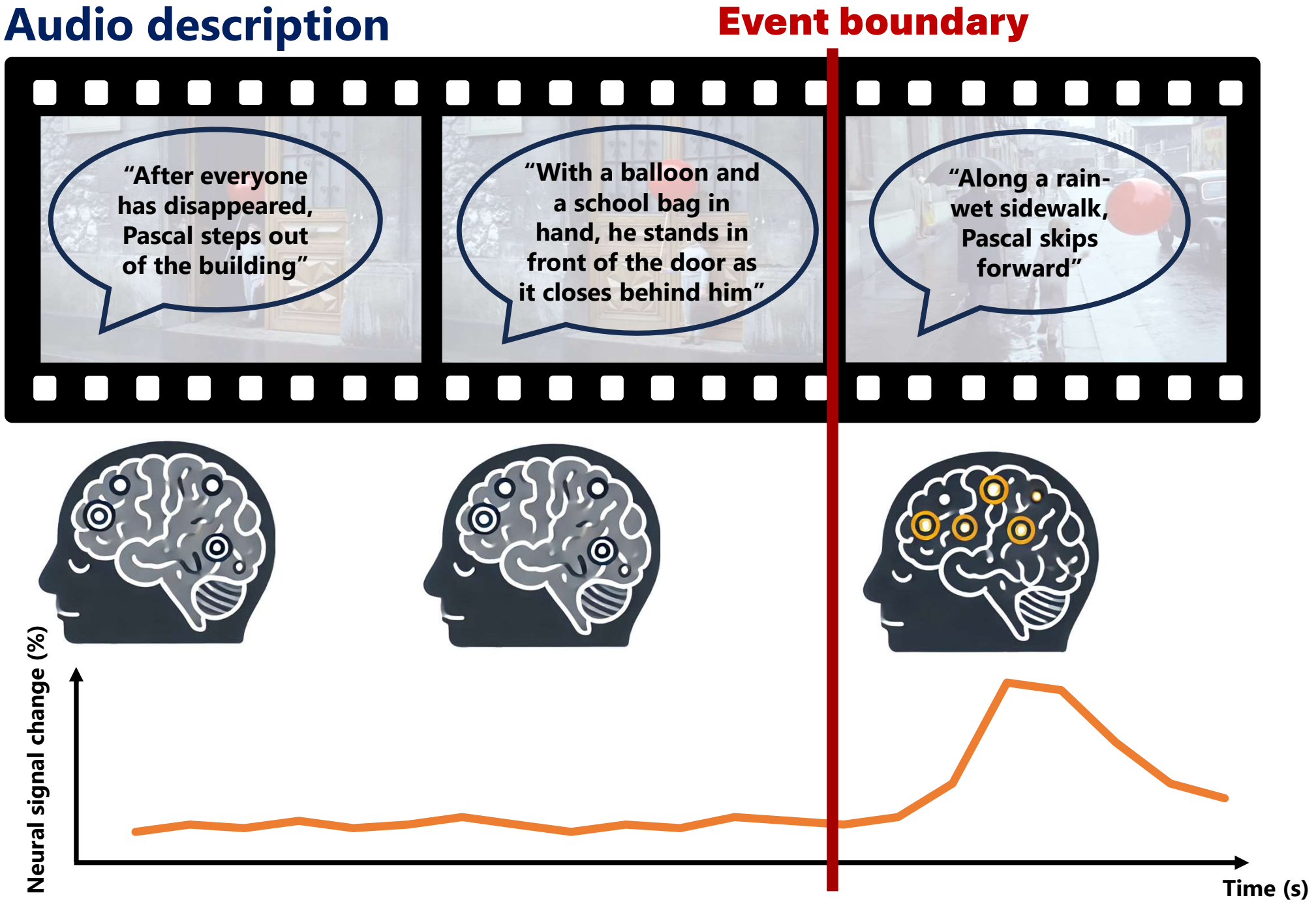
Study 2 – neurocognitive insights of event segmentation in audio described films

- Watching *"the red balloon"* while in an MRI scanner (Zacks et al., 2010)
- **Sighted** participants exposed to **the original film.**
- **Sighted** participants exposed to the **AD without visual input.**
- **Non-sighted** participants exposed to the film **with AD.**
- No task to indicate event boundaries.
- **Spontaneous event segmentation** captured by **neural signal changes** based on methods from Zacks et al. (2010).






Original Movie
Event boundary
Neural signal change (%)
Time (s)


Audio description
Event boundary
"After everyone has disappeared, Pascal steps out of the building"
"With a balloon and a school bag in hand, he stands in front of the door as it closes behind him"
"Along a rain-wet sidewalk, Pascal skips forward"
Neural signal change (%)
Time (s)

Automatic AI-generated video descriptions and event segmentation

- Large Language models, like GPT-3 and GPT-4 can segment narrative texts into events comparable to human coders (Michelmann, et al., 2023).
- However, for visual and audiovisual media, the case is very different…

<u>Possible to detect:</u>

- Scene cuts through distinct changes in visual information.
- Motion-based event boundaries through the “optical flow”.
- However, many false positives….

Automatic AI-generated video descriptions and event segmentation

- CLIP (Contrastive Language-Image Pretraining) model by open AI.
- Sample keyframes and turn into text to extract semantic meaning.
- I did not manage to get good performance though (multiple attempts and extensive effort).
- A lot of false positives….
- Might work with extensive tweaking and training!

Thank you for your attention!