

Spatiotemporal description of events in AD - the role of cognitive aspects and voice quality

Roger Johansson, Jana Holsanova, Viveka Lyberg-Åhlander, Erika Sombeck, Tina Rastegar

ARSAD 2023



LUND
UNIVERSITY



Åbo Akademi
University



Forskningsrådet för
hälsa, arbetsliv och välfärd

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

Ongoing project: Roger Johansson, Jana Holsanova, Viveka Lyberg-Åhlander

What?

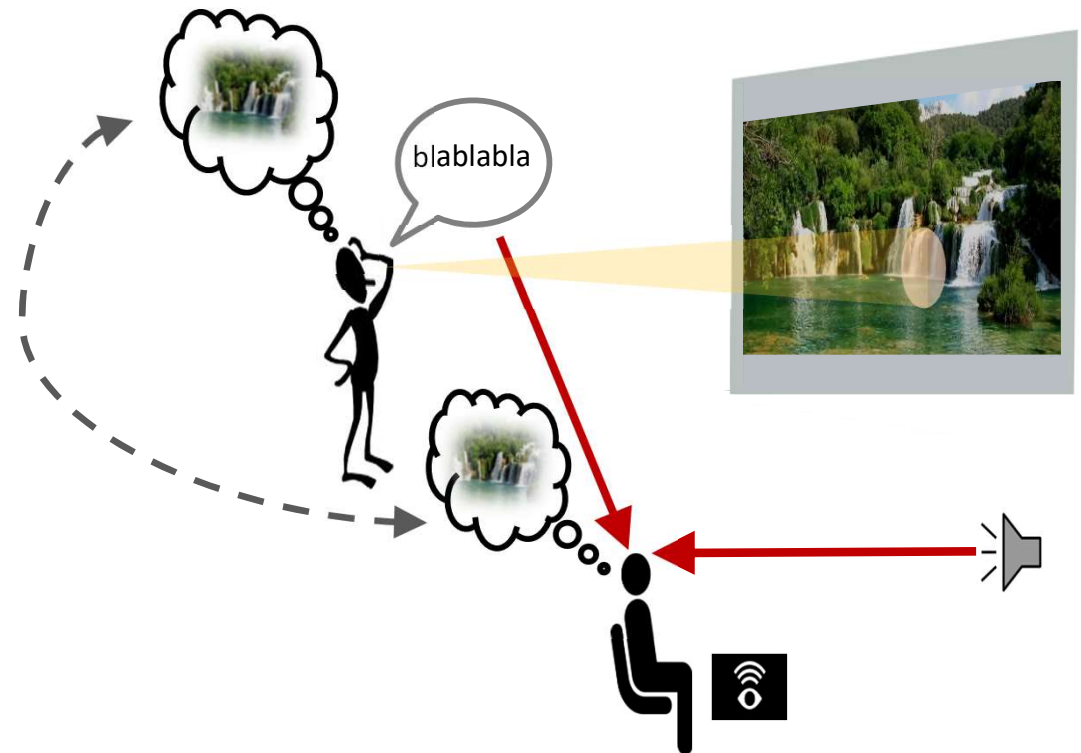
- Identify **perceptual** and **cognitive** factors underlying successful communication between the sighted and the blind during **audio descriptions (AD)** of visual events.

How?

- Experimental studies using methods from **cognitive science** and **experimental psychology**.

Why?

- Increase **knowledge** of how these factors affect communication between the sighted and the blind.
- 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.



The present study

- **Aim:** Systematically investigate how non-sighted people experience and understand **spatial relations** and **temporal change** of verbally described events - and to investigate what significance the describer's **voice quality** has in this interaction.
- Specific focus on how sighted vs non-sighted people **imagine** and **create mental models** of spatiotemporal content from verbal event descriptions. Similarities?
Differences?
- Critical for the **experience** and **understanding** of how described state-of-affairs relate to each other over **time** and **space**.

The Experiment

- 40 participants
- 20 sighted and 20 non-sighted (congenitally blind or lost sight early in life)
- Groups matched for verbal working memory (Competing Language Processing Task – CLPT)
- Conducted over Zoom
- Each participant listened to 50 event descriptions
- 20 Event descriptions of spatial relations
- 30 Event descriptions of motion changes
- High and Low specificity

Description Specificity

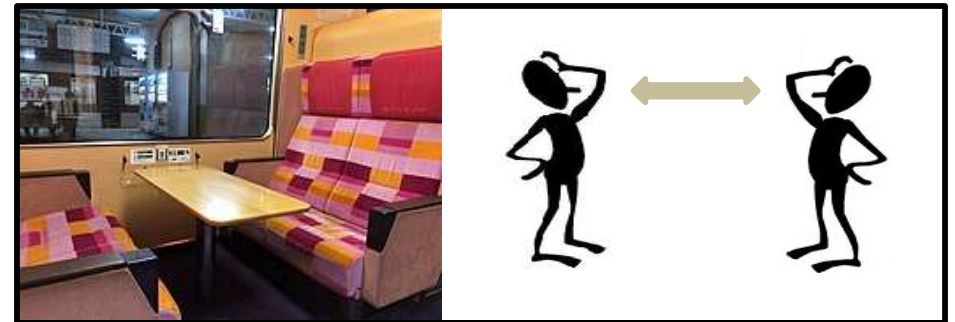
Event descriptions of spatial relations

Low Specificity

On the train. Lisa is in a train compartment. Lisa's sister Maja is also there. Lisa sits **in front of** Maja.

High Specificity

On the train. Lisa is in a train compartment. Lisa's sister Maja is also there. Lisa sits **opposite** Maja.



Description Specificity

Event descriptions of motion changes

Low Specificity

In school. It's Monday morning. Frank **enters** the classroom door.



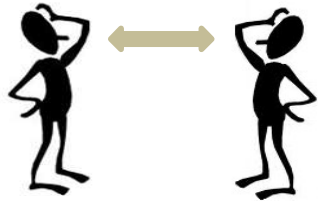
High Specificity

In school. It's Monday morning. Frank **rushes through** the classroom door.

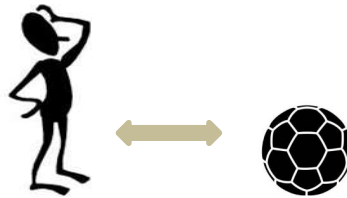


Types of Event Descriptions

Event descriptions of spatial relations



1. Person-Person

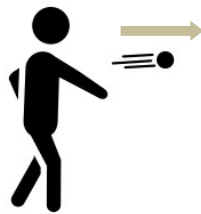


2. Person-Object

Event descriptions of motion changes



1. Person



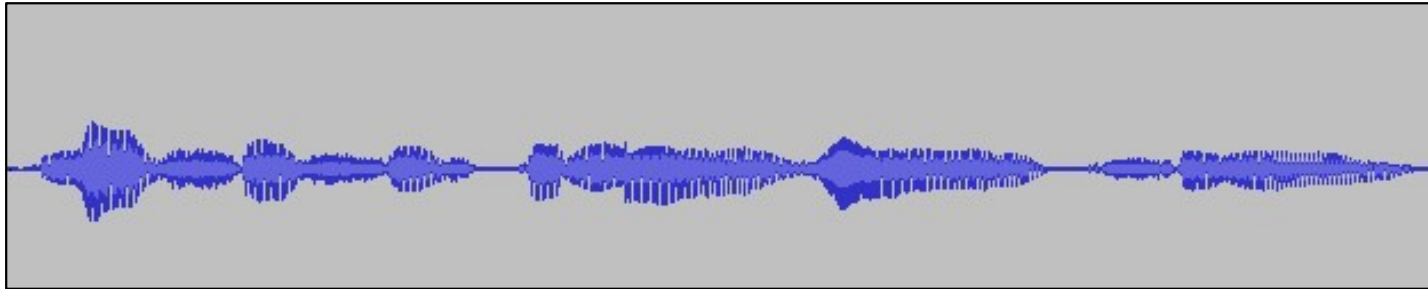
2. Person-Object



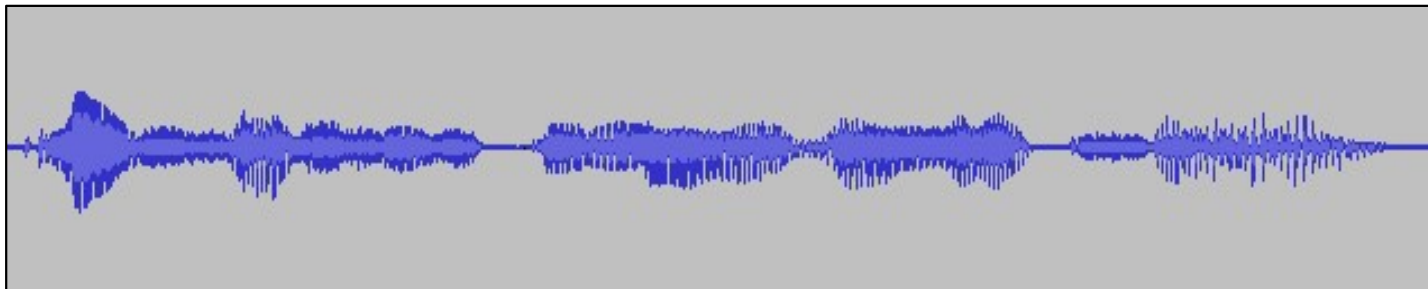
3. Person-Person

Voice Quality

Normal Voice



Dysphonic (hoarse) Voice



The Experiment

- Each participant listened to 50 event descriptions
- 20 Event descriptions of spatial relations
 - 10 Person-person, 10 Person-Object
- 30 Event descriptions of motion changes
 - 10 Person, 10 Person-Object, 10 Person-Person
- Equally distributed across high and low specificity
- Equally distributed across normal and dysphonic voice
- Data analysed with Generalised Mixed Effects Models

The Experiment

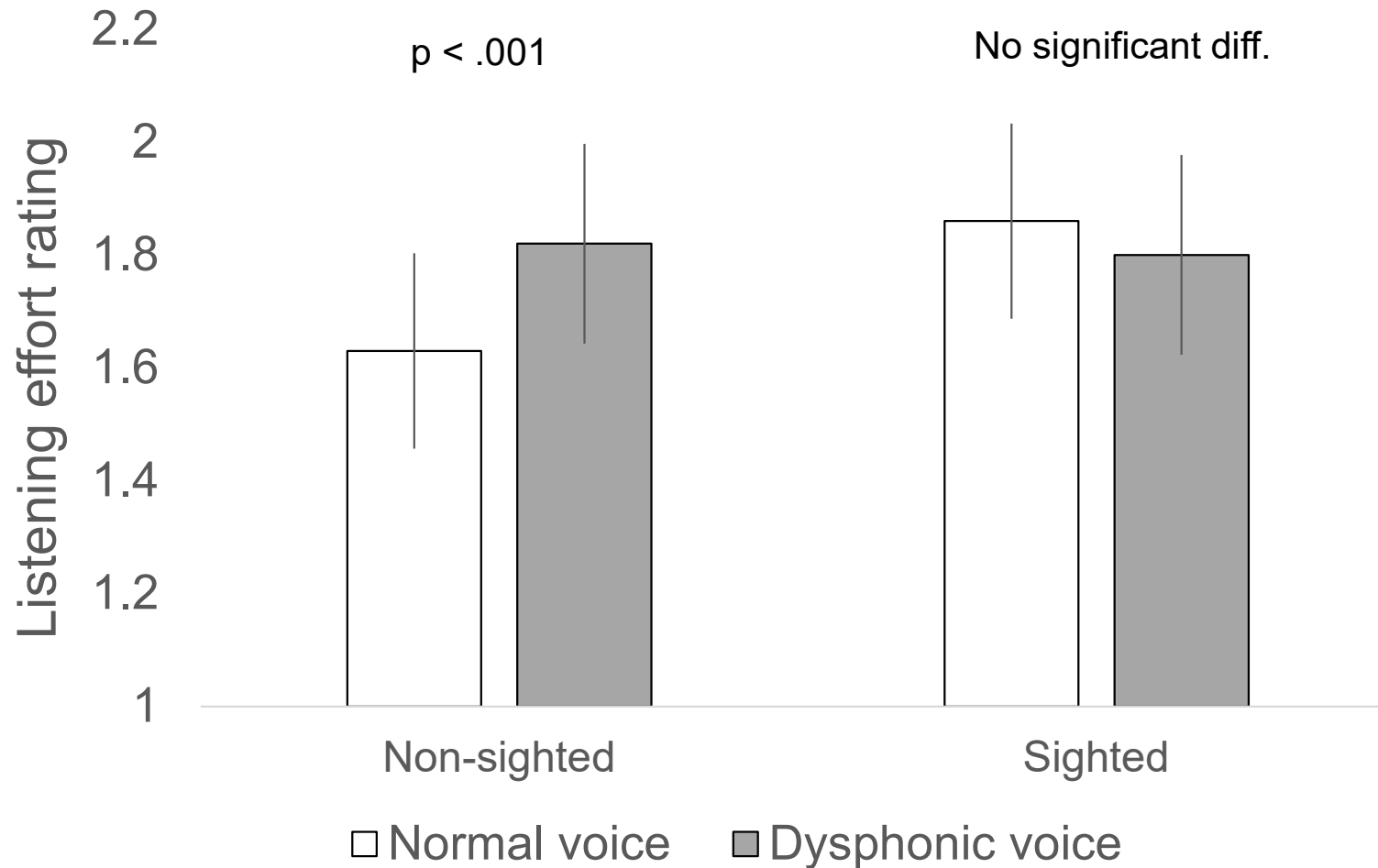
Task after listening to each event description

On a scale 1 to 6 rate:

1. How well you could **imagine** the content of the described scenario
2. How well you **understood** the described scenario
3. How **effortful** it was for you to listen to the described scenario
4. How **enjoyable** it was for you to listen to the described scenario

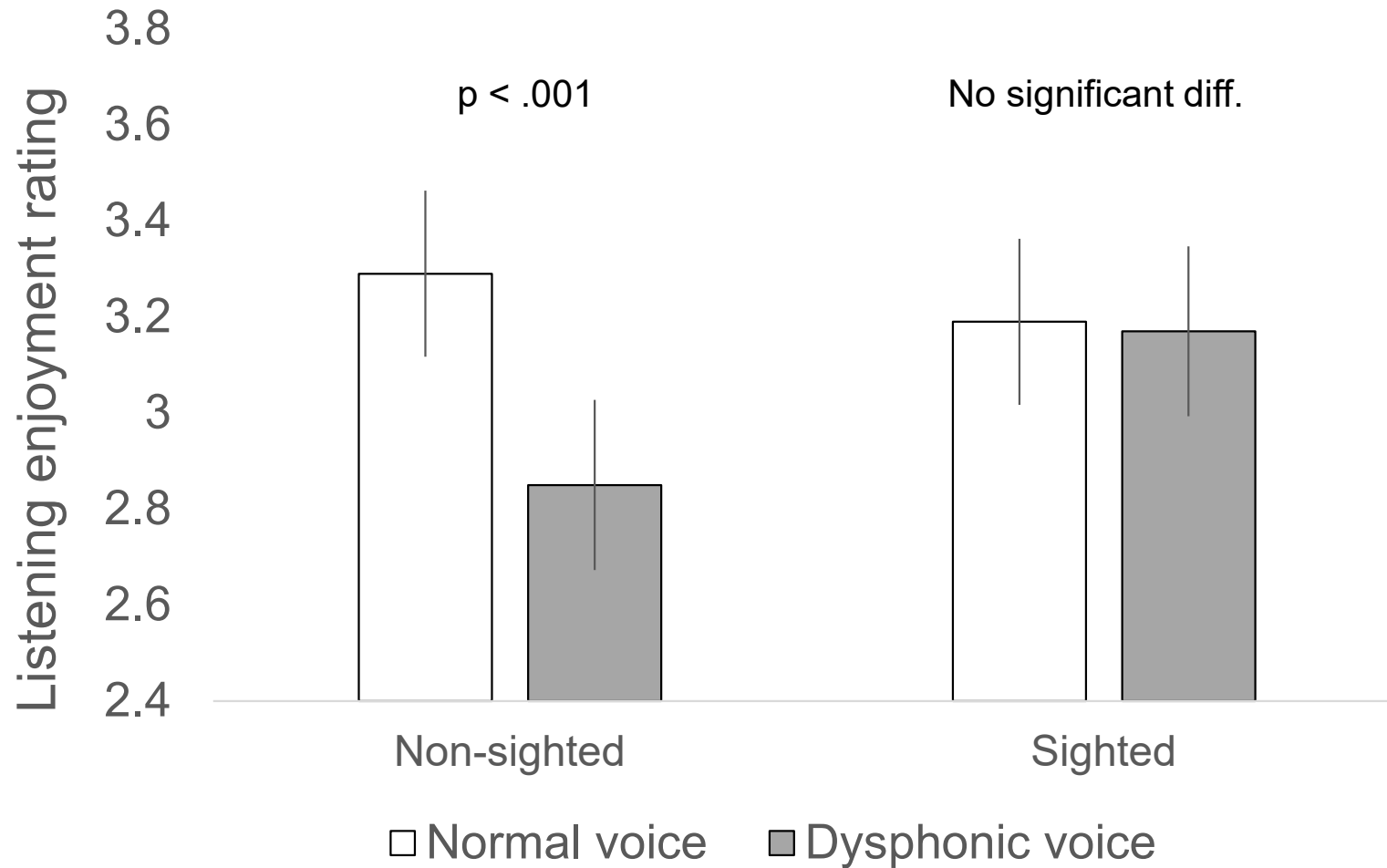
Results – Voice quality

Listening effort



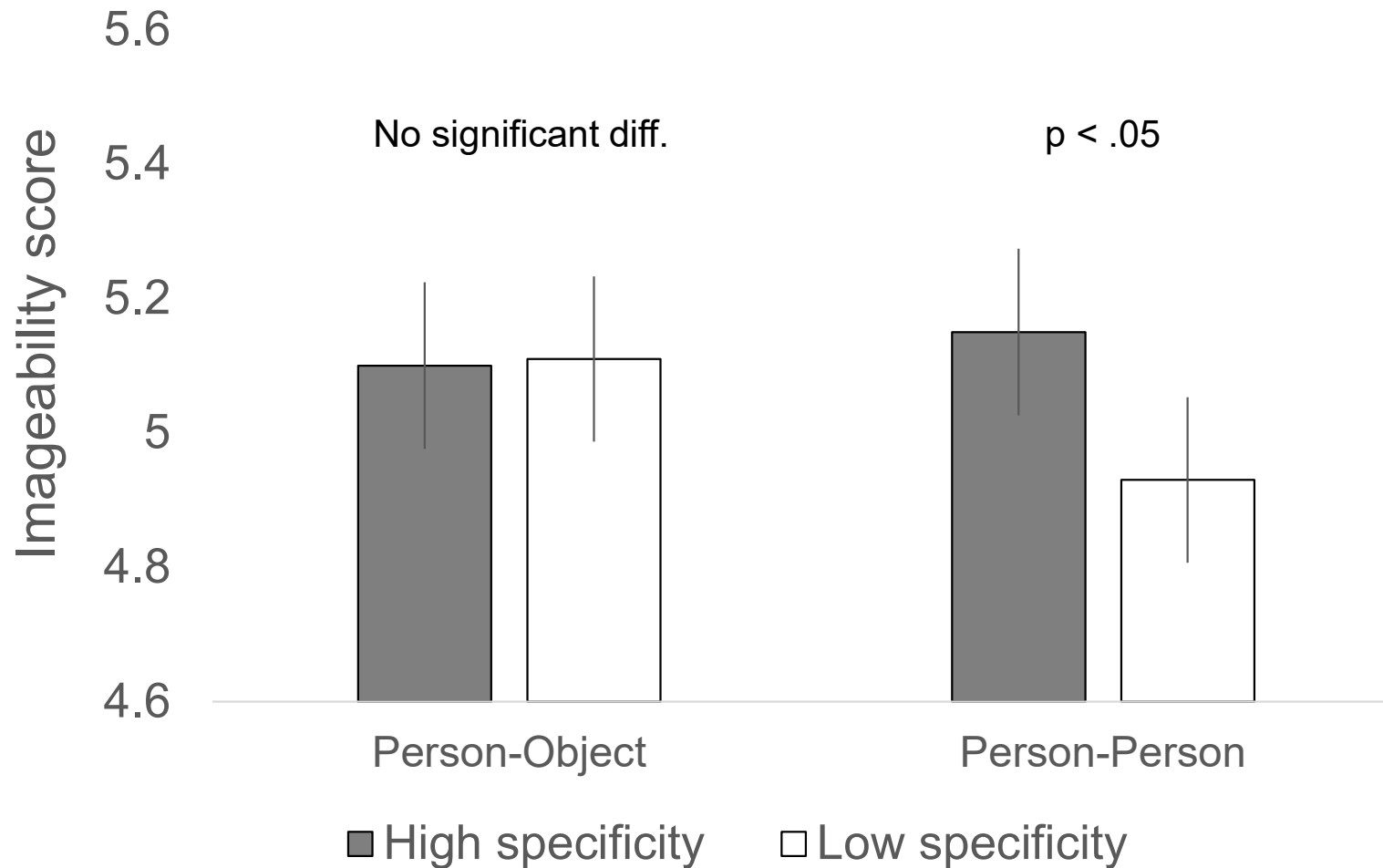
Results – Voice quality

Enjoyment



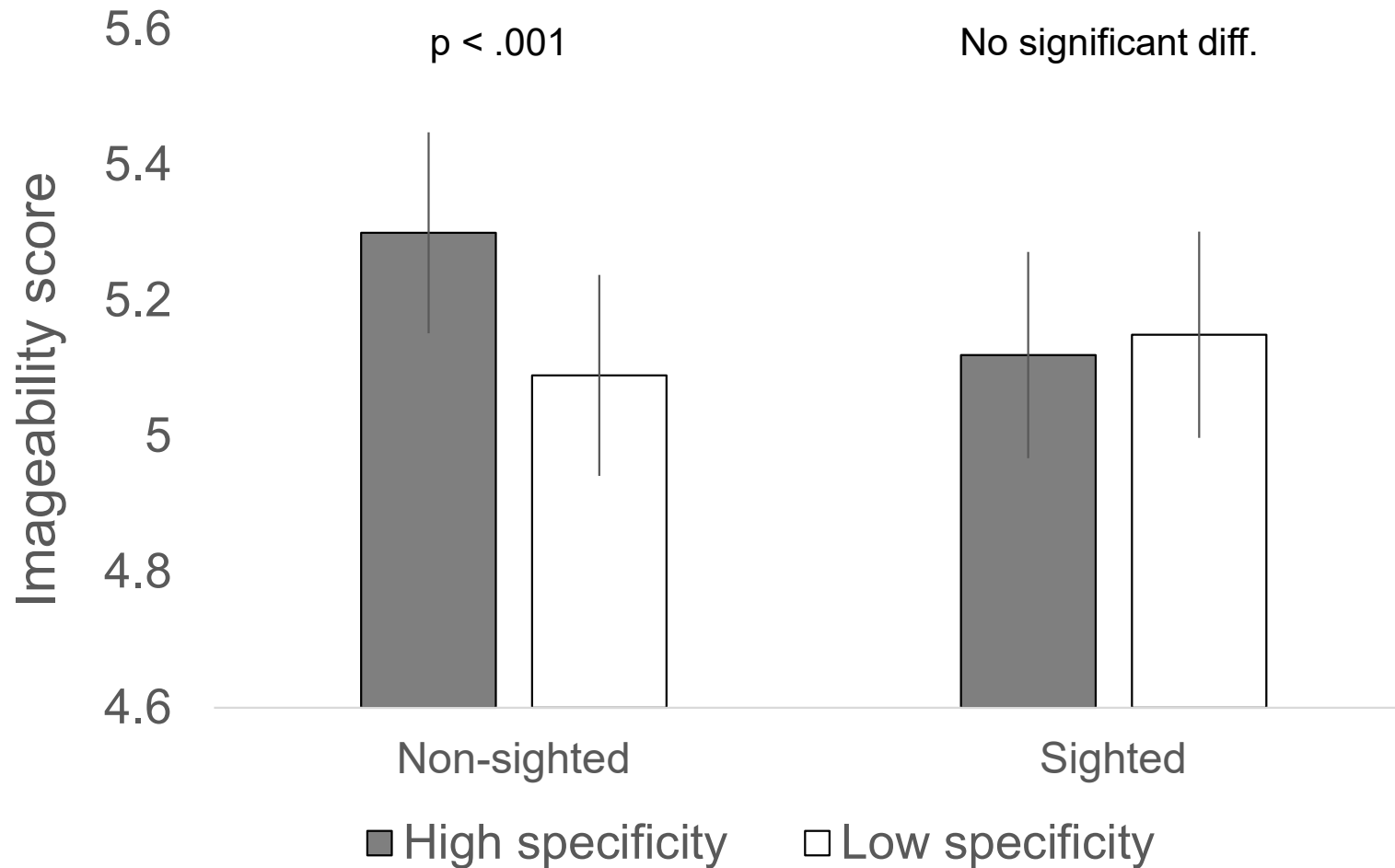
Results - imageability

Event descriptions of spatial relations



Results - imageability

Event descriptions of motion



Summary

- **Voice quality** in verbal narration of visual events play a major role in **listening effort** and **enjoyment** of listening.
 - Big difference for sighted and non-sighted listeners!
- **The specificity** in how **spatial relations** and **changes in motion** are described plays a **major role for the imageability** of the described content.
 - Primarily for the non-sighted group and especially prominent for motion changes!
- **Empirical evidence** that these factors are important for successful communication between the sighted and the blind.
- Important factors to consider in **audio descriptions** of visual events.

**Thank you for your
attention!**