

Visual-to-auditory conversion methods for sensory substitution: sound spatialization only *versus* cross-modal correspondence

Camille Bordeau¹, Florian Scalvini², Cyrille Migniot², Julien Dubois², Maxime Ambard¹

¹LEAD-CNRS UMR5022, Université de Bourgogne, Dijon, France

²ImViA EA 7535, Université de Bourgogne, Dijon, France



bordeau.camille@gmail.com

Sensory substitution

- Sensory substitution devices (SSDs) convey spatial information for the blind (Kristjánsson et al., 2017).
- Elevation of spatialized sound is misperceived (Wenzel et al., 1993).
- Cross-modal correspondence between pitch and visual height (Spence et al., 2013) is intuitive in a recognition task (Stiles et al., 2015).

Does the cross-modal correspondence between auditory pitch and visual height improve localization abilities with a SSD ?

Conversion schemes

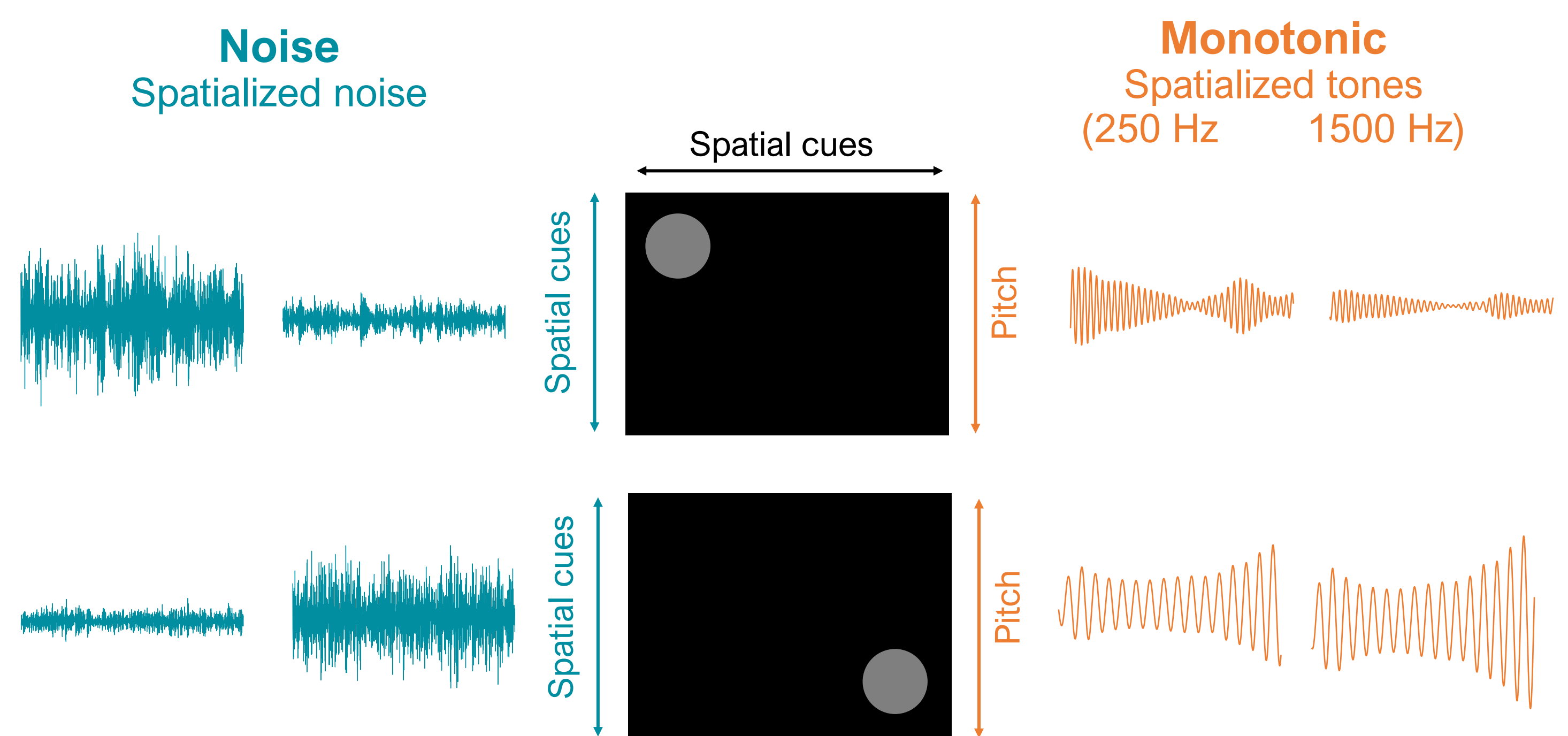
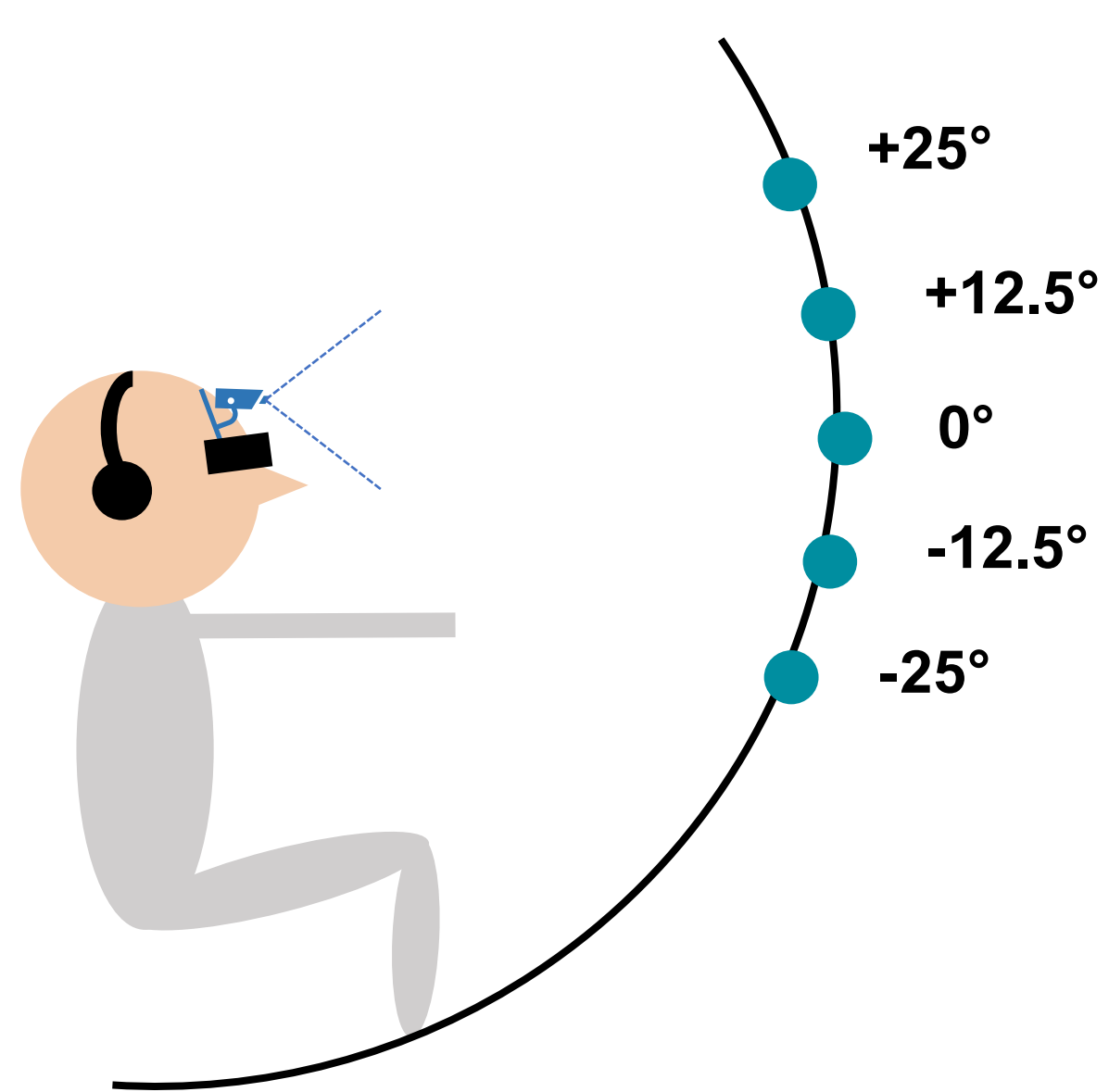
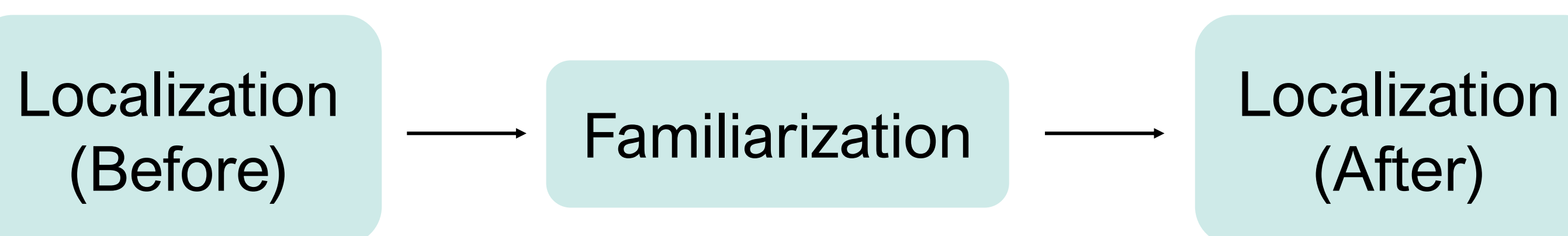


Figure 1. Examples of soundscapes associated to an image with the Noise (left) and Monotonic (right) conversion schemes.

Procedure

- 19 blindfolded participants (age: M = 25.5, SD = 3.04) tested the **Noise** and **Monotonic** conversion schemes



Metrics for localization abilities:

- Error-based (unsigned error)
- Regression-based (response position and gain)

Figure 2. Experimental set-up and possible locations of the target in the localization task.

Localization task

- **Elevation unsigned error was lower with the Monotonic than with the Noise conversion scheme**

Before: Monotonic (M = 31.5°, SD = 27.2) < Noise (M = 40.2°, SD = 37.0), p < .0001.

After: Monotonic (M = 19.8°, SD = 16.3) < Noise (M = 24.9°, SD = 18.4), p < .0001.

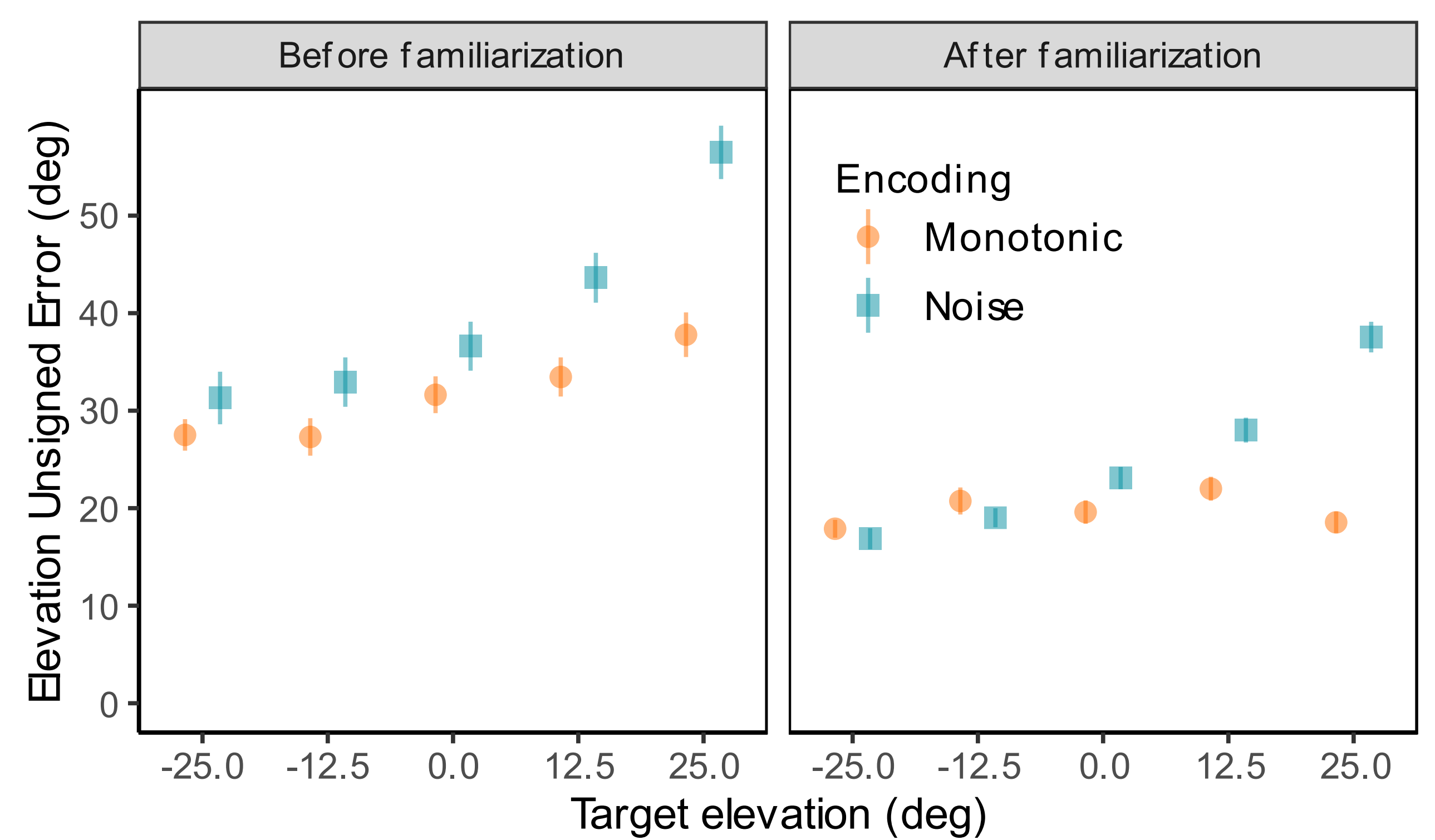
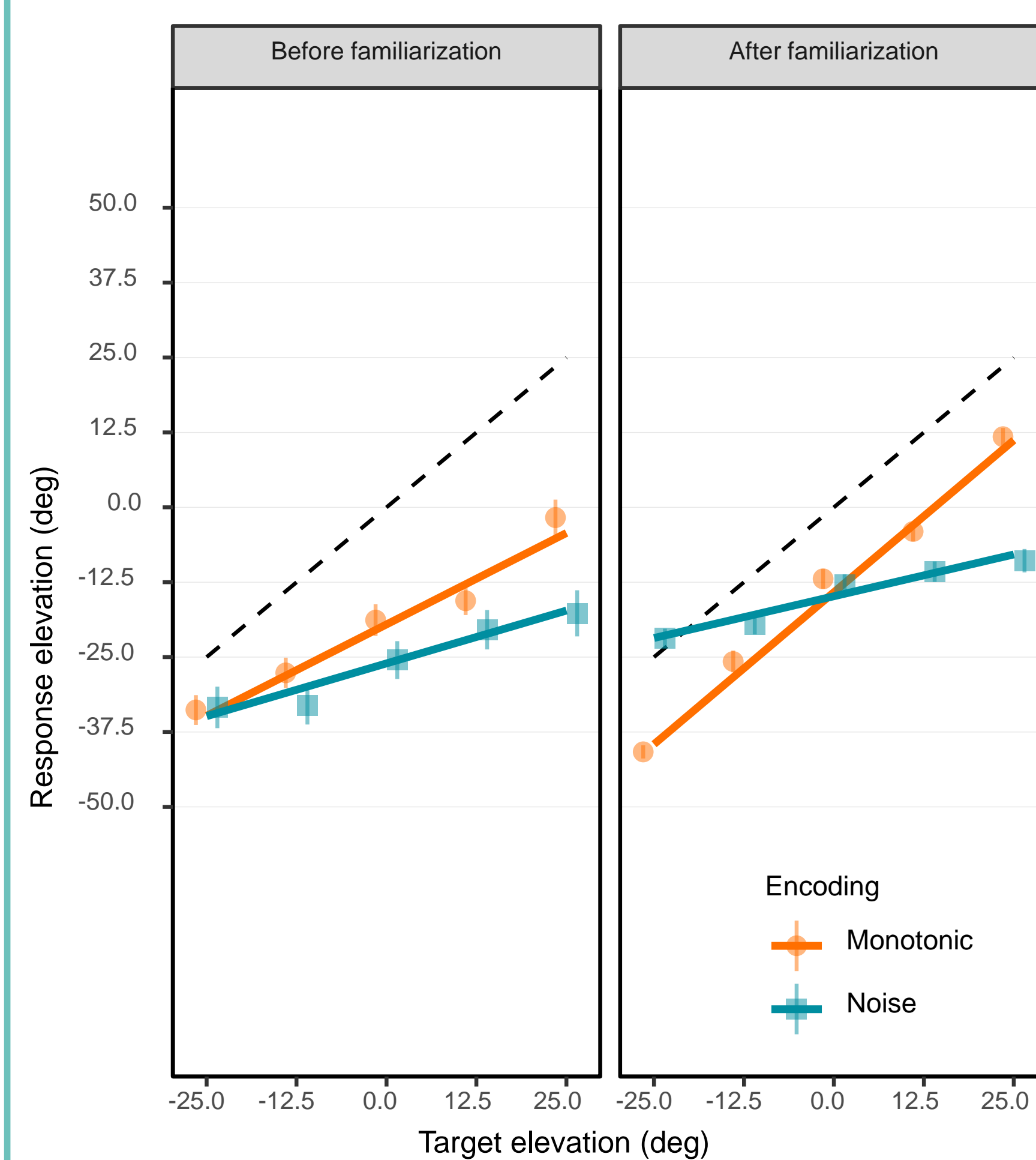


Figure 3. Mean unsigned error in elevation as a function of target elevation, before and after the familiarization.

- **A compression bias was observed? and elevation gains were closer to 1.0 with the Monotonic than Noise conversion scheme**



Before: Monotonic (0.61, 95% CI = [0.49, 0.73]) > Noise (0.35, 95% CI = [0.24, 0.47]), p = .003

After: Monotonic (1.02, 95% CI = [0.9, 1.13]) > Noise (0.28, 95% CI = [0.16, 0.4]), p < .0001

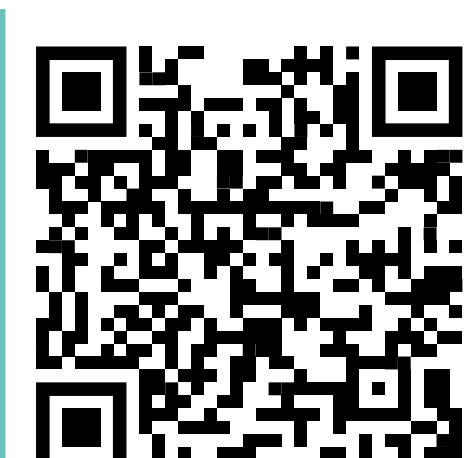
Figure 4. Mean elevation response as a function of target elevation, before and after the familiarization. Solid lines show the estimated trend from the Linear Mixed Model while the dotted line shows the optimal trend.

Discussion

- Facilitation effect of the cross-modal correspondence between auditory pitch and visual height in the early stage of use of a SSD.
- Cross-modal correspondence seems intuitive to localize an object with an SSD.

References

Kristjánsson, R., Moldoveanu, A., Jóhannesson, M. I., Balan, O., Spagnol, S., Valgeirsdóttir, V. V. & Unnthorsson, R. (2016). Designing sensory-substitution devices : Principles, pitfalls and potential. *Restorative Neurology and Neuroscience*, 34(5), 769-787. doi: 10.3233/rnn-160647. Spence, C., & Deroy, O. (2013). How automatic are crossmodal correspondences ? *Consciousness and Cognition*, 22(1), 245-260. doi: 10.1016/j.concog.2012.12.006. Stiles, N. R. B., & Shimojo, S. (2015). Auditory Sensory Substitution is Intuitive and Automatic with Texture Stimuli. *Scientific Reports*, 5(1). doi:10.1038/srep15628. Wenzel, E. M., Arruda, M., Kistler, D. J., & Wightman, F. L. (1993). Localization using nonindividualized head-related transfer functions. *Journal of the Acoustical Society of America*, 94(1), 111-123. doi: 10.1121/1.407089



Scan the QR code to read the full published article

Bordeau, C., Scalvini, F., Migniot, C., Dubois, J., & Ambard, M. (2023). Cross-modal correspondence enhances elevation localization in visual-to-auditory sensory substitution. *Frontiers in Psychology*, 14. doi: 10.3389/fpsyg.2023.1079998

Thanks to the Conseil Régional de Bourgogne Franche-Comté, France and the Fond Européen de Développement Régional (FEDER) which are supporting financially this research.