Exploring the Relation between Gesture Presentation Perspective and Children's Spatial Performance

The presence of gestures and visuospatial materials enhances the children's spatial performance.[1,2] While describing route descriptions to children using tangible visuospatial maps, using multimodal (i.e., speech-gesture combinations) description strategies enhances their recall compared to speech-only descriptions. Although it is an emergent context in online educational tools, no study so far investigated whether and how the perspective of a speaker's gesture presented to children relates to differences in their spatial performance. Here, we investigate whether the perspective of multimodal input relates differently to 5-year-old monolingual Turkish children's spatial performance by varying the perspective of gesture presentation with a virtual visuospatial map.

5-year-old monolingual Turkish children were engaged in the Directions Task presented on a tablet screen. The task included visuospatial maps and videos of speakers describing routes in the map by three conditions: Speech-Gesture combination with a front-facing view (SG-FF), Speech-Gesture combination with an upper back angle (SG-UB), and Speech only conditions with a front-facing view (SO-FF) for control (see Figure 1). Children were asked to recall the route described to them in the videos after each trial. Children's responses were coded for the total correct recall of target information in the route descriptions for action (e.g., jumping), location (e.g., house), and spatial (e.g., behind) information.

Three repeated measures of ANOVAs were conducted for each target information type to analyze the differences between conditions (SG-FF, SO-FF, and SB-UB). Children's performances in SG-FF are higher than SO-FF ($p = 0.013$) for spatial information (see Figure 2a). Moreover, children's performances in SG-UB condition are significantly higher compared to SG-FF ($p < 0.001$) and SO-FF ($p = 0.036$) conditions for action information (see Figure 2b). However, no significant differences were observed for location information across conditions (see Figure 2c).

Summarizing, the results suggest that the presentation perspective of gestures showed varying effects on children's spatial performance. That is, spatial and action encoding is higher in multimodal Speech-Gesture conditions than in Speech-Only. However, children's recall of action information benefits more from an upper-back presentation perspective, which aligns with their perspective of the map. The results of this study underline the importance of multimodal input and presentation perspective in enhancing children's spatial performance.

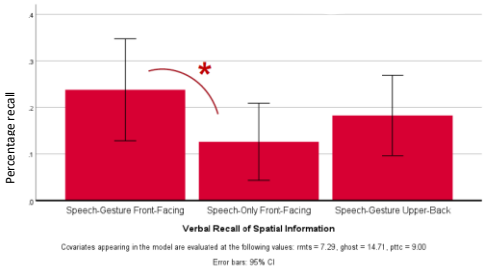
Keywords: Multimodal Input, Perspective, Spatial Understanding, Visuospatial Maps

Figure 1

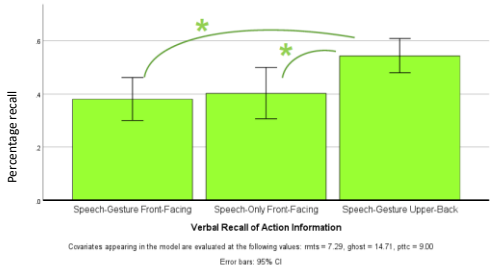
SO-FF **SG-FF** **SG-UB**

Figure 2

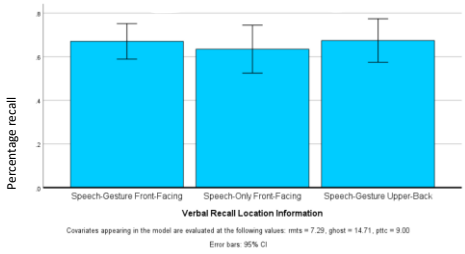
a. Verbal Recall of Spatial Information (SG-FF, SO-FF, SG-UB)



b.Verbal Recall of Action Information (SG-FF, SO-FF, SG-UB)



c. Verbal Recall of Location Information (SG-FF, SO-FF, SG-UB)



References

[1] Austin, E. E., & Sweller, N. (2014). Presentation and production: The role of gesture in spatial communication. *Journal of Experimental Child Psychology,* 122, 92–103. https://doi.org/10.1016/j.jecp.2013.12.008.

[2] Simms, N. K., & Gentner, D. (2019). Finding the middle: Spatial language and spatial reasoning. *Cognitive Development*, 50, 177-194. https://doi.org/10.1016/j.cogdev.2019.04.002