

Speech cues modulate gesture comprehension in a levels-of-congruence paradigm

Laura Pissani (laura.pissani@uni-saarland.de) & Vera Demberg

Department of Language Science & Technology, Saarland University

Introduction. Language and gesture are argued to form an integrated system, in which each modality contributes to the underlying meaning of utterances [1]. As McNeill puts it, gestures “are integrated with linguistic content into growth points and appear, to the speaker, as an unbroken package of semiosis with it.” The notion of an “unbroken package” implies a bidirectional relationship between speech and gesture, in which gestures influence how language is comprehended and language influences how gestures are processed. In a recent study, [2] developed a levels-of-congruence paradigm in which a speaker produced metaphors, such as *Men are oaks*, under three gesture conditions. In one condition, gestures matched the preferred interpretation of the metaphor (both arms flexed to convey robustness). In a second condition, gestures reflected a possible but less preferred interpretation (hand raised above the head to convey tall stature). In a third condition, no gesture was produced. After each video, participants were presented with three written interpretations (preferred, possible, impossible) and asked to rate how well the gesture matched each (Experiment 1) or to choose the best interpretation (Experiment 2). Across all gesture conditions, the preferred interpretation obtained the highest ratings and was selected most often. However, when gestures conveyed the possible interpretation, ratings and probabilities of being selected increased significantly for the possible interpretation. These findings indicate that co-speech gestures modulate metaphor comprehension—although gestures alone are not sufficient to override default interpretations. But does speech also influence the comprehension of gestures? In this study, we employed the levels-of-congruence paradigm to investigate whether the comprehension of gestures is similarly modulated by the presence of additional multimodal cues. **Method.** To examine the effect of multimodal cues, we created three multimodal conditions. In the baseline condition, clips were muted with blurred lips (Figure 1a). In the lip-movement condition, clips were muted with visible lips (Figure 1b). In the speech condition, clips had visible lips and audible speech (Figure 1c). After each clip, three interpretations were presented, each with a five-point rating scale. The gestured interpretation (e.g., *robust*) corresponded directly to the gesture, the alternative interpretation (e.g., *tall*) reflected another interpretation of the metaphor but was unrelated to the gesture, and the unrelated interpretation (e.g., *hikers*) was unrelated to both the gesture and the speech. Participants (N = 150, age 21–70 years, M = 41.98, SD = 12.03; 84 female) rated how well each interpretation matched the gesture in a single assigned multimodal condition. **Results.** The coefficients of the regression model are summarized in Table 1. There was a main effect of interpretation, with alternative and unrelated interpretations receiving lower ratings than gestured interpretations. There was also an interaction between multimodal condition and interpretation, with Bayes factors providing extreme evidence in favor of the full model including the interaction, $BF_{10} = 10^{16}$. As shown in Figure 2, the gestured interpretation received similar ratings across multimodal conditions; however, ratings for the alternative interpretation increased in the speech condition compared to the baseline condition with no additional cues ($\beta = 0.84$ log odds, 95% CrI [0.62, 1.06]). As predicted, the perceived match between the gesture and the alternative interpretation increased when the metaphor was audible. **Discussion.** This pattern is particularly interesting because it shows that participants incorporated the alternative figurative sense of the metaphor in their responses even when it was not needed to evaluate the gesture. When the actor performed the ‘robust’ gesture while the metaphor *Men are oaks* was audible, ratings for the alternative interpretation *tall* increased compared to when the metaphor was inaudible, despite *tall* being unrelated to the observed gesture. These findings indicate that participants not only attended to the gesture but also integrated it with both figurative senses of the speech. Altogether, these findings support the integrated-systems hypothesis [3], in which speech and gesture are mutually and obligatorily integrated. In this study, gestures presented without additional cues were already interpretable, but when speech cues were available, participants also considered alternative interpretations of the gesture that were consistent with the verbal content.

References

- [1] McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. U of Chicago P.
 [2] Reference omitted
 [3] Kelly, S. D., Özyürek, A., & Maris, E. (2010). Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension. *Psychological science*, 21(2), 260-267.

Figures and Tables

Figure 1: Video frames of the same ‘robust’ gesture across the three multimodal conditions.

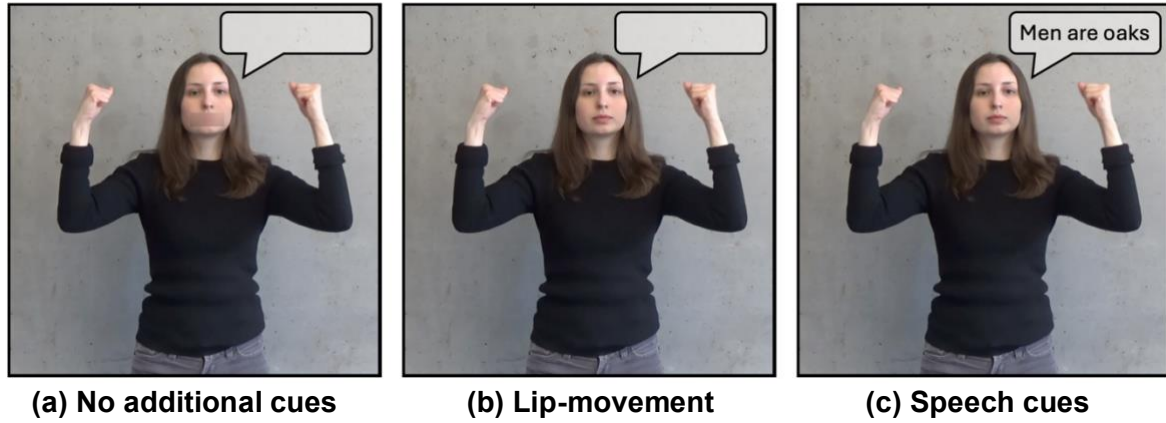


Figure 2: Ratings for the gestured interpretation, alternative interpretation of the metaphor, and unrelated interpretation across multimodal conditions.

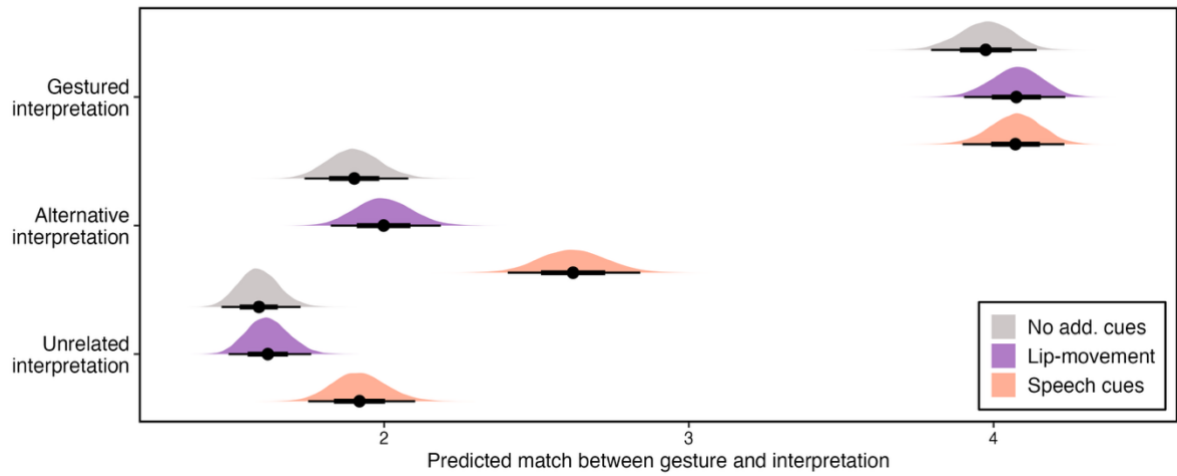


Table 1: Bayesian regression coefficients for log-odds changes in ratings.

Parameter	β	95% CrI
Alternative interpretation	-2.89	[-3.06, -2.73]
Unrelated interpretation	-3.49	[-3.66, -3.32]
Lip-movement	0.19	[-0.13, 0.50]
Speech cues	0.20	[-0.12, 0.51]
Alternative $\times$ Lip-movement	-0.04	[-0.26, 0.17]
Unrelated $\times$ Lip-movement	-0.14	[-0.37, 0.09]
Alternative $\times$ Speech	0.84	[0.62, 1.06]
Unrelated $\times$ Speech	0.41	[0.19, 0.64]