Does contrastive attention guidance facilitate action recall? An eye-tracking study

Amit Singh and Katharina J. Rohlfing

Faculty of Arts and Humanities, Paderborn University

1. Introduction

Studies suggest that language interferes with motor system which in turn influences action perception at a very early stage [2]. This top-down influence of language on non-linguistic processes can be used to guide observer's attention in an action demonstration [1]. Studies of this type often use a simple assertive verbal description concurrently with the action, and ignores the context of the motion event as it unfolds. Given that the current sub-event is interpreted in light of the past event, a simple assertive guidance might not be suitable to demonstrate an action varying along the perceptual dimension e.g., visual contrast. Till date the relationship between verbal and visual contrast in action understanding has barely been investigated. To address this question, in a recall task we evaluated the effect of verbal contrasts (assertion and negation) on the visually contrastive and non-contrastive motion. We use negation since it creates verbal contrast and provides a rich contextual information when interpreted against it's positive counterpart [3].

2. Aims

1. To test whether a verbal contrast - a sequence of assertive and negative description of the action path - enhances recall of a visual contrastive action as opposed to a sequence of simple assertive description.
2. Using an eye-tracker, we evaluated whether verbal instructions averted observer's attention from the goal, which otherwise is the focus of attention in a motion event [4]

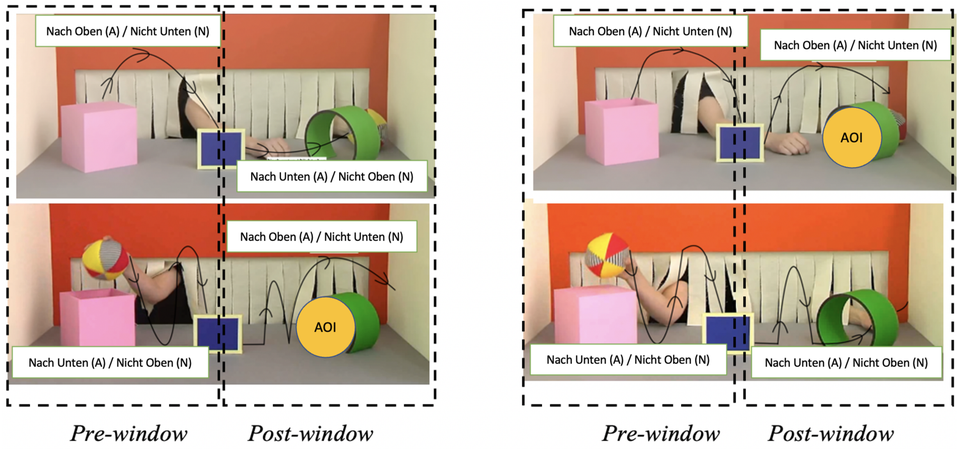

Figure 1: Contrastive and non-contrastive action paths, each window was preceded by either assertive, negative or no voice instruction

3. Methods

1. Participants: 30 students, age(mean) = 23.90
2. Stimulus: Four videos in which a ball was moved against three landmark objects creating non-contrastive (Up-Up/Down-Down) or contrastive (Up-Down/Down-Up) action sequence
3. Conditions: Each video segment (pre-and post) was accompanied by assertive or negative path description, creating a sequence of either two assertives (e.g., Up-Down), two negatives (not Down-not Up) or assertive-negative (Up-not Up) instructions, where a video without instruction (no voice) was treated as a baseline
4. Procedure: Participants saw the videos on an eye-tracking screen and then performed the action on a stage. All the trials were counterbalanced and presented in random order

4. Results

1. A significant main effect of action path such that recall for contrastive action was higher than non-contrastive action sequence(Fig. 2)
2. A significant main effect of voice such that recall for assertive-assertive was higher than no voice (Base), suggesting that assertive instructions were overall helpful for action recall

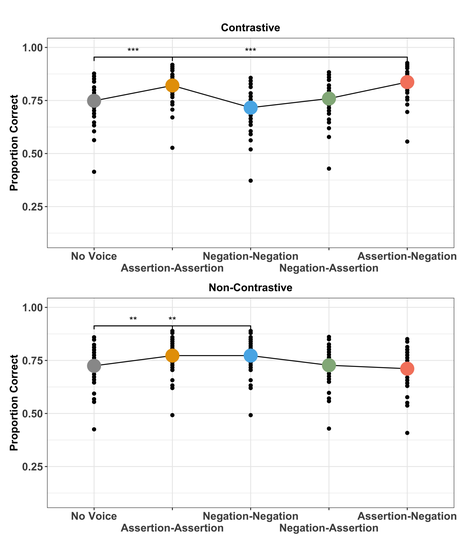

Figure 2: Recall task[1]

3. A significant interaction between path and voice; a pairwise comparison shows that the assertive-negative voice condition enhanced the recall for the contrastive action paths i.e., Up-Down or Down-Up
4. Eye-Tracking results: Voice significantly reduced the fixation on the goal such that it was maximum in no voice (Base) condition which reduced maximally in assertion-negation voice condition (Fig. 3)

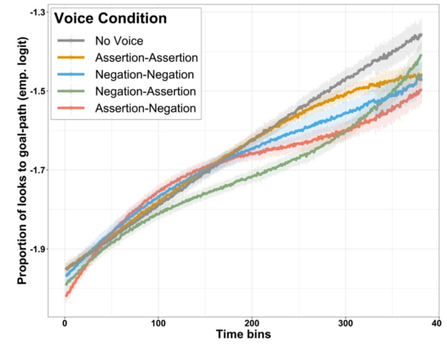

Figure 3: Mean fixation on goal (post-window region)[2]

5. Discussion

1. A contrastive action contains sub-actions where the later is the opposite of former, we show that such actions can be better demonstrated by combining assertion-negation instruction which reveals a rich contextual information that cannot be achieved alone by assertive instruction [3]
2. Most studies treat action as discrete sub-events by giving only assertive verbal instruction, we show that negation instead can be used to provide a contextual understanding of a contrastive action
3. Contrastive instructions avert attention from the goal, directs it back to the path of the action, hence facilitates its recall which otherwise remains relatively hidden in an action demonstration

References

[1] K. J. Rohlfing, J. Fritsch, B. Wrede, and T. Jungmann.
How can multimodal cues from child-directed interaction reduce learning complexity in robots.
Adv. Robot., 20:1183–1199, 2006.

[2] A. Sciutti, K.S. Lohan, G. Gredebäck, B. Koch, and K.J. Rohlfing.
Language meddles with infants' processing of observed actions.
Front. Robot. AI, 3(46), 2016.

[3] P. Wason.
The contexts of plausible denial.
Journal of Verbal Learning and Verbal Behavior, 4:7–11, 1965.

[4] J. M. Zacks and B. Tversky.
Event structure in perception and conception.
Psychol. Bull., 127:3–21, 2001.

Acknowledgements and Contact

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - SFB Transregio 318, project A05.

- Web: https://amits1ngh.github.io
- Email: amit.singh@uni-paderborn.de

[1] Significance codes: *** = $p \leq 0.001$, ** = $p \leq 0.01$

[2] For both contrastive and non-contrastive condition, fixation pattern on the goal object was same i.e., maximum in no voice (Base) condition which decreased with instructions