

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Grounding Action Verbs in Egocentric Visual Perception

Permalink

<https://escholarship.org/uc/item/9h7090tp>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors

Zhang, Yayun
Cain, Ellis
Crandall, David
et al.

Publication Date

2022

Peer reviewed

Grounding Action Verbs in Egocentric Visual Perception

Yayun Zhang

University of Texas - Austin, Austin, Texas, United States

Ellis Cain

University of California, Merced, Merced, California, United States

David Crandall

Indiana University, Bloomington, Indiana, United States

Chen Yu

University of Texas at Austin, Austin, Texas, United States

Abstract

It has been conjectured that verb learning is hard because verb meanings are not readily "packaged" from the physical world. To provide new empirical evidence on this account, we analyzed egocentric video collected from natural toy-play interaction and focused on the naming events when action verbs were uttered in parent speech. Using the Human Simulation Paradigm, we showed egocentric videos of those naming events to adult observers and asked them to guess the target verb in parent speech. We found that adult observers used many different verbs to describe the same visual event, and only one of them matched with the verb in parent speech. We analyzed mismatched verbs and identified several sources of mismatch, and found that all of the mismatched verbs are relevant to the target verb, but they capture different properties (temporal, semantic, etc) of the visual event. We also found that different naming events for the same verb also differ in terms of the degree of ambiguity. Taken together, the results in the present paper provided new evidence from the child's view, showing that verb learning is hard not only because multiple possible meanings are embedded in each learning situation, but also because these candidate meanings expand across multiple dimensions of the physical world, overlap with each other, and relate to the target meaning in many different ways.