

Lexical Access to Scene Function Relies on Anchor Object Presence

Poster Presentation 23.305: Saturday, May 18, 2024, 8:30 am – 12:30 pm, Banyan Breezeway

Session: [Scene Perception: Miscellaneous](#)

[Schedule of Events](#) | [Search Abstracts](#) | [Symposia](#) | [Talk Sessions](#) | [Poster Sessions](#)

Lea Alexandra Müller Karoza¹ (muellerkaroza@psych.uni-frankfurt.de), Sandro Luca Wiesmann¹, Melissa Le-Hoa Vo¹; ¹Goethe University Frankfurt, Scene Grammar Lab, Germany

Effortless engagement with our surroundings relies on the purposeful arrangement of functional elements within a room for efficient use. Central to these functional clusters are so-called anchor objects, predicting the location of local objects (e.g., soap and toothbrushes near the sink-anchor). Additionally, scene categorization relies on scene function, or affordance, more than on objects (Greene et al., 2016). To examine how anchors affect affordance understanding, we primed a lexical decision task (LDT) on action words (i.e., “washing hands”) with scene images lacking related (REL; sink) or unrelated (UNREL; shower) anchors, or random objects (RAND; shelves). Images from other categories (e.g., a kitchen) served as controls. In Experiment 1, stimuli comprised real photos with pixel masks hiding objects. Participants were quickest in the RAND condition and slower when an anchor was missing, regardless of its action relevance (REL or UNREL). In Experiment 2, using 3D-rendered scenes with whole objects removed, participants were fastest in the RAND condition. Notably, removing a related anchor impeded lexical decision more than removing unrelated anchors. To ensure that visually sparser rendered scenes were still identifiable when anchors were absent, participants categorized the scenes in a control experiment. Scene type (real versus 3D rendered) and missing object (random versus anchor) impacted categorization, with no interaction, suggesting that reported differences between experiments are not solely due to varying scene categorization, but might rely on realism or the information remaining in scenes (i.e., prevalent objects in rendered scenes versus clutter, context- and texture-cues in photographs of real scenes). Experiment 1 implies scene-level affordance understanding using photos and pixel masks, while Experiment 2 suggests object-specific affordance understanding with sparser 3D-rendered scenes. Conclusively, understanding scene affordances flexibly involves both specific object-level information and broader scene context depending on their diagnosticity for scene affordance assessment and the context provided by the scene.

Acknowledgements: This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project number 222641018 SFB/TRR 135 TP C7 granted to MLHV and the Hessisches Ministerium für Wissenschaft und Kunst (HMWK; project ‘The Adaptive Mind’).