

Children and Adults Integrate Complex Visual Contexts in Language Prediction

Linda Sommerfeld¹, Maria Staudte², Nivedita Mani³, Jutta Kray¹

¹Department of Psychology, Saarland University, Germany

²Department of Language Science and Technology, Saarland University, Germany

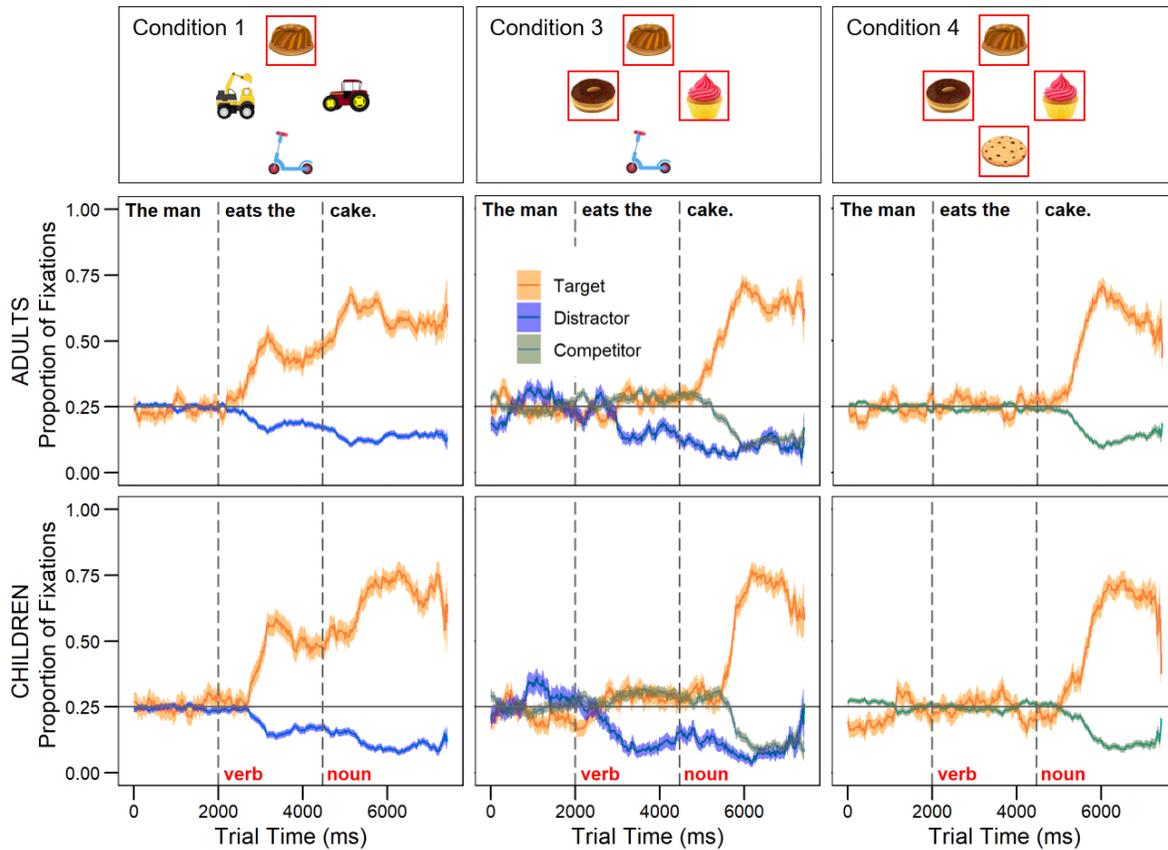
³Department of Psychology, University of Göttingen, Germany

(Correspondence should be addressed to Linda Sommerfeld, lindas@coli.uni-saarland.de)

Not only adults, but already children anticipate upcoming input in sentences with semantically constraining verbs. In the visual world, the sentence context is used to anticipatorily fixate the only object matching potential sentence continuations^{ab}. Adults process multiple visual cues in parallel when predicting language^c. Here, we examined whether young children can also process multiple visual objects fitting the constraining verb of a sentence. We also examined if their processing of multiple potential referents is affected by language skills, given that vocabulary size modulates children's prediction^b. German children (5-6 years; $n=26$) and adults (19-40 years; $n=33$) listened to 32 subject-verb-object sentences with semantically constraining verbs (e.g., The man *eats* the cake.), while looking at visual scenes of four objects. The number of objects being consistent with the verb's constraints (e.g., being *edible*) varied between 0, 1, 3, and 4. The semantic fit of verbs and visual objects as well as their name agreement was normed in a separate study (4-6 years; $N=40$). Results of a linear mixed effects model on the proportion of target fixations with the effect coded factors age, condition, and time showed that both groups (i) anticipatorily fixate the only object out of four visual stimuli matching the sentence constraints upon hearing the verb, and (ii) anticipatorily fixate three or four objects that match the verb, while (iii) objects not being in line with prediction (i.e. the distractors) are fixated less (see Fig. 1). This confirms that young children integrate complex visual contexts as extensive and differentiated as adults when predicting. But, children with higher vocabulary size, implied by the Peabody Picture Vocabulary Test, fixate potential targets more often upon hearing the verb than those with lower scores ($r=.13$, $p<.001$), hinting that language skills affect children's anticipatory language processing in the visual world.

Figure 1

Fixations to Targets, Competitors, and Distractors over Averaged Trials in Condition 1, 3, and 4.



Note. Either 1, 3, or 4 potential target objects (marked with squares) given the sentence *The man eats the cake.* Condition 1: Target fixations (orange line) increase, distractor fixations (blue line) decrease after hearing the verb. Condition 3 and 4: Target fixations and competitor fixations (green line) predominate upon the verb, whereas distractor fixations decrease (condition 3).

References

[°]Ankener, C., Drenhaus, H., Crocker, M. W., & Staudte, M. (2018). Multimodal surprisal in the

N400 and the Index of Cognitive Activity. In: *40th Ann. Conf. of the CogSci*, Madison, WI.

^bBorovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill

and not age is associated with anticipatory incremental sentence interpretation in

children and adults. *Journal of Experimental Child Psychology*, 112(4), 417.

^aReuter, T., Dalawella, K., & Lew-Williams, C. (2020). Adults and children predict in complex and

variable referential contexts. *Language, Cognition and Neuroscience*, 36(4), 1–17.