

# Multimodal predictors of early object noun recognition in Tseltal



Kennedy Casey<sup>1,2</sup> & Marisa Casillas<sup>1</sup>

<sup>1</sup>Department of Comparative Human Development, University of Chicago

<sup>2</sup>Department of Psychology, Princeton University

Contact: kcasey@princeton.edu

## Background

Caregiver object labeling and child object handling make unique contributions to early word learning in Western communities<sup>1,2</sup>

Ethnographic accounts of Tseltal (Mayan) early language environments describe them as non-child-centered and non-object-centered<sup>3-5</sup>

Early Tseltal vocabularies include an overrepresentation of verbs, and not nouns<sup>6</sup>

**What are the contributions of adult object labeling and child object handling to Tseltal children's early noun learning?**

## Looking-while-listening experiment in the field

### Participants

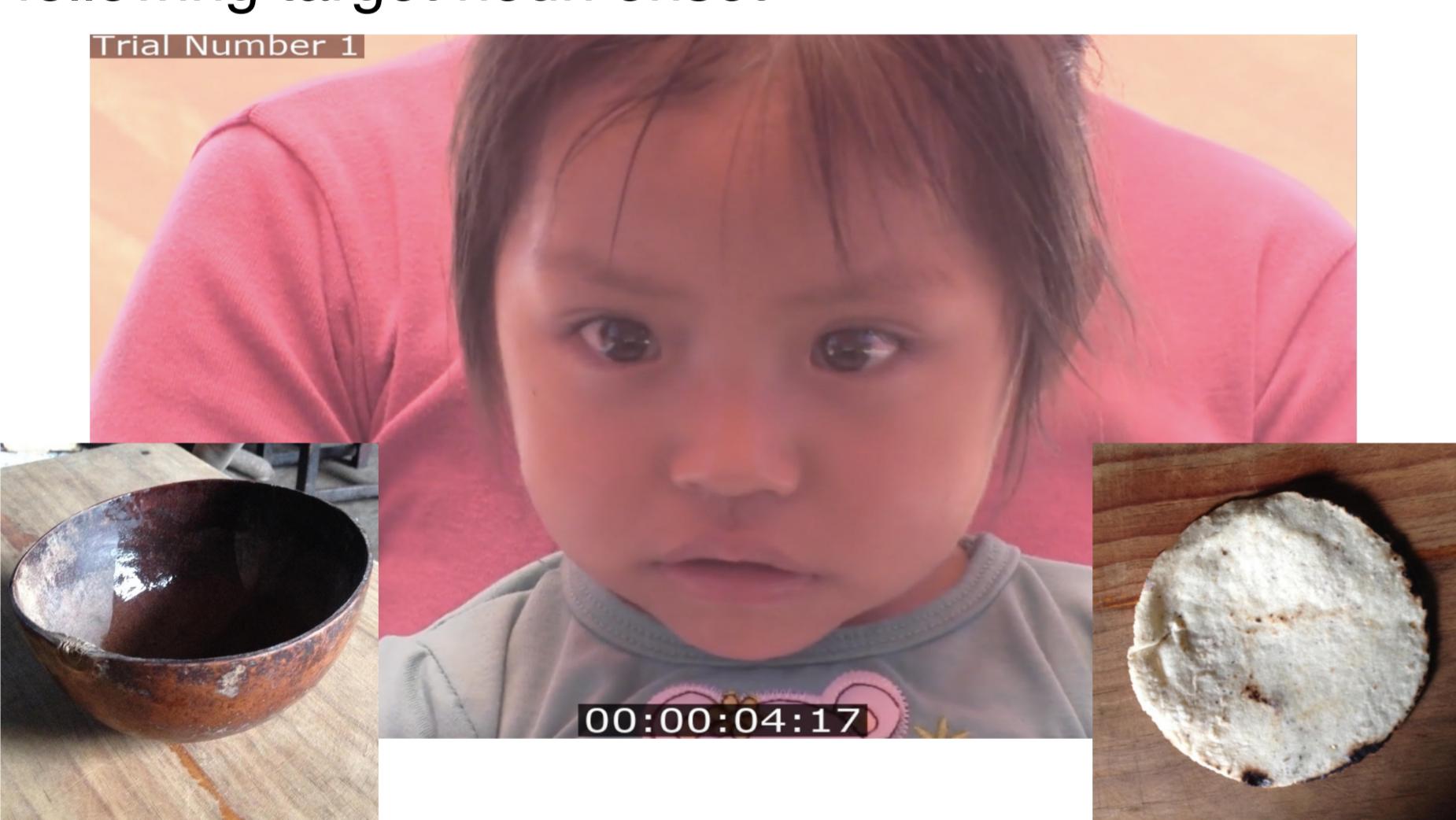
N = 62 Tseltal-learning children  
Age range = 0;09.09–3;11.24 (mean = 2;07.02)

### Design

30 common community-specific object nouns  
“Experi-tent” setup with stimulus images displayed on laptop



Gaze coded offline in ELAN<sup>6</sup> and peyecoder<sup>7</sup>, with analyses based on the window from 365–2005ms following target noun onset



## Input predictors estimated from daylong photo-linked audio recordings

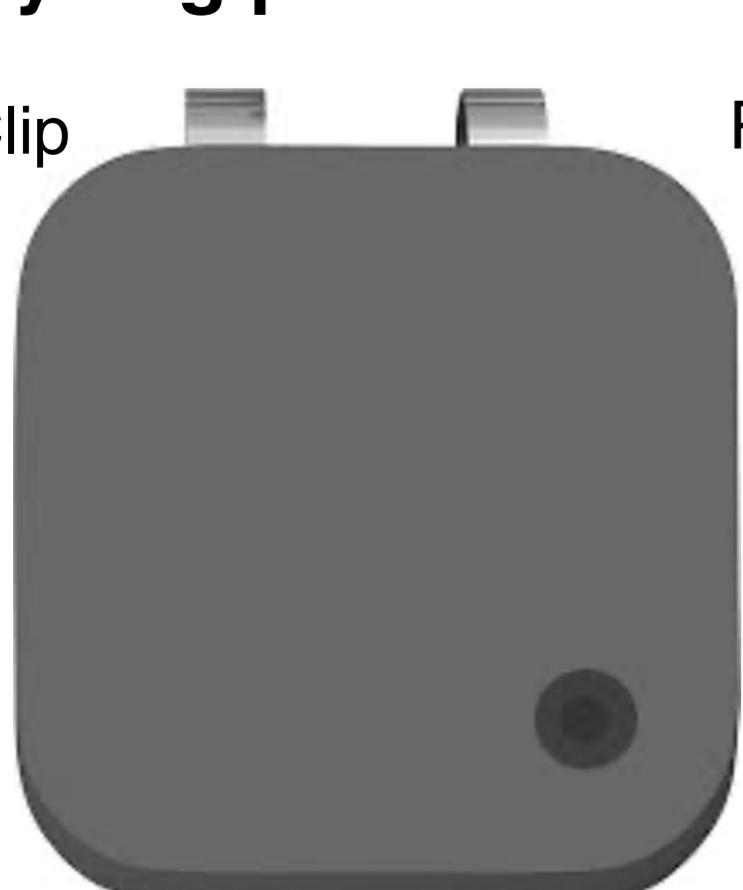
N = 43 children up to age 4;0 (31 overlapping with the LWL sample)

~9 hrs of continuous audio from a child-worn recorder  
Photos captured ~2 times per minute using a child-worn camera with a 180° fisheye lens

**Adult object labeling frequency estimated from 9 x 5-minute transcribed, randomly selected clips (target-child-directed language only)<sup>8</sup>**

**Child object handling frequency counted across all full daylong photo streams<sup>9</sup>**

Narrative Clip camera



Photojojo “super” fisheye lens



## Key findings

Overall above-chance salience-corrected LWL accuracy, measured as the relative increase in proportion looking to a given stimulus image when it appeared as the labeled target vs. unlabeled distracter ( $\beta = 0.19$ ,  $SE = 0.01$ ,  $p < 0.001$ )

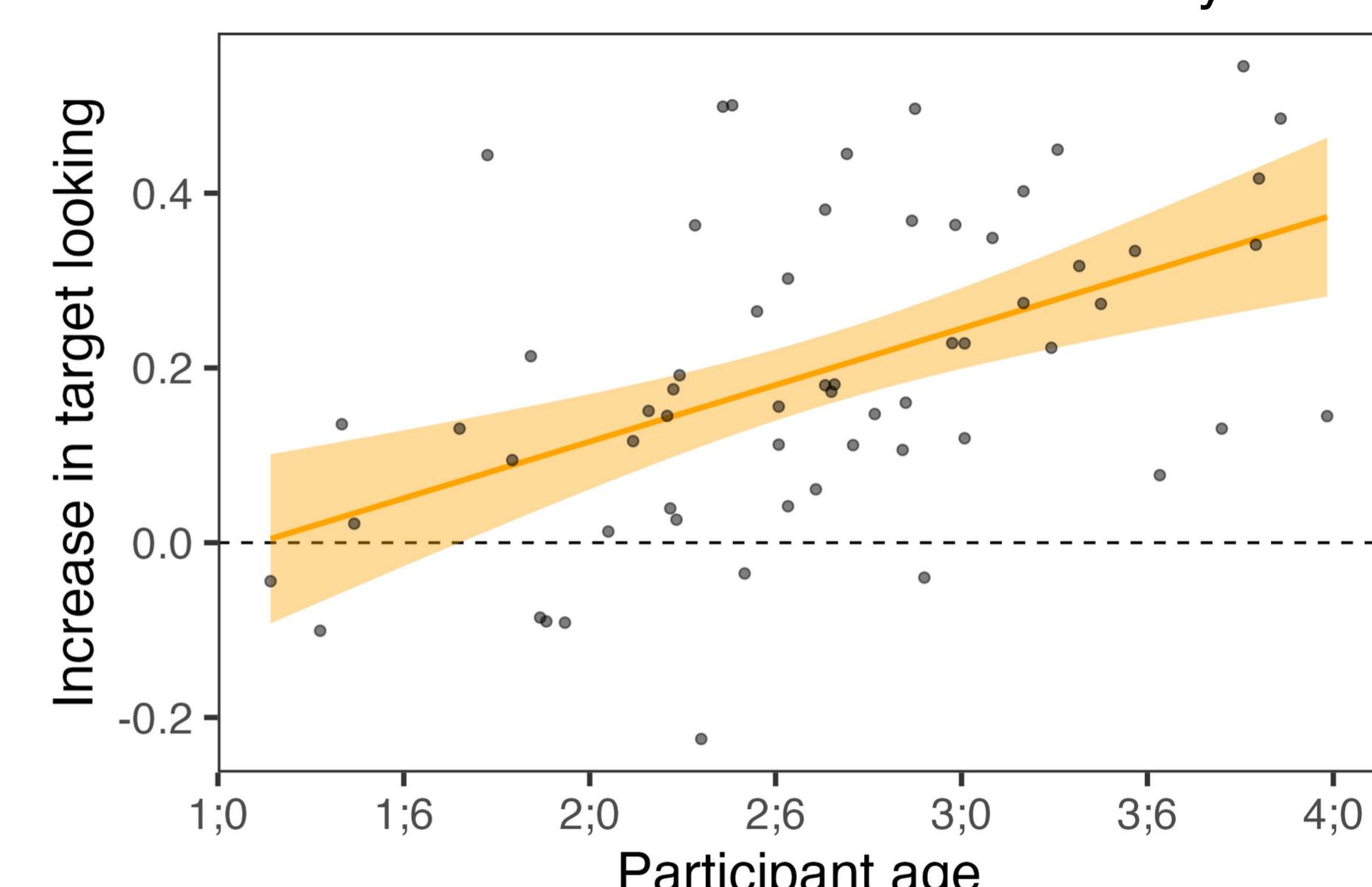
Significant age-related increases in LWL accuracy scores ( $\beta = 0.01$ ,  $SE = 0.002$ ,  $p < 0.001$ ; **Figure 1**) and speed of target noun recognition

No significant effect of **adult object labeling** frequency or **child object handling** frequency on overall accuracy scores ( $p > 0.05$ ; **Figure 2**)

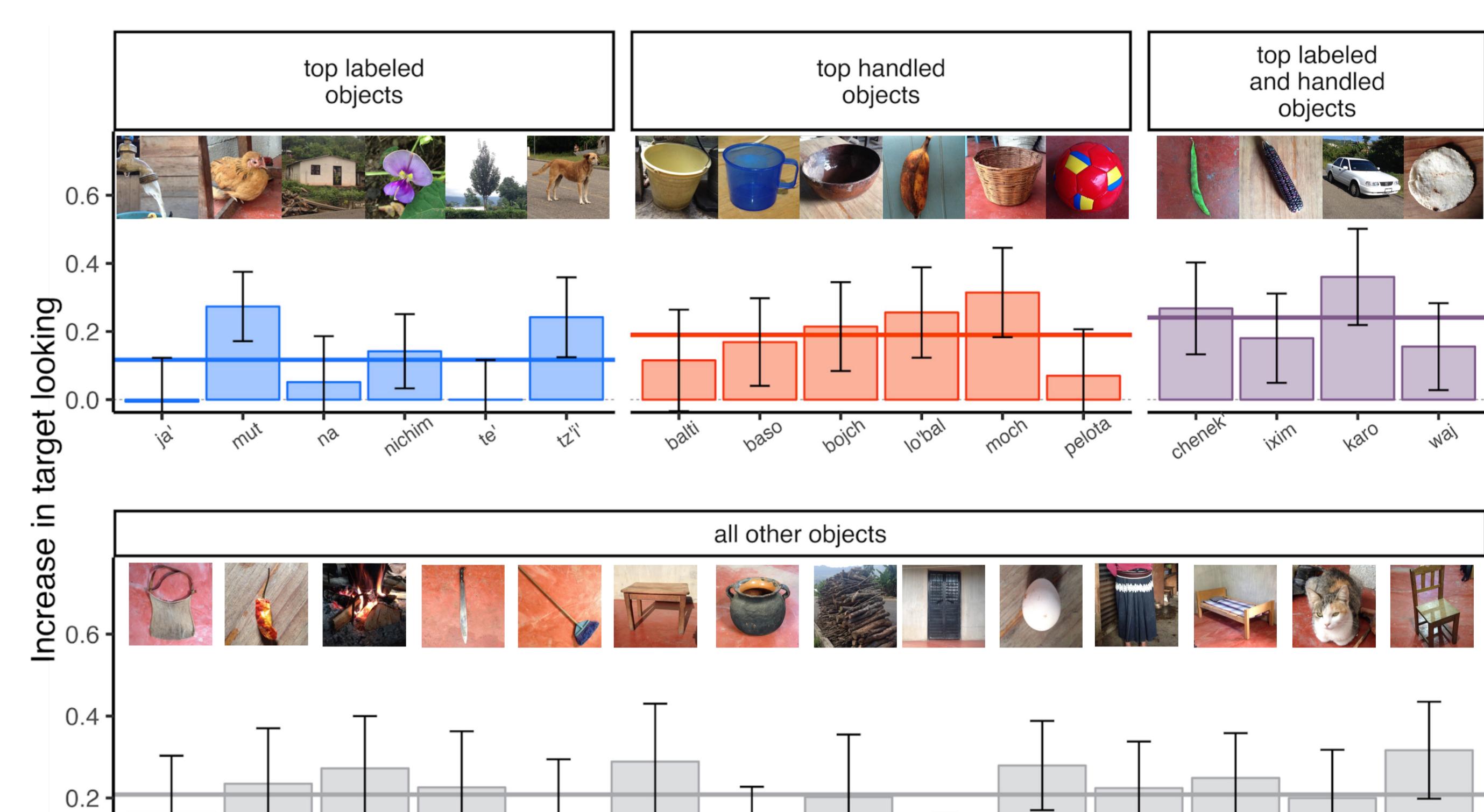
Exploratory growth-curve analyses<sup>10</sup> reveal mixed evidence for the effect of the two multimodal input predictors on looking trajectories (**Figure 3**), including the following notable results:

- Two- and three-way interactions between **labeling**/**handling** frequency and the linear and cubic time terms ( $p < 0.01$ )
- Negative two-way interaction between participant age and **labeling** ( $p < 0.001$ ) but not **handling** frequency ( $p = 0.816$ )

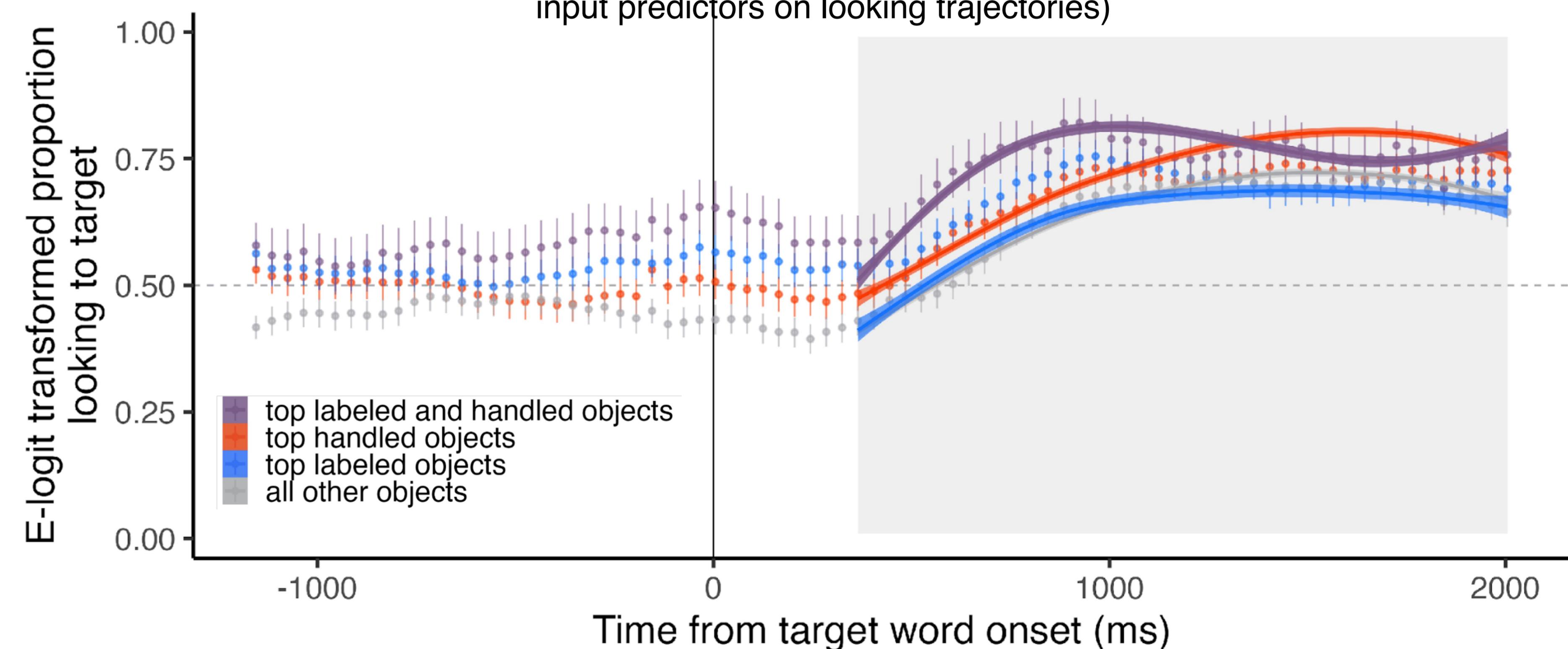
**Figure 1: Age-related increases in overall salience-corrected LWL accuracy**



**Figure 2: Item effects (no overall impact of multimodal input predictors on salience-corrected LWL accuracy)**



**Figure 3: Growth-curve analyses (mixed evidence for the effect of multimodal input predictors on looking trajectories)**



## Discussion

Reliable evidence of early object noun recognition in a non-Western community where child-directed language is relatively infrequent and object noun labeling is relatively rare see also 11

Equivocal evidence for the specific influence of **adult object labeling** frequency and **child object handling** frequency on Tseltal children’s online word recognition, but preliminary results suggest an advantage for objects that are among the most frequently labeled and handled in this Mayan community

Future research is needed to (a) consider the impact of overhearable language on Tseltal children’s word learning, (b) more densely sample children’s multimodal home language experiences, (c) understand the frequency of co-occurrence between object labeling and handling, and (d) test for links between individual children’s input patterns and their word learning outcomes

## References

- [1] Muraki, E. J., Siddiqui, I. A., & Pexman, P. M. (2022). Quantifying children’s sensorimotor experience: Child body-object interaction ratings for 3359 English words. *Behavior Research Methods*, 54, 2864–2877.
- [2] Suarez-Rivera, C., Linn, E., & Tamis-LeMonda, C. S. (2022). From play to language: Infants’ actions on objects cascade to word learning. *Language Learning*, 72(4), 1092–1127.
- [3] Brown, P. (2011). The cultural organization of attention. In A. Duranti, E. Ochs, & B. B. Schieffelin (Eds.), *Handbook of Language Socialization* (pp. 29–55). Malden, MA: Wiley-Blackwell.
- [4] Brown, P. (2014). The interactional context of language learning in Tzeltal. In I. Arnon, M. Casillas, C. Kurumada, & B. Estigarribia (Eds.), *Language in Interaction: Studies in Honor of Eve V. Clark* (pp. 51–82). Amsterdam, NL: John Benjamins.
- [5] Casillas, M., Brown, P., & Levinson, S. (2020). Early language experience in a Tseltal Mayan village. *Child Development*, 91(5), 1819–1835.
- [6] ELAN [Computer software]. (2023). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive.
- [7] Olson, R. H., Pomper, R., Potter, C. E., Hay, J. F., Safran, J. R., Ellis Weismar, S., & Lew-Wiliams, C. (2020). Peyecoder: An open-source program for coding eye movements. Zenodo.
- [8] Casillas, M., Brown, P., & Levinson, S. C. (2017). Casillas HomeBank Corpus.
- [9] Casey, K., Elliott, M., Mickiewicz, E., Silva Mandujano, A., Shorter, K., Duquette, M., Bergelson, E., & Casillas, M. (2022). Sticks, leaves, buckets, and bowls: Distributional patterns of children’s at-home object handling in two subsistence societies. In J. Culbertson, A. Perfors, H. Rabagliati, & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Meeting of the Cognitive Science Society* (pp. 927–933).
- [10] Mirman, D. (2014). *Growth Curve Analysis and Visualization using R*. Chapman; Hall/CRC.
- [11] Foushee, R., & Srinivasan, M. (under review). Evidence for early lexical knowledge in infants primarily exposed to overheard speech.

**Acknowledgements:** We are grateful to the participants and their families for contributing to this research. This work would not have been possible without annotation help from Mary Elliott, Carla Escalante, Cielke Hendricks, Jordyn Martin, & Ine Alvarez van Tussenbroek.