Sticks, leaves, buckets, and bowls: Distributional patterns of children's at-home object handling in two subsistence societies

Anonymous CogSci submission

Abstract

Keywords:

Introduction

The artifacts that we regularly pick up and handle at home—a coffee cup, a laptop, a baby bottle—tell about our routines, aspirations, relationships, and more. These objects offer a window into the physical, social, and cultural contexts that shape our understanding of the world. In this paper, we take a glimpse into everyday life at its beginnings by exploring children's at-home object handling from early infancy until age four. We contextualize our study with respect to the effects of object-centric interaction on word learning, though we note that different analyses of these same data could shed new light on other types of social learning, as well as motor development (see Herzberg, Fletcher, Schatz, & Tamis-LeMonda, 2021 on the latter point).

Object handling and word learning

For young learners, objects—along with their associated activities and surrounding language—form a critical source of input for word learning. Hands (and what they are handling) can be reliable indicators of what someone is doing and talking about during object play and can thus help learners map word forms onto their meanings in and across real-time interaction (e.g., Yu & Smith, 2013; Yurovsky, Smith, & Yu, 2013). The labels of present, attended-to objects are reflected in the babble of children who have acquired stable consonants (Laing & Bergelson, 2020). Further, caregivers' tendency to use nouns referring to objects in the here-and-now positively predicts their children's early word comprehension (Bergelson & Aslin, 2017).

How frequently do children engage in object-centric interactions? First, hands—others' and their own—are in good supply in young children's view of the world, especially after early infancy (Fausey, Jayaraman, & Smith, 2016; Jayaraman, Fausey, & Smith, 2017; Long, Kachergis, Agrawal, & Frank, 2020). Infants' own object handling is also relatively frequent: Herzberg and colleagues (2021) find that US infants handle objects $\sim$60% of the time during at-home play, Yu and colleagues (2013) find $\sim$70% when including joint handling with adults in US in-lab object play, and Casillas and Elliott (2021) find $\sim$15 and 17% object handling in daylong photo streams in a Papuan and a Mayan community, respectively. Concurrent with these events, children will sometimes encounter linguistic information relating to the focused-on object (e.g., its label and associated concepts). However, this critical additional ingredient for word learning may only occur during a small subset of total object handling time. We do not yet know how often objects in the here and now are typically talked about over the course of children's whole waking days at home, but we do know that such talk fluctuates across high and low activity periods (Bergelson, Amatuni, Dailey, Koorathota, & Tor, 2019).

Object handling across age and culture

Children's object handling input changes enormously across the first few years due to both maturational constraints and culture-specific caregiving practices. In early infancy, children have little ability to hold things or to control their posture, primarily experiencing objects through what others bring near to them. Faces, rather than objects, may make up a much greater proportion of their social input early on (Fausey, Jayaraman, & Smith, 2016; Jayaraman, Fausey, & Smith, 2017; but see also Long, Kachergis, Agrawal, & Frank, 2020). However, later gains in manual dexterity and gross motor skill (e.g., sitting, crawling, walking) increasingly widen children's ability to seek, reach, and grab a diversity of objects in their environment. Increasing motor development not only gives children greater control over what objects they handle, but also *how* they elicit social information relating to objects and for how long (Adolph, Karasik, & Tamis-LeMonda, 2010; Gaskins, 2000; Herzberg, Fletcher, Schatz, & Tamis-LeMonda, 2021; Kretch, Franchak, & Adolph, 2014; Sanchez, Long, Kraus, & Frank, 2018).

Early access to objects is also shaped by culture-specific practices for carrying children, keeping them safe and warm, and scaffolding the development of locally valued capacities (e.g., word learning in many US families, walking in Kenyan Kipsigis families: Super, 1976; see Adolph, Karasik, & Tamis-LeMonda, 2010, for an overview). The array of objects available to children will also vary in type and prevalence across cultures. Objects spread via globalization (e.g., plastic bags) and objects with a basic functional role that has arisen similarly across many groups (e.g., spoon-like things for eating) are likely to appear widely, while other objects re-main specific to people and places (e.g., the gourd and bombilla for drinking mate in much of South America, stemming from Indigenous Guaraní and Tupí tradition). Take, for example, middle-class US family homes, which have been noted for their large quantities of possessions ("clutter"), much of which is designed specifically for children [e.g., toys and books; Arnold, Graesch, Ochs, & Ragazzini (2012)]. We might infer, based on this description alone, that much of what children do and talk about at home is tailored to what particularly interests them. Thus, children's worlds, in this sense, look very different from their caregivers'. Recent work by Herzberg and colleagues (2021) underscores this point with data from infants (13–23 months old) who spent nearly 70% of their time in object play with toys or a mix of toys and non-toys, with $\sim$100% of infants playing with children's books and stuffed animals, and a total of 32 toy types appearing in $\geq$25% of infants' play. Non-toy play was also common, but still appeared to predominantly include infant-specific objects (e.g., sippy cups, baby spoons, high chairs, pacifiers). We would expect many of these items to be rare in other parts of the world, with much greater overlap between objects for infants and objects for adults (e.g., Karasik, Schneider, Kuchirko, & Tamis-LeMonda, 2018).

The current study

Using daylong photo streams from child-worn cameras, we analyze object handling by children under age four in two rural, small-scale subsistence farming communities from opposite sides of the globe: Rossel Island ("Rossel"; Milne Bay Province, Papua New Guinea) and Tenejapa ("Tseltal"; Chiapas, Mexico). While these communities are comparable in many ways (e.g., rural, swidden horticulturalist, housed in multi-generation family complexes), prior work has established substantial differences in the organization of young children's daily lives, child carrying practices, and each community's level of market integration (e.g., the greater availability of synthetic materials in Tenejapa), leading us to expect differences in what children handle across the day and early lifespan (Brown & Casillas, 2021; Casillas, Brown, & Levinson, 2020, 2021; Casillas & Elliott, 2021). We first establish how often children handle objects from different categories (e.g., food vs. tools), both by the total amount of handling and by number of unique objects per hour in each category across sites. We explore the top individual objects in each site and how overlap exists between sites. Finally, we investigate how the rate and characteristics of object handling change with developmental age, as predicted in prior work (Casillas & Elliott, 2021).

Our findings reveal relative consistency in the broad composition of objects handled by children, both between sites and across age, with a few important exceptions: a greater diversity of synthetic objects per hour for Tseltal children (e.g., relating to greater market integration), more time spent with immovable objects for Rossel children (e.g., relating to socializing time on/near household verandas), and a greater diversity of held objects with developmental age. We discuss open questions and potential implications of these findings for early word learning.

Method

Corpus

Daylong photo streams consisted of images captured approximately every 15 (Rossel) to 30 (Tseltal) seconds over the course of 8 (Rossel) to 9 (Tseltal) waking hours at home. Children wore a recording vest equipped with a camera (Narrative Clip 1) and miniature fisheye lens (Photojojo Super Fisheye) that provided a 180° view of the environment. For younger infants who were not yet walking, the camera was instead worn by the primary caregiver. Previously, 83 daylong photo streams (113668 photos) had been comprehensively manually annotated for the presence or absence of child object handling (Casillas & Elliott, 2021). Here, we further annotate and analyze the subset of 16368 with object handling in the present study.

We included one daylong photo stream from each of 74 children (Rossel: 39, Tseltal: 35), ranging in age from 0 to 48 months (M_{Rossel} = 22.2, $M_{Tseltal}$ = 23.3). The amount of object handling and thus the number of photos annotated varied across children, ranging from 1 to 626 (M_{Rossel} = 223.5, $M_{Tseltal}$ = 188.2).

Figure 1: Example images with object and category labels.

Manual annotation

Photos were annotated with IMCO, an open-source Python program adapted for efficient coding of photo streams (Casey, Fisher, Tice, & Casillas, 2022). Annotators provided labels for the handled object(s) in each photo (e.g., "twig") and selected among predefined categories to characterize each type of object (e.g., "Natural"). Categories included food, mealtime tools ("Tool-M"), toys, clothing, tools for working or cleaning ("Tool-W"), immovable objects (e.g., furniture and housing structures), natural objects, and miscellaneous synthetic objects (see Figure 1 for example images and Table 1 for example objects from each category).

Data preparation and reliability

Images were excluded if they were too dark, bright, blurry, or covered for annotators to identify handled objects (738 images, 4.51% of the data set), if annotators were otherwise unsure about what objects were being handled (133, or 0.81%), if there was no handled object (210, or 1.28%), or if the researcher was still present when the image was captured (3, or 0.02%). To avoid unnecessary data loss, all excluded photos were checked by at least one other annotator and re-included for analysis if objects were identifiable. In total, 15305 images were deemed usable by annotators (8717 for Rossel, 6588 for Tseltal).

XX% of photo streams were double coded. Reliability annotations were equally spread across sites and ages and included a total of XXXX images. At the category level, annotators agreed on XX.X% of decisions (Rossel: XX.X%, Tseltal: XX.X%). At the object label level, annotators agreed on XX.X% of decisions (Rossel: XX.X%, Tseltal: XX.X%).

		Rossel		Tseltal	
	Object Category	**N**	**Top objects**	**N**	**Top objects**
1	Food	37	betelnut, coconut, tuber	54	bean, tortilla, chips
2	Synthetic	75	blanket, plastic bottle, woven basket	77	plastic bag, blanket, bucket
3	Toy	20	ball, book, swing	44	toy car, ball, book
4	Natural	21	stick, leaf, rock	13	stick, plant, tree
5	Clothing	16	shirt, purse, skirt	22	shirt, pants, shoe
6	Mealtime Tool	20	bowl, spoon, knife	12	bowl, cup, mug
7	Immovable	19	stairs, wall, floor	19	chair, door, fence
8	Work Tool	14	knife, broom, scrub brush	27	embroidery ring, knife, broom

Table 1: Unique object counts and top non-study-related objects for each category across sites.

Results

Overall frequency statistics

Children handled an average of 26.84 unique objects per day (median = 27.5, *SD* = 15.73, range = 1–59), with no significant differences across sites (M_{Rossel} = 26.38, $M_{Tseltal}$ = 27.34, W = 666, p = 0.862). The distribution of handled objects was highly right-skewed within and across children. Each child's distribution was skewed such that a small group of objects was handled in a majority of their images but most objects were handled for only short periods of time (Figure 2). Across children, common objects followed a similar Zipfian distribution: some objects were handled by many children, but most objects were only handled by 1-2 children in each site (Rossel: 57.21%, Tseltal: 61.57%).

Comparing across sites, 32.7% of objects were present in both communities, and several shared objects were among the most frequently handled by children in both sites. In fact, among the top 25 most common objects, 10 were shared across sites (Figure 3). Of note, the study camera was the object that was handled by the most children in both sites (Rossel: 69.2%, Tseltal: 91.4% of children) but accounted for a relatively small percentage of each child's object handling time, on average (M_{Rossel} = 3.8%, $M_{Tseltal}$ = 6.5% of images). The camera and other study-related objects (i.e., vest and privacy cover for the camera), were retained in our analyses; however, inclusion of these items did not qualitatively change any of the reported results.

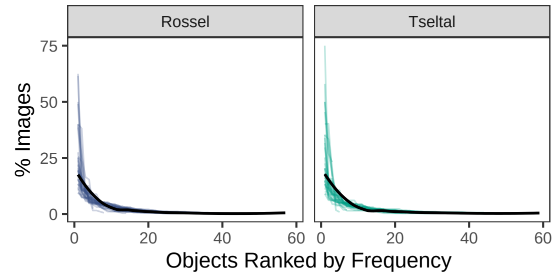

Figure 2: Zipfian distribution of objects handled by each child across sites. For each child, the top object was defined as the object appearing in the greatest number of images; thus, the identity of the top object does not match across all children.

Effects of object category

The frequency of object categories was similarly divided across sites (Figure 4A). Children primarily handled food items (M_{Rossel} = 27.31%, $M_{Tseltal}$ = 31.52% of handling) and miscellaneous synthetic objects (e.g., rope, woven basket, bottle, etc., M_{Rossel} = 24.3%, $M_{Tseltal}$ = 31.37% of handling). For 51 of 75 children, their top category was either food or synthetic objects. Two-tailed Wilcoxon tests revealed only one significant category-level difference between sites: children's handling of large or immovable objects, where Rossel children handled these objects more frequently than Tseltal children (M_{Rossel} = 10.79%, $M_{Tseltal}$ = 3.55%, adjusted p = 0.001, but these objects were still among the least frequently handled in both sites. The initial predicted difference in synthetic object handling between sites was significant; however, after correcting for multiple comparisons, this effect did not persist (adjusted p = 0.104).

During any given hour, children handled 6.5 objects from 3.66 different categories, on average (median = 6 objects, *SD* = 4.62, range = 1–27). To test for differences across sites and categories, we ran individual linear mixed-effects models for each of the six object categories, with category membership dummy coded (i.e., objects belonging to the target category for a given model = 1, objects belonging to other categories = 0). Models included fixed effects of site, category, number of images, and site-by-category interaction as well as random intercepts for individual children. After correcting for multiple comparisons, we found a significant main effect of the synthetic object category (β = 0.39, *SE* = 0.09, t = 4.42, $p < 0.001$) and a significant site-by-synthetic interaction (β = 0.37, *SE* = 0.12, t = 2.98, p = 0.056) such that children handled more unique synthetic objects per hour than objects from other categories, and this effect was stronger for Tseltal children than for Rossel children. Additionally, we found negative main effects for the toy (β = -0.45, *SE* = 0.14, t = -3.34, p = 0.017) and work tool (β = -0.73, *SE* = 0.19, t = -3.92, p = 0.002) categories, indicating that children han-

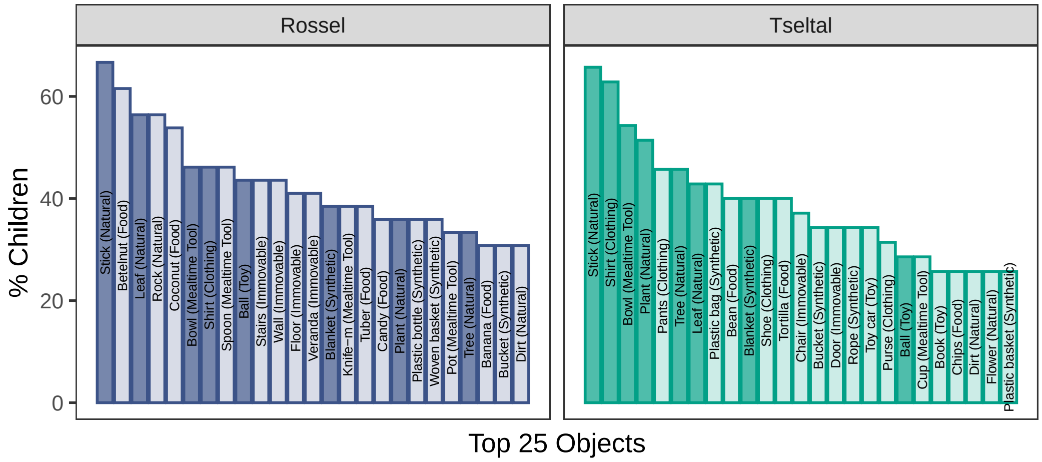

Figure 3: Non-study-related objects handled at least once by the most children in each site. Filled bars represent objects that were among the top 25 for both sites.

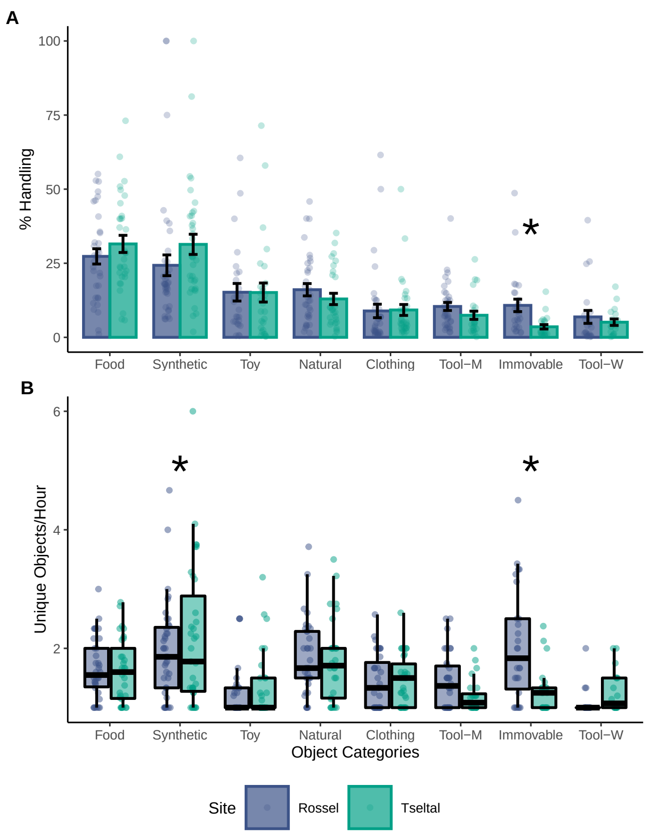

Figure 4: (A) Overall frequency of handling by object category. Points reflect percentages for individual children. (B) Count of unique objects handled per hour by object category. Points reflect means for individual children across all hours of recording. Asterisks indicate significant differences between sites after correcting for multiple comparisons.

dled fewer unique objects from these categories per hour. Finally, a significant main effect of the immovable object category (β = 0.46, *SE* = 0.11, t = 4.01, p = 0.001) and site-by-immovable interaction (β = -0.88, *SE* = 0.18, t = -4.96, $p <$ 0.001) revealed that children handled more unique immovable objects per hour than objects from other categories, and this effect was stronger for Rossel children than for Tseltal children (Figure 4B).

Effects of age

The frequency of object categories was largely stable over age (Figure 5). The only exception was a decreasing proportion of synthetic object handling across developmental time for children in both sites (β = -0.01, 0, *SE* = 0, 0, t = -2.28, -2.52, p = 0.026, 0.014). [explain that this is driven by the fact that synthetic objects are the most common and younger kids handle fewer objects]. Thus, the broad composition of handled objects did not change significantly as a function of age.

However, children's overall *rate* of object handling increased significantly with age (Figure 6A). That is, older children handled more unique objects per hour (β = 0.11, *SE* = 0.03, t = 3.18, p = 0.002). Additionally, with increasing age, children handled more objects from different categories per hour (Figure 6B; β = 0.05, *SE* = 0.01, t = 4.14, p = 0). These effects were consistent across sites; we found no main effects of site or interactions between site and age (all ps $>$ 0.05).

To further explore developmental changes in the characteristics of children's object handling, we measured transitions between different objects and different object categories. We modeled the relative number of transitions per hour (i.e., number of transitions divided by the number of possible objects or categories for the hour) as a function of age and site, plus their interaction. We found no overall age-related increase in object transitions (β = 0.01, *SE* = 0, t = 1.43, p = 0.157). However, there was a significant main effect of site (β = -0.44, *SE* = 0.14, t = -3.26, p = 0.002) as well as a site-by-age interaction (β = 0.01, *SE* = 0, t = 2.56, p = 0.013) such that Tseltal children made fewer transitions between objects

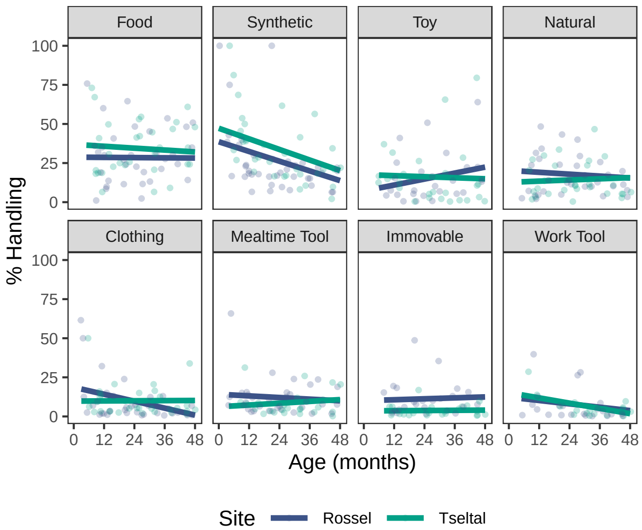

Figure 5: Frequency of handling by object category across age. Individual points show raw percentages per hour for each child, and lines reflect model-predicted percentages.

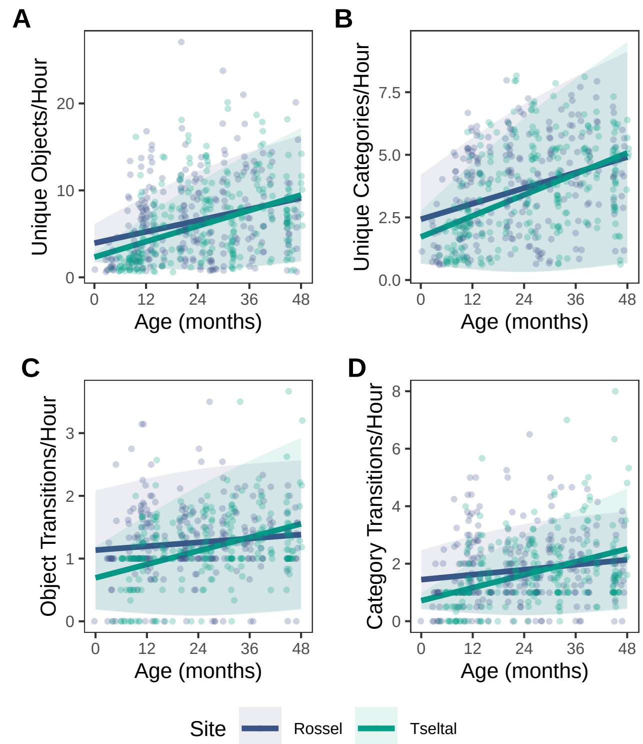

Figure 6: (A) Unique objects and (B) object categories handled per hour as a function of age. (C) Relative number of transitions between objects and (D) object categories per hour as a function of age. Points reflect raw hourly counts for each child, and lines reflect model predictions with shaded standard error regions.

per hour than Rossel children but showed a steeper increase across age (Figure 6C). At the category level, we found that, with increasing age, children made significantly more transitions between object categories per hour ($\beta = 0.01$, $SE = 0.01$, $t = 1.77$, $p = 0.08$), with no detectable differences across sites (Figure 6D).

Discussion

Just some miscellaneous notes for now

Cross-cultural similarities and differences

Summary sentence of the broad overlap between communities on many of the measures reported above, likely owing to the many characteristics that are shared across these two cultural contexts (e.g., rural, swidden horticulturalist, housed in multi-generation family complexes). However, the differences that emerge between sites reveal the influence of market integration, the organization of daily life, and infant carrying practices (Brown & Casillas, 2021; Casillas, Brown, & Levinson, 2020, 2021; Casillas & Elliott, 2021).

Marginally more unique synthetic objects for Tseltal children. Discussion of the effects of market integration on the availability of objects in Tenejapa vs. Rossel Island.

More handling of large or immovable objects for Rossel kids. Discussion of differences in the organization of young children's daily lives. More time spent socializing on/near household verandas

Greater age-related increases in transitions between objects for Tseltal children. Discussion of infant carrying practices. More time in sling for Tseltal babies so fewer opportunities to seek out different objects at the youngest ages.

Other sources of consitency and variability

Possible changes with time of day and activity context as predicted in prior work (e.g., morning mealtimes, Casillas, Brown, & Levinson, 2020, 2021). Open question: are handled objects a good index for activity context?

Changes from day to day: We don't know how much overlap there is from day to day. While the exact objects may change, it's likely that they'd follow a similar right-skewed distribution since we see this for most kids in our sample and across mamy measures of children's input, including visual (Clerkin, Hart, Rehg, Yu, & Smith, 2017; Long, Kachergis, Bhatt, & Frank, 2021) and linguistic (Montag, Jones, & Smith, 2018).

Age: The broad composition of handled objects was largely stable across age (consistent with Long, Kachergis, Bhatt, & Frank, 2021 for visually present categories) but few categories here and also the possibility that individual types may change. More here about manual dexterity and possibly attention affecting object handling bout durations, which we don't quantify here, but the transitions per hour analysis is a first-pass way of (hopefully) getting at something similar

Consequences for learning

Zipfian distributions can be helpful for learning (e.g., Carvalho, Chen, & Yu, 2019). Overlap between kids can get us to start thinking about which object names and object-relevant words and learned earlier. But important to note that there's less overlap between kids (within sites) than we see for US context (e.g., Herzberg, Fletcher, Schatz, & Tamis-LeMonda, 2021), possibly owing to cultural differences and sampling strategies (e.g., two-hour videos vs. daylong photos). Another point is that while there may be less overlap between kids, there's likely more overlap between kids and adults (fewer child-specific items)

Notes for near the end: Future plans to link to daylong audio recordings. We can speculate about children's learning now that we have some understanding of the distribution of object handling and the identities of commonly handled objects. However, without knowing how often children are also hearing *talk* about the objects they're handling, we can't fully address the word learning piece.

References

10 Adolph, K. E., Karasik, L. B., & Tamis-LeMonda, C. S. (2010). Motor skill. In M. H. Bornstein (Ed.), *Handbook of cultural developmental science* (pp. 61–88). Psychology Press: New York, NY.

Arnold, J. E., Graesch, A. P., Ochs, E., & Ragazzini, E. (2012). *Life at home in the twenty-first century: 32 families open their doors*. ISD LLC.

Bergelson, E., Amatuni, A., Dailey, S., Koorathota, S., & Tor, S. (2019). Day by day, hour by hour: Naturalistic language input to infants. *Developmental Science*, *22*(1), e12715.

Bergelson, E., & Aslin, R. N. (2017). Nature and origins of the lexicon in 6-mo-olds. *Proceedings of the National Academy of Sciences*, *114*(49), 12916–12921.

Brown, P., & Casillas, M. (2021). *Childrearing through social interaction on Rossel Island, PNG*. (A. J. Fentiman & M. Goody, Eds.). New York, NY: Berghahn.

Carvalho, P., Chen, C., & Yu, C. (2019). *Rethinking the input: Skewed distributions of exemplars result in broad generalization in category learning*.

Casey, K., Fisher, W., Tice, S. C., & Casillas, M. (2022). ImCo: A python tkinter application for coding lots of images (Version 2.0). Retrieved from `https://github.com/kennedycasey/ImCo2`

Casillas, M., Brown, P., & Levinson, S. C. (2020). Early language experience in a Tseltal Mayan village. *Child Development*, *91*(5), 1819–1835.

Casillas, M., Brown, P., & Levinson, S. C. (2021). Early language experience in a papuan community. *Journal of Child Language*, *48*(4), 792–814.

Casillas, M., & Elliott, M. (2021). Cross-cultural differences in children's object handling at home. PsyArXiv. http://doi.org/10.31234/osf.io/43db8

Clerkin, E. M., Hart, E., Rehg, J. M., Yu, C., & Smith, L. B. (2017). Real-world visual statistics and infants' first-learned object names. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160055.

Fausey, C. M., Jayaraman, S., & Smith, L. B. (2016). From faces to hands: Changing visual input in the first two years. *Cognition*, *152*, 101–107.

Gaskins, S. (2000). Children's daily activities in a Mayan village: A culturally grounded description. *Cross-Cultural Research*, *34*(4), 375–389.

Herzberg, O., Fletcher, K. K., Schatz, J. L., & Tamis-LeMonda, C. S. (2021). Infant exuberant object play at home: Immense amounts of time-distributed, variable practice. *Child Development*, *XX*, 1–15.

Jayaraman, S., Fausey, C. M., & Smith, L. B. (2017). Why are faces denser in the visual experiences of younger than older infants? *Developmental Psychology*, *53*(1), 38.

Karasik, L. B., Schneider, J., Kuchirko, Y. A., & Tamis-LeMonda, C. S. (2018). Not so WEIRD object play in Tajikistan. Presentation to the International Conference on Infant Studies, Philadelphia, PA. http://doi.org/10.31234/osf.io/43db8

Kretch, K. S., Franchak, J. M., & Adolph, K. E. (2014). Crawling and walking infants see the world differently. *Child Development*, *85*(4), 1503–1518.

Laing, C., & Bergelson, E. (2020). From babble to words: Infants' early productions match words and objects in their environment. *Cognitive Psychology*, *122*, 101308.

Long, B., Kachergis, G., Agrawal, K., & Frank, M. C. (2020). *Detecting social information in a dense database of infants' natural visual experience*.

Long, B., Kachergis, G., Bhatt, N., & Frank, M. C. (2021). *Characterizing the object categories two children see and interact within a dense dataset of naturalistic visual experience*.

Montag, J. L., Jones, M. N., & Smith, L. B. (2018). Quantity and diversity: Simulating early word learning environments. *Cognitive Science*, *42*, 375–412.

Sanchez, A., Long, B., Kraus, A. M., & Frank, M. C. (2018). Postural developments modulate children's visual access to social information. In *Proceedings of the 40th annual conference of the cognitive science society* (pp. 2412–2417).

Super, C. M. (1976). Environmental effects on motor development: The case of 'African infant precocity.' *Developmental Medicine & Child Neurology*, *18*(5), 561–567.

Yu, C., & Smith, L. B. (2013). Joint attention without gaze following: Human infants and their parents coordinate visual attention to objects through eye-hand coordination. *PloS One*, *8*(11), e79659.

Yurovsky, D., Smith, L. B., & Yu, C. (2013). Statistical word learning at scale: The baby's view is better. *Developmental Science*, *16*(6), 959–966.