

**Learning language *in vivo***

Marisa Casillas

University of Chicago

**Author Note**

Marisa Casillas  <https://orcid.org/0000-0001-5417-0505>

Correspondence concerning this article should be addressed to Marisa Casillas,  
Department of Comparative Human Development, University of Chicago, Rosenwald Hall  
3rd floor, 1101 E. 58th Street, Chicago, IL 60637. E-mail: [mcasillas@uchicago.edu](mailto:mcasillas@uchicago.edu)

**Abstract**

The language *in vivo* framework pushes for a substantially deeper view of children's linguistic input, with the ultimate aim of building sustainable and tangible links between theoretical models of language development and families' everyday lived experiences. Children's language experiences constrain theoretical models in ways that may illuminate universal learning biases. However, more than that, these experiences provide a staggering array of test cases and demonstrate the stage-setting effects of situational, familial, and societal context on language use and uptake. Centering on activity type as an entrypoint into "context", I motivate and outline an approach that draws on both big- and small-data methods to generate grounded, informative, and representative descriptions of children's input. Minimally, this approach complements recent work that uses more bundled input measures. Maximally, an *in vivo* approach could inspire new questions, insights, and innovations to broaden the coverage and application of theoretical models to individual communities and families.

*Keywords:* activity, language development, input, daylong recordings, LENA, culture

## **Learning language *in vivo***

### **Introduction**

How do children learn language? The last half century has seen immense growth in the data, methods, and theory attempting to answer this question (Slobin, 2014). Developmental psycholinguistics, for its part, has arrived at a common, three-piece framework for conceptualizing the core puzzle: input, processor, and output (e.g., Kachergis et al., 2022). In one piece of this framework, children are exposed to language patterns in their ambient environment. In another piece, this linguistic “input” is processed via a variety of cognitive mechanisms supporting learning. The third and final piece recapitulates input patterns in the child’s own linguistic behaviors, i.e. the “output”.

Undergirding each piece of this framework are ways of thinking that are taken for granted, but actually reflect disciplinary and historical artifacts relevant to psychological inquiry. For example, hypothesis-driven design purposefully constrains the knowledge we create. We strip down and operationalize complex phenomena in order to reliably quantify, justifiably compare, and adequately control their presence in our datasets. These disciplinary traditions reflect both our field’s origins and where it is headed, and have resulted in an impressive slate of findings about how language development proceeds in some contexts. We have forged an epistemic path that, by its very trajectory, implies sensible ways forward.

Among these ways forward are approaches to further illuminate the nature of the linguistic input and processor, including: a continued focus on child-directed language from adults, automated measures of language behavior, and (relatedly) indices of linguistic behavior that scale to whole days or longer. As I will argue, these approaches have been tremendously fruitful in recent years, but present a highly limited view of the learner and therefore limited depth of the theories that arise from them. To keep a grip on the individual phenomena that add up to the wonder of childhood language learning, we require a complementary approach. In this paper I gaze away from the beaten path, toward other

possible approaches, to identify shortcomings among our otherwise-sensible ways forward.

I propose an approach that explicitly engages with context rather than attempting to average it away. I argue that analytical approaches to understanding linguistic input and knowledge need to center around learners' interactional and cultural contexts in order to make meaningful connections between mainstream theoretical frameworks and families' everyday lived experiences. In what follows I review how we have typically measured input, then I sketch an alternative, language *in vivo* approach, closing with what we stand to gain.

### Past and current approaches to linguistic input

I first address three facets of the current approach to analyzing input—all three need further consideration under a language *in vivo* approach: (1) which *sources* of input we include, (2) what is captured in recordings, and (3) how we *measure* input.

#### *Sources of input*

There is general consensus that child-directed language<sup>1</sup> from adults (CDL<sub>A</sub>) is the linguistic input *par excellence*. CDL<sub>A</sub> is associated with earlier and cumulatively larger gains in vocabulary and speed of lexical retrieval, and has implications for aspects of syntactic development (e.g., Bates & Goodman, 1999; Frank et al., 2021). That said, other sources of input (e.g., observable talk between others) likely make critical contributions beyond vocabulary (e.g., Benigno et al., 2007; Foushee et al., 2021; Oshima-Takane, 1988).

CDL itself varies immensely between communities in form and function. As linguistic anthropologists have warned: in many communities, language directed to young children is infrequent, multi-party, or more focused on social practices than referential communication. The mere presence of CDL<sub>A</sub> in and of itself also doesn't guarantee the scaffolding features presumed to make it helpful in the first place (e.g., McClay et al.,

---

<sup>1</sup> Unless referring specifically to speech, I use the terms child- and adult-directed 'language' (CDL/ADL) to refer to spontaneous language use.

2022). For example, the pitch variation prototypically associated with CDL<sub>A</sub> is especially salient in North American English, compared both to other English varieties and to other languages (including Lebanese Arabic, K'iche, and Tseltal; Farran et al., 2016; Pye, 1986; Soderstrom et al., 2021)). Even within North American English contexts, CDL use is moderated by the age, gender, and affect of its producer (e.g., Bergelson, Casillas, et al., 2019; Kitamura et al., 2001).

Singular emphasis on CDL<sub>A</sub> misses a critical insight: *how* others talk to children provides essential information about their expected roles in society, including how they are expected to use language. For example, EuroAmerican caregivers tend to linguistically interpret and respond to their infants' expressions of negative and positive inner experience. Such behavior positions infants as individuals with mature psychological states and creates a foundation for children's continued outward sharing of their inner experiences (Gaskins, 2006). These caregivers use CDL<sub>A</sub> to socially center children and to pull their inner experiences outward. In contrast, Yucatec Mayan children are encouraged to be more conservative in verbally expressing inner experiences. A lack of CDL<sub>A</sub> in response to infant discomfort is an important socialization tool; caregivers rely on bodily contact to swiftly, effectively, and (often) nonverbally attend to infants' negative inner experiences (Gaskins, 2006). In order to chart the contours of the learner's linguistic environment, we must go beyond a focus on CDL<sub>A</sub> and engage with a variety of different input sources (e.g., different people who interact with the child), diverse interaction types (e.g., dealing with distress, social routines), and multiple modalities (see, e.g., Abu-Zhaya et al., 2017; Capirci et al., 2022; West et al., 2022).

### ***Capturing language in recordings***

Methods for recording child language environments have seen significant change. The foundation of our field rests on careful analysis of naturalistic interactions, recorded for relatively short periods at the child's home (or in a laboratory playroom) and then

painstakingly annotated before analysis. Then, the LENA system (Greenwood et al., 2011) launched our field into the domain of big data research with its 16-hour audio-recording capacity and software for automated acoustic analysis.<sup>2</sup> While LENA has been an enormous boon to the field, the associated cost of going ‘big’ is a loss of depth. The critical issues are that (a) the automated system produces highly limited information about the language used<sup>3</sup> and (b) the data are audio only (e.g., no sign, gaze, or gestural data). This trade-off between scale and depth should make one wary of how feasibility and convenience vs. theoretical ideals weigh into study design.

Some heroic efforts to annotate daylong data in more detail have demonstrated both the richness of what could be studied were it manually annotated (e.g., Bergelson, Amatuni, et al., 2019; Montag, 2020) and the potential downfalls of relying exclusively on automated output (e.g., Ferjan Ramírez et al., 2021). Encouragingly, protocols and guides for manual annotation of daylong data have started to emerge, some with an eye toward maximizing efficiency (Casillas et al., 2017; Cychosz et al., 2021; Mendoza & Fausey, 2021). Researchers going down this path can expect significant investments of time and money (Figure 2). The payoff, however, is worthwhile: languages, embedded in their social and material contexts, offer an inspiring array of interrelated phenomena whose learning processes must be explained. We need to inspect the input (and output) closely to capture most of them.

### ***Measures of input***

Transcribed data are unbeatable in demonstrating how children actually encounter linguistic units (e.g., words/signs, gestures, syntactic structures, communicative acts, etc.),

---

<sup>2</sup> Large cross-corpus analyses of “small data” began long before then, largely thanks to CHILDES (MacWhinney, 2000).

<sup>3</sup> Annotations include: when basic speaker types (woman, man, key child, other child) talk, when the key child and a nearby adult talk one after the other, and high-level speech content estimates (e.g., number of nearby adult words). Open-source alternatives somewhat improve this outlook (see Räsänen et al., 2021).

making resources such as CHILDES indispensable (MacWhinney, 2000, an open online database of naturalistic child language data). However, rather than looking at individual phenomena (e.g., frequency of passive syntactic constructions, phonetic properties of spontaneously produced vowels, etc.), many researchers choose to examine input as an aggregated bundle of such features, as in “CDL” (e.g., Bergelson, Amatuni, et al., 2019; Bunce et al., 2022; Casillas et al., 2020, 2021; Cristia et al., 2019; Greenwood et al., 2011; Sperry et al., 2019).

While useful for general models of language development, such high-level measures guarantee little in explaining how children acquire the specific phenomena relevant for their language communities (e.g., how to inflect a verb in minimally vs. highly verb-inflecting languages, like English/Mandarin vs. Inuktitut/Murrinhpatha). While we *can* characterize child language development as a general process, theory must also cover the specific and interacting subprocesses encompassed by this general process. Precisely *because* these subprocesses may be realized differently across diverse contexts, they provide useful theoretical levers into the nature of the learner’s cognitive toolkit. Diversifying the way we measure the input is essential to understanding how these general and specific processes of learning interrelate.

Research using aggregate input rates has, instead, recently focused on a different issue: how to examine ‘quality’ over ‘quantity’ input (e.g., Masek et al., 2021). The apparent tension between ‘quality’ and ‘quantity’, however, may ultimately be a red herring. Some ‘quantity’ measures (e.g., CDL<sub>A</sub>) capture multiple input characteristics that are canonically considered to be of high ‘quality’. Further, many of these same characteristics, e.g., contingent turn-taking with the key child, are likely culturally specific in what they do for the child’s budding competence as a member of society. Returning to the US vs. Mayan contrast above, different provision of caregiver response and encouragement of child contributions predicts that turn-taking behaviors invoke different linguistic and social productions in the two groups—thus, the underlying mechanisms

relating turn-taking behavior in the input to the child's learned behavior in the output are likely to differ across populations.

The biggest issue at stake then seems to be: how can we focus our research on meaningful linguistic practices in the child's environment while *also* maximizing the naturalness, reliability, and generalizability of our findings? Not taking this question seriously can lead to limited or even misleading findings, as I will demonstrate.

### ***Universals in the macro and potential dead ends***

The wide application of aggregate- or indicator-style language measures can reveal crosslinguistic and crosscultural patterns in development (i.e., potential universals). For example, similar factors predict parent-reported early vocabulary across a diverse group of languages. This suggests that there are a handful of biases that similarly impact learning across these groups (Frank et al., 2021). In general, however, when aggregate measures are exported from their original context (e.g., from urban industrial societies to rural subsistence ones), there is significant risk of discovering consistencies that overestimate similarity and underestimate diversity.

Recent work using manual measures of CDL input rates across diverse cultural contexts stumbled upon one such possible universal—children under 3;0 encounter infrequent CDL, an average of ~2.5–5 min/hour, across their daytime hours (Bunce et al., 2022; Casillas et al., 2020, 2021). This macro-level consistency obscures tremendous crosscultural variation in how caregivers deploy CDL. For example, ethnographic work characterizes caregiver-child interaction as child-centric on Rossel Island (Papua New Guinea) and non-child-centric in a Tseltal Mayan community (Mexico; Brown, 2014; see also Shneidman and Goldin-Meadow, 2012). Yet these communities appear nearly identical in CDL baseline input rates (Casillas et al., 2020, 2021). Where have the differences gone? The differences, I argue, will be primarily visible at the micro scale of moment-to-moment interaction. Across individual interactions, we should see the differences via the people

children engage with and the things they do with language. At the macro scale, however, averaging a single measure over a wide variety of contexts and scenarios, we cannot see these differences.

Both macro and micro scales provide informative angles on language development. The macro scale tends to highlight the processes that relate to robust, high-level phenomena occurring across groups. The micro scale is equally important in its ability to reveal how the content and style of the input leads children to resemble other language users in their community. Without centering both perspectives, models of child language will not reach their full theoretical potential. In this vein I propose the language *in vivo* approach to studying children's linguistic experiences.

### **The language *in vivo* approach**

The core intuition behind the *in vivo* approach is that we cannot extract language acquisition from the social context of everyday life (see Adolph & Sternberg, 2019; de Barbaro, 2019, for related calls). The way in which children learn to attend to and produce language is entwined with the things they do with and around others. Even from the bottom-up perspective of internalizing the statistical properties of the surrounding language, linguistic representations and processes are not fenced off from social experiences (e.g., when to use “blankie” vs. “blanket”). Children do not just learn language, but learn it to communicate. What they talk about and to whom is shaped by local social and cultural context (e.g., Bruner, 1985; Ochs & Schieffelin, 1984). The theoretical impact of these ideas is that the language learning system must be adaptive and highly flexible. While there may be core mechanisms that work similarly across most children (e.g., some types of statistical learning), others likely vary depending on context (e.g., propensity to learn from others' interactions; Rogoff et al., 2003) or may be specialized to certain domains (e.g., words vs. facts vs. phonology; Foushee et al., 2021; Knightly et al., 2003).

How can we take a truly socially centered approach to studying early language

experience? Language socialization research achieves this aim via anthropological engagement with the community under study (Ochs & Schieffelin, 1984). However, such an approach is infeasible for most developmental psychologists, so it's unclear how they should proceed. I propose we start by scaling down the child's social language environment to a more bite-sized phenomenon that we can reliably identify and that has meaningful similarities and differences across diverse homes: activity context. From the starting point of activity, we can unfold and integrate many of the concepts, methods, and analytical approaches native to developmental psychology with more contextualized considerations of the child's developmental milieu.

## **Activity context**

Activity as an entryway into language study—indeed as an entryway to language *development* by the child—has deep theoretical roots in our field (e.g., Bruner, 1985). We know that activity affects the language children encounter and use with others (e.g., Glas et al., 2018; Tamis-LeMonda et al., 2019). The language *in vivo* approach builds on these ideas with one main difference: it demands explicit exploration of how the studied activities are meaningful to participants in their situated context and the broader societal, research, and applied contexts used to frame the study. By holding ourselves accountable for the *local* relevance of studied activities, researchers can greatly enrich their work, grounding broad theories in children's moment-to-moment language experiences. That is, their language experience *in vivo* (see Figure 1). Two challenges in taking this approach are how to identify activities and what to do with activity information.

### ***Identifying activities***

I focus on long-form naturalistic data (e.g., LENA recordings) because: (1) multiple activity contexts, often overlapping, are captured from the child's perspective as they go throughout their day and (2) there are no set conventions for finding and choosing activity

types (see Mendoza & Fausey, 2021, for more)<sup>4</sup>, leaving room to test out a preliminary language *in vivo* approach.

Researchers using an *in vivo* approach should elaborate the purpose and relevance of the studied activities according to participants. Therefore, even just identifying activities to study may require significant investment, e.g., manual annotation of existing data and engagement with research outside one's typical domain. Commonly studied activities like 'free play' and 'book reading' are not exempted. They meaningfully differ in their relevance and familiarity across communities, and researchers should explicitly address how these activities are understood by participants, along with implications for data interpretation. The proposal is not to wholesale manually annotate recordings prior to data analysis. Beginning with hypotheses or selective attention to certain activity types is encouraged—just together with clear data- or community-internal evidence of activity relevance along with explicit recognition of the limits of generalization.

I give two concrete examples of how to identify activities in line with an *in vivo* approach. The first is to begin with a theme that implies relevance to participants but requires local specification. For example, asking about how language is (and isn't) used by caregivers in dealing with infants' bodily essentials: eating, sleeping, crying, bathing, and dealing with urine/feces. A combination of observational experience and familiarity with the participant community is critical to identifying the specific activities relevant to this theme. A second way to begin is to set out particular goals or roles of interest, e.g., typical daily journeys, spontaneous informal pedagogy, encouraging socially appropriate behavior, etc.. Again, the specific relevant activities will differ according to context. Counting activities in the US look rather different from recitation of kin ancestry on Rossel Island in Papua New Guinea, but comparison of the two provides a more truthful examination of informal pedagogy in each context than would exclusively examining one of those activities

---

<sup>4</sup> Current automated tools can indicate two activities: nearby electronic noise (e.g., TVs) and infant sleep (Bang et al., 2021; Greenwood et al., 2011).

across both (e.g., asking US parents to engage their children in kin ancestry pedagogy; Figure 1).

### ***Using activity information***

Once collected, what do we do with these recordings? If the goal is to arrive at contextually validated findings that are sensitive to both local frames and broader discourses, there is really just one answer: detailed inspection and annotation of the data. Such work is costly (Figure 2), but the returns are invaluable in highlighting critical gaps in current knowledge and potential ways forward for bridging them. Such data can also typically be used again in future work, especially if they are made available to others (e.g., MacWhinney, 2000).

While this work is descriptive, it contributes substantively to theory. I mention here three examples of theoretical work that would be done best on the basis of an activity-centered approach: (a) how lexical development is guided by everyday activities at both the domain level (e.g., automobiles, foods) and at the feature level (e.g., deixis, animacy, register); (b) the conditions under which spontaneous informal pedagogy occurs and what material and interactional frames it involves; and (c) how societal expectations about children as language users shape what children attend to and (re)produce during interaction.

For each of these topics, there would still be significant need for further examination via experimental and computational work, elaborating on the proposed processes. The extended *in vivo* challenge will be to find innovative ways to keep more controlled investigations like experiments in touch with the features of interaction that made the descriptive work meaningful in the first place.

### **What we stand to gain**

The purpose of the language *in vivo* approach is to celebrate continuing advances in our ‘big data’ ability to document developmental environments without losing sight of

what children and their interactants actually *do* with language. I have envisioned this work from the perspective of (**positionality statement redacted for review**). But the *in vivo* framework, firmly grounded in families' everyday social desires and obligations, is also useful from other perspectives, including those that aim to document or support language communities undergoing change (de León, 2019) or those explicitly intervening on parent or child behavior (see Rowe & Weisleder, 2020, for a review).

The *in vivo* approach can help develop new perspectives on classic questions and delve deeper into less-researched questions. For example, current research aims to explain the processes by which vocabulary arises in the way that it does by proposing mechanisms ranging from statistical tendencies to analogy-based reasoning. An *in vivo* approach would follow this path, but emphasize the role of cultural, situational, and historical context with the hope of better understanding how this early word knowledge informs children's development as competent language users. Even when a word is learned early across diverse groups, its representation will be built up from very different experiences (e.g., "aunt" in an isolated urban context vs. a multi-generational family settlement).

Regarding less-studied research questions, a prime example comes from peer- and other-directed linguistic input. Whether children are repeating a secret, a swear word, or a character name from a movie they have never seen, caregivers are frequently reminded that their children's input goes beyond what they contribute. Establishing how children attend to and reproduce others' linguistic practices is tightly linked to what children see as relevant—an issue addressable with an *in vivo* approach.

## Conclusions

While it is tempting to scientifically isolate language as a symbolic system (viz. grammar and lexicon) from social systems, we would limit our ability to draw direct connections between theories and families' everyday lived experiences. Theory that is constructed without context (or in ignorance of its implicit context) will not touch ground

naturally or equally when applied in diverse domains. To ordinary folks, language is how one trades in social, cultural, and world knowledge. People primarily use language to do things with others, even the smallest language users. Not just leveraging but *embedding* social context at the heart of our work would pay off immensely in these respects.

### Acknowledgements

Redacted for review.

## References

- Abu-Zhaya, R., Seidl, A., & Cristia, A. (2017). Multimodal infant-directed communication: How caregivers combine tactile and linguistic cues. *Journal of Child Language*, 44(5), 1088–1116.
- Adolph, K. E., & Sternberg, R. (2019). Ecological validity. In R. J. Sternberg (Ed.), *My biggest research mistake: Adventures and misadventures in psychological research* (pp. 187–190). SAGE Publications.
- Bang, J., Kachergis, G., Weisleder, A., & Marchman, V. (2021). An automated classifier for child-directed speech from LENA recordings. *Presented at the 46th Boston University Conference on Language Development*.
- Bates, E., & Goodman, J. C. (1999). On the emergence of grammar from the lexicon. *The emergence of language* (pp. 29–80). Mahwah, NJ: Lawrence Erlbaum.
- Benigno, J. P., Clark, L., & Farrar, M. J. (2007). Three is not always a crowd: Contexts of joint attention and language. *Journal of Child Language*, 34(1), 175–187.
- 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., Casillas, M., Soderstrom, M., Seidl, A., Warlaumont, A. S., & Amatuni, A. (2019). What do North American babies hear? A large-scale cross-corpus analysis. *Developmental Science*, 22(1), e12724.
- 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.
- Bruner, J. (1985). Child's talk: Learning to use language. *Child Language Teaching and Therapy*, 1(1), 111–114.
- Bunce, J. P., Soderstrom, M., Bergelson, E., Rosemberg, C. R., Stein, A., Alam, F., Migdalek, M., & Casillas, M. (2022). A cross-cultural examination of young

- children's everyday language experiences. *PsyArXiv*.  
<https://doi.org/10.31234/osf.io/723pr>
- Capirci, O., Caselli, M. C., & Volterra, V. (2022). Interaction among modalities and within development. In A. Morgenstern & S. Goldin-Meadow (Eds.), *Gesture in language: Development across the lifespan* (pp. 113–133). De Gruyter Mouton; American Psychological Association. <https://doi.org/https://doi.org/10.1037/0000269-005>
- Casillas, M., Bergelson, E., Warlaumont, A. S., Cristia, A., Soderstrom, M., VanDam, M., & Sloetjes, H. (2017). A new workflow for semi-automatized annotations: Tests with long-form naturalistic recordings of children's language environments. *Proceedings of Interspeech 2017*, 2098–2102.
- 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.
- Cristia, A., Dupoux, E., Gurven, M., & Stieglitz, J. (2019). Child-directed speech is infrequent in a forager-farmer population: A time allocation study. *Child Development*, 90(3), 759–773.
- Cychosz, M., Villanueva, A., & Weisleder, A. (2021). Efficient estimation of children's language exposure in two bilingual communities. *Journal of Speech, Language, and Hearing Research*, 64(10), 3843–3866.
- de Barbaro, K. (2019). Automated sensing of daily activity: A new lens into development. *Developmental Psychobiology*, 61(3), 444–464.
- de León, L. (2019). Playing at being bilingual: Bilingual performances, stance, and language scaling in Mayan Tzotzil siblings' play. *Journal of Pragmatics*, 144, 92–108.
- Farran, L. K., Lee, C.-C., Yoo, H., & Oller, D. K. (2016). Cross-cultural register differences in infant-directed speech: An initial study. *PloS One*, 11(3), e0151518.

- Ferjan Ramírez, N., Hippe, D. S., & Kuhl, P. K. (2021). Comparing automatic and manual measures of parent–infant conversational turns: A word of caution. *Child Development*, 92(2), 672–681.
- Foushee, R., Srinivasan, M., & Xu, F. (2021). Self-directed learning by preschoolers in a naturalistic overhearing context. *Cognition*, 206, 104415.
- Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2021). *Variability and consistency in early language learning: The Wordbank project*. MIT Press.
- Gaskins, S. (2006). Cultural perspectives on infant-caregiver interaction. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 279–298). Oxford:Berg.
- Glas, L., Rossi, C., Hamdi-Sultan, R., Batailler, C., & Bellemouche, H. (2018). Activity types and child-directed speech: A comparison between French, Tunisian Arabic and English. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 63(4), 633–666.
- Greenwood, C. R., Thiemann-Bourque, K., Walker, D., Buzhardt, J., & Gilkerson, J. (2011). Assessing children’s home language environments using automatic speech recognition technology. *Communication Disorders Quarterly*, 32(2), 83–92.
- Kachergis, G., Marchman, V. A., & Frank, M. C. (2022). Toward a “standard model” of early language learning. *Current Directions in Psychological Science*, 31(1), 20–27.
- Kitamura, C., Thanavishuth, C., Burnham, D., & Luksaneeyanawin, S. (2001). Universality and specificity in infant-directed speech: Pitch modifications as a function of infant age and sex in a tonal and non-tonal language. *Infant Behavior and Development*, 24(4), 372–392.
- Knightly, L. M., Jun, S.-A., Oh, J. S., & Au, T. K.-f. (2003). Production benefits of childhood overhearing. *The Journal of the Acoustical Society of America*, 114(1), 465–474.

- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk. Third Edition.* Mahwah, NJ: Lawrence Erlbaum Associates.
- Masek, L. R., Ramirez, A. G., McMillan, B. T., Hirsh-Pasek, K., & Golinkoff, R. M. (2021). Beyond counting words: A paradigm shift for the study of language acquisition. *Child Development Perspectives*, 15(4), 274–280.
- McClay, E. K., Cebioglu, S., Broesch, T., & Yeung, H. H. (2022). Rethinking the phonetics of baby-talk: Differences across Canada and Vanuatu in the articulation of mothers' speech to infants. *Developmental Science*, 25(2), e13180.
- Mendoza, J. K., & Fausey, C. M. (2021). Quantifying everyday ecologies: Principles for manual annotation of many hours of infants' lives. *Frontiers in Psychology*, 12.
- Montag, J. L. (2020). New insights from daylong audio transcripts of children's language environments. *Proceedings of the 42nd Annual Virtual Meeting of the Cognitive Science Society (CogSci 2020)*, 3005–3011.
- Ochs, E., & Schieffelin, B. (1984). Language acquisition and socialization: Three developmental stories and their implications. In R. A. Shweder & R. A. LeVine (Eds.), *Culture theory: Essays on mind, self, and emotion* (pp. 276–320). Cambridge University Press New York.
- Oshima-Takane, Y. (1988). Children learn from speech not addressed to them: The case of personal pronouns. *Journal of Child language*, 15(1), 95–108.
- Pye, C. (1986). Quiché Mayan speech to children. *Journal of Child Language*, 13(1), 85–100. <https://doi.org/10.1017/S0305000900000313>
- Räsänen, O., Seshadri, S., Lavechin, M., Cristia, A., & Casillas, M. (2021). ALICE: An open-source tool for automatic measurement of phoneme, syllable, and word counts from child-centered daylong recordings. *Behavior Research Methods*, 53(2), 818–835.
- Rogoff, B., Paradise, R., Arauz, R. M., Correa-Chávez, M., & Angelillo, C. (2003). Firsthand learning through intent participation. *Annual Review of Psychology*, 54(1), 175–203.

- Rowe, M. L., & Weisleder, A. (2020). Language development in context. *Annual Review of Developmental Psychology*, 2, 201–223.
- Shneidman, L. A., & Goldin-Meadow, S. (2012). Language input and acquisition in a mayan village: How important is directed speech? *Developmental Science*, 15(5), 659–673.
- Slobin, D. I. (2014). Before the beginning: The development of tools of the trade. *Journal of Child Language*, 41(S1), 1–17.
- Soderstrom, M., Casillas, M., Gornik, M., Bouchard, A., MacEwan, S., Shokrkon, A., & Bunce, J. (2021). English-speaking adults' labeling of child-and adult-directed speech across languages and its relationship to perception of affect. *Frontiers in Psychology*, 12.
- Sperry, D. E., Sperry, L. L., & Miller, P. J. (2019). Reexamining the verbal environments of children from different socioeconomic backgrounds. *Child Development*, 90(4), 1303–1318.
- Tamis-LeMonda, C. S., Custode, S., Kuchirko, Y., Escobar, K., & Lo, T. (2019). Routine language: Speech directed to infants during home activities. *Child Development*, 90(6), 2135–2152.
- West, K. L., Fletcher, K. K., Adolph, K. E., & Tamis-LeMonda, C. S. (2022). Mothers talk about infants' actions: How verbs correspond to infants' real-time behavior. *Developmental Psychology*, 58(3), 405.

**Figure 1**

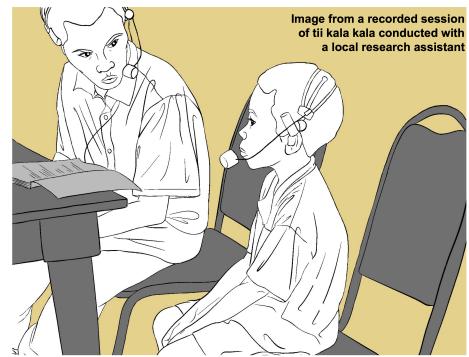
*Example set of researcher considerations when contextualizing candidate activities. We here compare informal home pedagogy activities in the US and on Rossel Island (Papua New Guinea) across a number of relevant and interrelated contextual dimensions.*

## How is this activity meaningful to participants?

**Example:**  
Considerations for  
“informal home pedagogy”  
activity types in the US vs.  
Rossel Island, Papua New  
Guinea



**Counting activities in the US**



**Tii Kala Kala on Rossel Island, PNG**

<b>General background</b>	<i>Broad domain/themes</i>	Numeracy; education	Kinship; ancestry
	<i>Key concept</i>	“Home is the first classroom”	Your kin shape your day-to-day and your life course
	<i>Relevant ages</i>	Early toddlerhood through primary school	Early toddlerhood through adolescence
	<i>Frequency</i>	Routine and sometimes everyday	Routine but not everyday
	<i>Modality notes</i>	Linguistic, gestural, orthographic; related to societal emphasis on symbolic/written systems	Linguistic only; related to an oral tradition of knowledge transmission—written language was only created recently
<b>Purpose and motive</b>	<i>Why it's worthwhile</i>	Academic/educational/occupational potential; serious consequences for everyday life if basic knowledge is not mastered	Intellectual and social virtue; no serious consequences if not mastered because of kin network support when this knowledge is relevant
	<i>Ideal outcomes</i>	Accurate imitation, then abstraction	Accurate imitation, then abstraction; fluent performance
<b>Implementation</b>	<i>Activity-specific features</i>	Many available learning aids; conventional basic knowledge/skills (counting and symbol mastery, simple arithmetic); Assessments where errors are explicitly corrected	No materials—purely verbal practice; conventional basic knowledge/skills (names, ordering, fluency); Assessments where errors are explicitly corrected
	<i>Interactional frame</i>	Often dyadic (parent-child); roles of teacher (adult) and student (child)	Dyadic (parent-child) or multiparty (multiple adults and/or children); roles of teacher (adult) and student (child)
	<i>Linguistic notes</i>	Recitation, quiz questions, special vocabulary	Recitation, quiz questions, special vocabulary
<b>Economic context</b>	<i>Application to...</i>	Industrial economy; STEM development at scale	Shell money system with bi-lineal accounting
	<i>Relevance of skills (ex.)</i>	Material exchange; timekeeping; employment	Daily greetings; marriage, death; conflict
<b>Historical context</b>	<i>Foundational systems</i>	Shift to industrialized economy and rise of the middle class in 19th and 20th centuries	Unwritten history; however, the shell money system is ancient and spread throughout the region
<b>Cultural context</b>	<i>Associations</i>	Via math, tied to concepts of intelligence; causes anxiety in many individuals; gendered outcomes in adolescents and adults	Relates to a variety of virtuous linguistic performance skills that can be developed, some of them gendered
<b>Social context</b>	<i>Among intimates</i>	Source of potential pride in child’s ability/caregiver’s teaching	Source of potential pride in child’s ability/caregiver’s teaching
	<i>In society at large</i>	Perceived status of family/individual via child economic success; individual participation in interactions reliant on numeracy (e.g., games)	Calculation of relationship to (nearly) anyone—required for knowing mutual obligations
<b>Institutional context</b>		Required knowledge to participate in most institutions	Minimal (few formal institutions)
<b>Other</b>	<i>Variation across class</i>	Structural economic inequality (often interacting with structural racism) leads to variation in outcomes	Limited input and relevance for children without paternal land connections (e.g., non-Rossel fathers)
⋮	⋮	⋮	⋮

**Figure 2**

*Data preparation and expected resource investment for systematic annotation of language events in daylong data. PI = Primary Investigator; LM = Lab Manager; RA = Undergraduate Research Assistant.*

