A R T I C L E

Child Development Perspectives

Learning language in vivo

Marisa Casillas

Department of Comparative Human Development, University of Chicago, Chicago, Illinois, USA

Correspondence
Marisa Casillas, Department of Comparative Human Development, University of Chicago, Rosenwald Hall, 3rd Floor, 1101 E. 58th Street, Chicago, IL 60637, USA.
Email: mcasillas@uchicago.edu

Abstract

In this article, I advocate for an enriched view of children's linguistic input, with the aim of building sustainable and tangible links between theoretical models of language development and families' everyday 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 entry point into context, I outline an approach—which I refer to as language in vivo—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, culture, daylong recordings, input, language development, LENA

INTRODUCTION

How do children learn language? In the last half century, data, methods, and theory attempting to answer this question have grown significantly (Slobin, 2014). Developmental psycholinguistics 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 that 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 to reliably quantify, justifiably compare, and adequately control their presence in our datasets. These disciplinary traditions reflect both our field's origins and where the field 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 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 fruitful in recent years but present a very

limited view of the learner and therefore limit the depth of the theories that arise from them. Understanding the individual phenomena that add up to the wonder of childhood language learning requires a complementary approach. In this article, I move away from the beaten path toward other possible disciplinary traditions to identify shortcomings among these ways forward that may otherwise be considered sensible.

I propose to explicitly engage 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 to make meaningful connections between mainstream theoretical frameworks and families' everyday experiences. In this article, I review how we have typically measured input, sketch an alternative approach—the language in vivo approach—and then close with what we will gain from engaging with context.

Past and current approaches to linguistic input

Now, I address three facets of the current approach to analyzing input: (1) which *sources* of input are included, (2) what are captured in recordings, and (3) how input is *measured*. All three need further consideration in the language in vivo approach.

Sources of input

There is general consensus that child-directed language[1] from adults (CDL_A) is the linguistic input *par excellence*. CDL_A is associated with earlier and cumulatively larger gains in vocabulary and speed of word recognition, 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 contribute significantly beyond vocabulary (e.g., Benigno et al., 2007; Foushee et al., 2021; Oshima-Takane, 1988; Schick et al., 2022).

Child-directed language itself varies immensely between communities in form and function. As linguistic anthropologists have warned, in many communities, language directed to young children is infrequent, involves multiple parties, or focuses more on social practices than on referential communication. The mere presence of CDL_A in and of itself also does not guarantee the scaffolding features presumed to make it helpful in the first place (e.g., McClay et al., 2022). For example, the pitch variation prototypically associated with CDL_A 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_A 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, Euro-American caregivers tend to linguistically interpret and respond to infants' expressions of negative and positive inner experience. Such behavior positions infants as individuals with mature psychological states and creates a foundation for children to continue sharing their inner experiences with others (Gaskins, 2006). These caregivers use CDL_A to socially center children and 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_A 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). To chart the contours of the learner's linguistic environment, we must go beyond a focus on CDL_A and engage with a variety of sources of input (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 children's language environments have changed significantly over the last half century. The foundation of our field was built 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 Language ENvironment Analysis (LENA) system (Greenwood et al., 2011) launched the field into the domain of big-data research with its 16-h audio-recording capacity and software for automated acoustic analysis.[2] While LENA has been an enormous boon to the field, the associated cost of going big is a loss of depth. Specifically, the automated system produces very limited information about the language used[3] and the data are recorded only in audio (e.g., sign language, gaze, and gesture cannot be captured). This tradeoff between scale and depth should make researchers wary of how feasibility and convenience versus theoretical ideals factor 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 drawbacks of relying exclusively on automated output (e.g., Ferjan Ramírez et al., 2021). Encouragingly, protocols and guides for manually annotating daylong

data have started to emerge, some with an eye toward maximizing efficiency (Casillas, Bergelson, et al., 2017; Cychosz et al., 2021; Mendoza & Fausey, 2021). Researchers choosing this path can expect large investments of time and money (see Figure 1). However, the payoff 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 encounter linguistic units (e.g., words and signs, gestures, syntactic structures, communicative acts), making resources such as Child Language Data Exchange System indispensable (MacWhinney, 2000). However, rather than looking at individual phenomena (e.g., frequency of passive syntactic constructions, phonetic properties of spontaneously produced vowels), 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).

Although 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 change verb inflections to appropriately express specific meanings like "I do" vs. "she does" and "she does" vs. "she did") in minimally versus highly verb-inflecting languages, like English/Mandarin versus in Inuktitut/

Murrinhpatha. While we *can* characterize children's language development as a general process, theory must also cover the specific and interacting subprocesses of 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.

Researchers using aggregate input rates have recently focused on a different issue: how to examine *quality* over quantity input (e.g., Masek et al., 2021). However, the apparent tension between quality and quantity may ultimately be a red herring. Some measures of quantity (e.g., CDL_A) capture numerous input characteristics that are canonically considered to be of high quality. Furthermore, 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 U.S. versus Mayan contrast mentioned earlier, different caregiver responses and encouragement of children's contributions predict that turn-taking behaviors will 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 will likely differ across populations.

The biggest issue at stake seems to be how we can focus our research on meaningful linguistic practices in the child's environment while *also* maximizing the naturalness, reliability, and generalizability of our findings. Failure to take this question seriously can lead to limited or even misleading findings, as I will demonstrate.
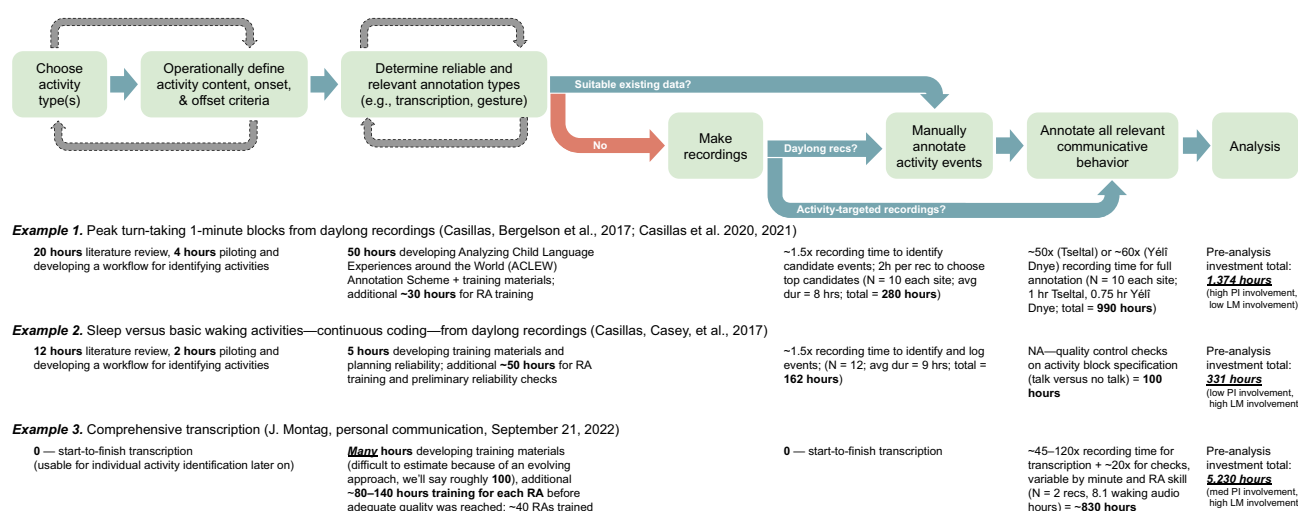

FIGURE 1 Data preparation and expected resource investment for systematic annotation of language events in daylong data. LM, lab manager; PI, primary investigator; RA, undergraduate research assistant. See also, e.g., Ferjan Ramírez et al. (2021) on annotation of automatically selected segments; Cychosz et al. (2021) on clip classification; Bergelson, Amatuni, et al. (2019) on noun annotation; and Mendoza and Fausey (2021) for an in-depth discussion of workflow

Universals in the macro level and potential dead ends

The wide application of aggregate- or indicator-style measures of language can reveal cross-linguistic and cross-cultural 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 a handful of biases affect learning across these groups similarly (Frank et al., 2021). However, in general, when aggregate measures are exported from their original context (e.g., from urban industrial societies to rural subsistence ones), researchers risk discovering consistencies that overestimate similarity and underestimate diversity.

In recent work using manual measures of CDL input rates across diverse cultural contexts, my colleagues and I stumbled on one such possible universal—children under age 3 encounter infrequent CDL, an average of ~2.5–5 min/h, across their daytime hours (Bunce et al., 2022; Casillas et al., 2020, 2021). This macro-level consistency obscures tremendous cross-cultural variation in how caregivers use CDL. For example, ethnographic work characterizes caregiver-child interaction in two subsistence farming communities as child-centric on Rossel Island (in Papua New Guinea) and non-child-centric in a Tseltal Mayan community (in Mexico; Brown, 2014; see also Shneidman & Goldin-Meadow, 2012). Yet, these communities appear nearly identical in CDL baseline input rates (Casillas et al., 2020, 2021). Where have the differences gone? I argue that the differences will be visible primarily 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. But at the macro scale, when averaging a single measure over a wide variety of contexts and scenarios, we cannot see these differences.

Both macro and micro scales provide informative perspectives 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 children's language cannot reach their full theoretical potential. In this vein, I propose what I term the *language* in vivo *approach* to studying children's linguistic experiences.

THE LANGUAGE IN VIVO APPROACH

The core concept behind the language 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 appeals.) The way children learn to attend to and produce language is entwined with what 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 a language, they 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 core mechanisms (e.g., some types of statistical learning) may work similarly across most children, other mechanisms likely vary depending on context (e.g., given the 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 socially center our approach to studying early language experience? Research on language socialization achieves this aim by engaging ethnographically with the community being studied (Ochs & Schieffelin, 1984). However, such an approach is infeasible for most developmental psychologists, who typically have no training in ethnographic methods, and so it is unclear how they should proceed. I propose that we start by scaling down the child's social language environment to a smaller phenomenon that we can identify reliably 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 that are native to developmental psychology with more contextualized considerations of the child's developmental milieu.
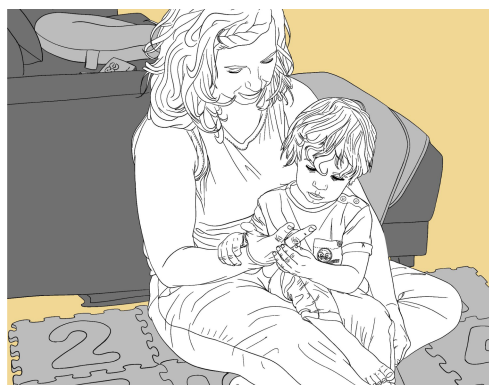
Activity context

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 in 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 enrich their work, grounding broad theories in children's moment-to-moment language experiences—that is, their language experience in vivo (see Figure 2). Two challenges in taking this approach

How is this activity meaningful to participants?

Example:
Considerations for "informal home pedagogy" activity types in the United States versus Rossel Island, Papua New Guinea

Image from a recorded session of tii kala kala conducted with a local research assistant

		Counting activities in the United States	**Tii Kala Kala on Rossel Island, PNG**
General background	Broad domain/themes	Numeracy; education	Kinship; ancestry
	Key concept	"Home is the first classroom"	Your kin shape your day-to-day and your life course
	Relevant ages	Early toddlerhood through primary school	Early toddlerhood through adolescence
	Frequency	Routine and sometimes everyday	Routine but not everyday
	Modality notes	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
Purpose and motive	Why it's worthwhile	Academic/educational/occupational potential; serious consequences for everyday life if basic knowledge is not mastered	Intellectual and social virtu; no serious consequences if not mastered because of kin network support when this knowledge is relevant
	Ideal outcomes	Accurate imitation, then abstraction	Accurate imitation, then abstraction; fluent performance
Implementation	Activity-specific features	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
	Interactional frame	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)
	Linguistic notes	Recitation, quiz questions, special vocabulary	Recitation, quiz questions, special vocabulary
Economic context	Application to…	Industrial economy; STEM development at scale	Shell money system with bi-lineal accounting
	Relevance of skills (ex.)	Material exchange; timekeeping; employment	Daily greetings; marriage, death; conflict
Historical context	Foundational systems	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
Cultural context	Associates	Via math, tied to concepts of intelligence; source of 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
Social context	Among intimates	Source of potential pride in child's ability/caregiver's teaching	Source of potential pride in child's ability/caregiver's teaching
	In society at large	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
Institutional context		Required knowledge to participate in most institutions	Minimal (few formal institutions)
Other	Variation across class	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 Examples of researchers' considerations when contextualizing candidate activities. This chart compares informal home pedagogy activities in the United States and on Rossel Island (Papua New Guinea) across several relevant and interrelated contextual dimensions

are how to identify activities and what to do with information about the activities.

Identifying activities

Here, I focus on identifying activities in long-form naturalistic data (e.g., LENA recordings) because multiple activity contexts, often overlapping, are captured from the child's perspective as they go about their day. Moreover, because this type of recording is still relatively new, the field has not yet developed a set of conventions for finding and choosing activity types (see Mendoza & Fausey, 2021, for more information),[4] leaving room to test a preliminary language in vivo approach.

Researchers using this 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., manually annotating existing data, engaging with work outside the researcher's typical domain). Commonly studied activities like free play and book reading are not exempted. They differ meaningfully 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 idea is not to wholesale manually annotate recordings prior to data analysis. Researchers should begin with hypotheses or selective attention to certain activity types, together with clear data- or community-internal evidence of the activity's relevance and some explicit recognition of the limits of generalization.

How do researchers identify activities in line with an in vivo approach? They might begin with a theme

that implies relevance to participants but requires local specification—for example, asking about how language is (and is not) used by caregivers 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 are critical to identifying the specific activities relevant to this theme. Or researchers might start by establishing the specific participant goals or roles of interest (e.g., typical daily journeys, spontaneous informal pedagogy, encouragement of socially appropriate behavior). Again, the specific relevant activities will differ according to context. Counting activities in the United States differ from recitations of kin ancestry on Rossel Island in Papua New Guinea, but comparing the two provides a more truthful examination of informal pedagogy in each context than would occur by examining one of those activities across both contexts (e.g., asking U.S. parents to engage their children in kin ancestry pedagogy; Figure 2).

Using activity information

Once these recordings are collected, what do we do with them? If the goal is to collect contextually validated findings that are sensitive to both local frames and broader discourses, there is just one answer: detailed inspection and annotation of the data. Such work is costly (Figure 1), but the returns are invaluable in highlighting critical gaps in knowledge and potential ways to bridge them. Such data can also usually be re-used, especially if they are made available to others (e.g., MacWhinney, 2000).

While this work is descriptive, it contributes substantively to theory. Three examples of theoretical work that would benefit from an activity-centered approach are (1) how lexical development is guided by everyday activities at both the domain level (e.g., cars, foods) and the feature level (e.g., deixis [word referent dependence on context], animacy [perceived sentience], register/style of talk), (2) the conditions under which spontaneous informal pedagogy occurs and what material and interactional frames it involves, and (3) how societal expectations about children as language users shape what children attend to and (re)produce during interactions. For each of these topics, researchers would need to do more study via experimental and computational work, elaborating on the proposed processes. The extended in vivo challenge is to find innovative ways to keep more controlled investigations like experiments in touch with the features of interaction that made the descriptive work meaningful.

WHAT WE WILL GAIN

The purpose of the language in vivo approach is to celebrate continuing advances in big data's ability to document developmental environments without losing sight of what children and their families actually *do* with language. I have envisioned this work from the perspective of a Western researcher who engages in cross-cultural comparison and has the privilege of focusing on basic scientific questions about human language-learning capacity. 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 that explicitly intervene on parents' or children's behavior (see Rowe & Weisleder, 2020, for a review).

The in vivo approach can help researchers develop new perspectives on classic questions and delve deeper into less-studied questions. For example, current research aims to explain the processes by which vocabulary arises the way 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 to understand more fully how this early word knowledge informs children's development as competent language users. After all, even when a word is learned early across diverse groups, its representation is built from very different experiences (e.g., the word *aunt* in an isolated urban context vs. in a multigenerational 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 adults intend to contribute. Establishing how children attend to and reproduce others' linguistic practices is closely linked to what children see as relevant—an issue that is addressable in an in vivo approach.

CONCLUSION

While it is tempting to scientifically consider language as a symbolic system (viz., grammar and lexicon) isolated from social systems, doing so limits our ability to directly connect theories with families' everyday experiences. Theory that is constructed without context (or in ignorance of its implicit context) will not apply naturally or equally across diverse domains—it will be limited in its impact. By embracing the social context of language from the viewpoint of activity, researchers using an in vivo approach can equip themselves to construct rich theories of early language learning that have significant potential for practical impact and that foster greater interdisciplinarity and innovation in the field of developmental language science. In summary, to most people, language is how one trades in social,

cultural, and world knowledge. People—including young children—primarily use language to do things with others. Considering social context is critical to building robust theories of how and why we learn and use language. Not just leveraging but *embedding* social context at the heart of our work would pay off immensely in these respects.

ORCID

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

ENDNOTES

[1] I use the term CDL (i.e., not child-directed speech) to refer to spontaneous language use directed to a child.

[2] Large cross-corpus analyses of small data began long before then, largely thanks to the Child Language Data Exchange System, an open online database of naturalistic child language data (MacWhinney, 2000).

[3] Among the annotations are 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 estimates of speech content (e.g., number of nearby adult words). Open-source alternatives improve this outlook somewhat (see Räsänen et al., 2021).

[4] Current automated tools can indicate two activity-related events: nearby electronic noise (e.g., TVs) and infant sleep (Bang et al., 2021; Greenwood et al., 2011).

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. https://doi.org/10.1017/S0305000916000416

Adolph, K. E., & Sternberg, R. (2019). Ecological validity: Mistaking the lab for real life. In R. J. Sternberg (Ed.), *My biggest research mistake: Adventures and misadventures in psychological research* (pp. 187–190). SAGE Publications. https://doi.org/10.4135/9781071802601.n58

Bang, J., Kachergis, G., Weisleder, A., & Marchman, V. (2021). *An automated classifier for child-directed speech from LENA recordings* [Conference presentation]. 46th Boston University Conference on Language Development, Boston, MA.

Bates, E., & Goodman, J. C. (1999). On the emergence of grammar from the lexicon. In B. MacWhinney (Ed.), *The emergence of language* (pp. 29–80). Erlbaum. https://doi.org/10.4324/9781410602367

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. https://doi.org/10.1017/S0305000906007732

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. https://doi.org/10.1111/desc.12715

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 10.1111/desc.12724

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). https://doi.org/10.1075/tilar.12.07bro

Bruner, J. (1985). Child's talk: Learning to use language. *Child Language Teaching and Therapy*, *1*(1), 111–114 10.1177/026565908500100113

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 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. https://doi.org/10.21437/Interspeech.2017-1418

Casillas, M., Brown, P., & Levinson, S. C. (2020). Early language experience in a Tseltal Mayan village. *Child Development*, *91*(5), 1819–1835 10.1111/cdev.13349

Casillas, M., Brown, P., & Levinson, S. C. (2021). Early language experience in a Papuan community. *Journal of Child Language*, *48*(4), 792–814. https://doi.org/10.1017/S0305000920000549

Casillas, M., Casey, K., Norgan, M., Guillu, A., Hsieh, E., & Scaff, C. (2017). *Daylong infant sleep and waking annotations from three societies* [Unpublished raw data].

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. https://doi.org/10.1111/cdev.12974

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. https://doi.org/10.1044/2021_JSLHR-20-00755

de Barbaro, K. (2019). Automated sensing of daily activity: A new lens into development. *Developmental Psychobiology*, *61*(3), 444–464 10.1002/dev.21831

de León, L. (2019). Playing at being bilingual: Bilingual performances, stance, and language scaling in Mayan Tzotzil siblings' play. *Journal of Pragmatics*, *144*(April), 92–108. https://doi.org/10.1016/j.pragma.2018.02.006

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. https://doi.org/10.1371/journal.pone.0151518

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 10.1111/cdev.13495

Foushee, R., Srinivasan, M., & Xu, F. (2021). Self-directed learning by preschoolers in a naturalistic overhearing context. *Cognition*, *206*(January), 104415. https://doi.org/10.1016/j.cognition.2020.104415

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2021). *Variability and consistency in early language learning: The Wordbank project*. MIT Press. https://doi.org/10.7551/mitpress/11577.001.0001

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). Berg. https://doi.org/10.4324/9781003135517

Glas, L., Rossi, C., Hamdi-Sultan, R., Batailler, C., & Bellemmouche, 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. https://doi.org/10.1017/cnj.2018.20

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. https://doi.org/10.1177/1525740110036782

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. https://doi.org/10.1177/09637214211057836

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. https://doi.org/10.1016/S0163-6383(02)00086-3

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. https://doi.org/10.1121/1.1577560

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Erlbaum. https://doi.org/10.4324/9781315805672

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. https://doi.org/10.1111/cdep.12425

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. https://doi.org/10.1111/desc.13180

Mendoza, J. K., & Fausey, C. M. (2021). Quantifying everyday ecologies: Principles for manual annotation of many hours of infants' lives. *Frontiers in Psychology*, *12*, 710636. https://doi.org/10.3389/fpsyg.2021.710636

Montag, J. L. (2020). New insights from daylong audio transcripts of children's language environments. In S. Denson, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd annual virtual meeting of the Cognitive Science Society* (pp. 3005–3011). Cognitive Science Society.

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. https://www.cambridge.org/us/academic/subjects/anthropology/anthropological-theory/culture-theory-essays-mind-self-and-emotion?format=PB&isbn=9780521318310

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. https://doi.org/10.1017/S0305000900012071

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. https://doi.org/10.3758/s13428-020-01460-x

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. https://doi.org/10.1146/annurev.psych.54.101601.145118

Rowe, M. L., & Weisleder, A. (2020). Language development in context. *Annual Review of Developmental Psychology*, *2*(1), 201–223 10.1146/annurev-devpsych-042220-121816

Schick, J., Fryns, C., Wegdell, F., Laporte, M., Zuberbühler, K., van Schaik, C. P., Townsend, S. W., & Stoll, S. (2022). The function and evolution of child-directed communication. *PLoS Biology*, *20*(5), e3001630. https://doi.org/10.1371/journal.pbio.3001630

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. https://doi.org/10.1111/j.1467-7687.2012.01168.x

Slobin, D. I. (2014). Before the beginning: The development of tools of the trade. *Journal of Child Language*, *41*(S1), 1–17. https://doi.org/10.1017/S0305000914000166

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*, 708887. https://doi.org/10.3389/fpsyg.2021.708887

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. https://doi.org/10.1111/cdev.13072

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 10.1111/cdev.13089

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), 405–416. https://doi.org/10.1037/dev0001285

How to cite this article: Casillas, M. (2022). Learning language in vivo. *Child Development Perspectives*, *00*, 1–8. https://doi.org/10.1111/cdep.12469