Journal of Child Language (2023), 1–23
doi:10.1017/S0305000922000617

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Combining observational and experimental approaches to the development of language and communication in rural samples: Opportunities and challenges

Alejandrina CRISTIA[1,*], Ruthe FOUSHEE[2], Paulina ARAVENA-BRAVO[3,4], Margaret CYCHOSZ[5,6], Camila SCAFF[1,7] and Marisa CASILLAS[8]

[1]Laboratoire de Sciences Cognitives et de Psycholinguistique, Département d'études cognitives, ENS, EHESS, CNRS, PSL University, Paris, France
[2]Department of Psychology, University of Chicago, Chicago, IL, USA
[3]Departamento de Fonoaudiología, Facultad de Medicina, Universidad de Chile, Santiago, Chile
[4]Escuela de Psicología, Pontificia Universidad Católica de Chile, Santiago, Chile
[5]Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland, USA
[6]Department of Linguistics, University of California, Los Angeles, California, USA
[7]Institute of Evolutionary Medicine, University of Zurich, Zurich, Switzerland
[8]Department of Comparative Human Development, University of Chicago, Chicago, IL, USA
*Corresponding author. Alejandrina Cristia, Laboratoire de Sciences Cognitives et de Psycholinguistique, Département d'études cognitives, ENS, EHESS, CNRS, PSL University, Paris, France. Email: alecristia@gmail.com

(Received 13 August 2022; revised 29 September 2022; accepted 04 October 2022)

Abstract

Multiple approaches – including observational and experimental – are necessary to articulate powerful theories of learning. Our field's key questions, which rely on these varied methods, are still open. How do children perceive and produce language? What do they encounter in their linguistic input? What does the learner bring to the task of acquisition? Considerable progress has been made for the development of spoken English (especially by North American learners). Yet there is still a great deal to discover about how children in other populations proceed, especially populations in rural settings. To examine language learning in these populations, we need a multi-method approach. However, adapting and integrating methods, particularly experimental ones, to new settings can present immense challenges. In this paper, we discuss the opportunities and challenges facing researchers who aim to use a multimethodological approach in rural samples, and what the field of language acquisition can do to promote such work.

Keywords: methods; observational; experimental; rural; small-scale; culture

Introduction

Discussions regarding the lack of racial, socioeconomic, and linguistic diversity in both developmental psychology (e.g., Singh, Cristia, Karasik & Oakes, 2021) and language

acquisition research (Kidd & Garcia, 2022; and commentaries) have gained recent prominence; in line with many recommendations for promoting greater diversity in European and Anglo-American samples (Sugden & Moulson, 2015). Meanwhile, discussion on how to extend diversity to include other regions and cultural groups has received much less attention. This paper brings together six researchers with expertise in the use of both observational and experimental child language research methods in rural settings. Young children from rural settings are under-represented in developmental research (e.g., Cristia, 2022), particularly in experimental research. Here we provide an overview of selected work that (a) combines observational and experimental perspectives on the same linguistic or communicative phenomenon in children and which is (b) based in a rural community. We then highlight methodological and theoretical challenges, ending with recommendations that could facilitate this type of research in the future.

We use the key terms, "observational" and "experimental," in a broad, encompassing manner. OBSERVATIONAL methods include behavioral observations that are carried out through live, in-the-moment annotations or via video or audio-recording devices for later analysis, provided that they are systematic and do not involve any specific elicitation or imposed constraint with respect to conversational topic or situation. An example is naturalistic video recordings of caregivers and children in their home. EXPERIMENTAL methods include any situation where investigators curate the child's context. Common experimental approaches include non-word or sentence repetition tasks, elicitations of behavioral responses (e.g., pointing) or linguistic utterances, and more recent gaze-direction and looking-time based studies. We also include in this category studies involving constrained interactions – for example, where the child and a single caregiver are provided specific toys to engage with. We focus here on cases where these two broad types of methods have been used together to examine a linguistic or communicative phenomenon in children under 5 years of age in the same population. While these criteria exclude several important and relevant bodies of work (see Supplementary Table 1 in Cristia, Foushee, Aravena-Bravo, Cychosz, Scaff & Casillas, 2022), they provide us with a well-defined conceptual and theoretical foundation to consider the scientific advantages and challenges of child language research in rural settings.

Why focus on rural populations?

We do not believe that, in principle, language is affected by whether a child is growing up in a "rural" area as opposed to an "urban" one. We also do not mean to suggest that "rural" populations form a single group – indeed, these populations are extremely diverse in how aspects of their rurality shape everyday communicative interactions and language development. For example, the language experiences of children in rural and agricultural areas of Canada versus subsistence farming communities in Papua New Guinea differ along a number of dimensions, including access to electricity, use of motor vehicles, family housing organization, geographical terrain, and many others. To understand how these aspects of environment and culture influence language experiences, we must conduct research with diverse populations. Ideally, the present paper would be more specific in identifying meaningful subgroups of rural communities along dimensions potentially relevant for language development. At present, however, the reality is that research in rural contexts is so limited that we lack a sophisticated or systematic approach for distinguishing between types of rurality and how they impact language development.

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

Some of the common features of the communities reviewed here include lesser access to institutional and infrastructural services such as electricity, telecommunications, health services, education, and industrial products (e.g., books, children's clothes, etc.). Many of these populations are also undergoing significant change as they integrate into the global market economy, which also has implications for, e.g., educational enrollment, exposure to non-heritage languages, and the distribution of child caregiving responsibilities. However, the communities presented here differ from each other in a number of ways.

Some, for example, are small-scale communities and indigenous societies. Small-scale communities can be defined as "a society of a few dozen to several thousand people who live by foraging wild foods, herding domesticated animals, or non-intensive horticulture on the village level" (O'Neill, 2008). Because much of the population's sustenance comes from local natural resources, children may spend the bulk of their time in the same spaces where adults work, leading to greater cross-generational integration. Childcare, and thus language input, is also likely to be distributed across multiple family members (Loukatou, Scaff, Demuth, Cristia & Havron, 2021). These characteristics have consequences for how language acquisition can and should be studied.

Other communities discussed here are not small-scale societies (e.g., Abels, 2020; Fibla Reixachs, 2021; see also discussion regarding Quechua samples in Cychosz's studies below). In addition, many could be described as coming from state-minoritized indigenous societies (e.g., Tseltal Mayan sample in Casillas, Brown & Levinson, 2020), but that is not the case for all of them (e.g., the Awadhi sample in Fibla Reixachs, 2021; the Tsimane' sample in Cristia, Dupoux, Gurven & Stieglitz, 2019). Both small-scale and indigenous societies may speak languages that are less well-documented than state languages, but this paper does not focus on under-described languages *per se*.[1] Our rural versus non-rural distinction has only partial overlap with the WEIRD vs. non-WEIRD dichotomy, in part because we do not believe this dichotomy is useful and in part because we would consider some communities in, e.g., Australia, the US, and Canada to be included in our scope. We therefore use rurality as a proxy feature that describes, in very broad terms, the communities that can be found in this review.

Importantly, this review does *not* include the immense, rich body of literature using quantitative and qualitative observational methods in rural communities from anthropology, linguistics, and beyond, because it specifically aims to advocate for the COMBINATION of observational and experimental methods in the same setting. We believe most readers of this work will be convinced (as we are) of the interest of employing observational methods in rural settings. But it is important to remember that observations cannot isolate causal factors underlying learning. In this sense, experimental methods can help us focus on specific structures in perception or production, allowing us to more closely establish causality or connections between language learning settings and developmental processes. In contrast, a common criticism of experimental work in rural settings is that the need for experimental control results in artificial and sometimes culturally or linguistically inappropriate study designs that inaccurately reflect participants' true language behaviors and understanding (a tension articulated also for work in rural contexts on the purported "word gap", e.g., see Ochs & Kremer-Sadl, 2020). This criticism is legitimate – the value of an experimental investigation is limited by its local

[1]In fact, one of the studies we review includes no mention of the language spoken by the represented community (Abels, 2020).

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

relevance and validity. However, experimental work that is done hand-in-hand with observational work can help develop theories benefiting from the merits of both methods: richer, contextualized descriptions of behavior bolstered by targeted tests shedding light on the nature of correlational and causal relationships identified in the naturalistic data. For this reason, we see the COMBINATION of observational and experimental work as having the strongest potential for advancing our knowledge of child language development in rural contexts. Researchers taking this dual approach must master multiple methods and often carry them out in new populations, rendering this work challenging and relatively rare. The present paper highlights successful cases of observational-experimental approaches to inspire future work. Take, for example, Clifton Pye's work on morphosyntactic development in K'iche' (Core Kichean, Mayan): Pye and Quixtan Poz (1988) used a simple picture-choice task to establish that K'iche' children understood constructions that were extremely rare in spontaneous conversation (Pye, 1980). Without this experimental data, it would have been difficult to make a solid case for children's comprehension of these constructions.

This paper reviews how past and ongoing work has managed to contend with the fact that current methodologies have been developed for urban settings and are thereby often difficult – sometimes impossible – to adapt to rural settings. Our efforts here are aimed at inspiring additional future work combining observational and experimental methods in understudied contexts to shed light on language development.

Overview of relevant work

The literature in the present overview is synthesized in Tables 1 and 2. It is the result of the authors' combined personal libraries, previous literature reviews, searches in scholar.google.com and SciELO.org (using English, and Spanish keywords), as well as emailing potential authors and relevant email listservs in developmental psychology and child language.

While many researchers work on multiple linguistic levels, this overview organizes studies by individual linguistic level to help the reader digest this broad body of work (e.g., gestures, morphosyntax, lexicon, and phonology). We present these lines of work in rough chronological order by date of publication as a way of embedding the review in historically relevant methodological, conceptual, and technological changes. Some of the oldest combinations of experiments and observations are those bearing on morphosyntactic development, which could rely on transcriptions of children's language production using audio tapes, a technology that has been available for over 70 years, in combination with picture selection methods that are even older. The second set of studies, bearing on social cues and gestures, has relied on observations, especially video recordings, which were already widely available in the 1990s. Next summarized are phonological investigations, which received a boost with the increased availability of equipment for doing fine-grained acoustic analyses, even under field conditions. The final set of studies bear on the lexicon. Although some studies employ picture selection or naming techniques that well predate contemporary work, other studies now use portable screens and eye trackers with looking-while-listening and similar language processing methods. Both phonological and lexical analyses also incorporate measures of children's production and linguistic input – some recent work even uses long-form recordings, a relatively novel technique, that allows for both manual and automatic acoustic analysis. For basic information on each community covered, see Table 2.

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

Table 1. Research combining observational and experimental approaches to describe early language acquisition (included if at least some of the children were aged 0-5 years) in rural populations (see main text for definition). See Table 2 for information on the communities. Age range is given in years. See Supplementary materials on other language and communication related phenomena not covered in this main text

Aspects studied	Community	Age range	Observational	Experimental
Morphosyntax	K'iche'-speaking village	3–5	audio-recordings (Pye, 1980)	picture identification (Pye, 1988)
	Sesotho-speaking village & towns in Mokhotlong	1–3	long-form audio-recordings (Demuth, 1989)	picture identification, picture description, & novel verb generalization (Demuth et al., 2010); picture identification (Demuth et al., 2003, 2005)
	Yucatec Maya-speaking villages	1–3	corpus analysis of existing long-form transcription (Espinosa Ochoa, 2022)	video recordings with controlled toys (Espinosa Ochoa, 2017; 2022)
Social cues & gestures	Yucatec Maya-speaking villages	~1	scan sampling & video-recordings (Salomo & Liszkowski, 2013)	pointing comprehension (Liszkowski, 2011), elicited pointing (Liszkowski et al., 2012)
	Yélî-speaking villages	~1	observations, video-recording, systematic sampling (Brown, 2011)	elicited pointing (Liszkowski et al., 2012)
	Tseltal-speaking community	~1	observations, video-recording, systematic sampling (Brown, 2011)	elicited pointing (Liszkowski et al., 2012)
	Villages in Vadodara	~1	observations, video-recordings (Abels, 2020)	elicited pointing (Abels, 2020)
Phonology	Tsimane' villages	0–11	long-form audio-recordings (Scaff et al., 2022), time sampling (Cristia et al., 2019)	non-word repetition (Cristia et al., 2020)
	Yélî-speaking villages	0–11	long-form audio-recordings (Casillas et al., 2021)	tablet-based word recognition (Casillas & Levinson, personal communication), non-word repetition (Cristia & Casillas, 2022), habituation-based phoneme discrimination (Casillas, personal communication)

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

Table 1. *(Continued)*

Aspects studied	Community	Age range	Observational	Experimental
	Quechua-speaking communities near a town	0–10	long-form audio-recordings (Cychosz, 2021a)	word repetition (Cychosz, 2021a, b)
Lexicon	Awadhi-speaking village	0–2	long-form audio-recordings (Fibla, 2021)	looking-while-listening (Fibla, 2021)
	Tsimane' villages	0–11	long-form audio-recordings (Scaff et al., 2022), time sampling (Cristia et al., 2019)	tablet-based word comprehension (unpublished), looking-while-listening (unpublished)
	Tseltal-speaking community	0–4	long-form photo-linked audio recordings (Casillas et al., 2020)	looking-while-listening (Casillas, personal communication), looking-while-listening (Foushee & Srinivasan, 2022)
	Yucatec Maya-speaking villages	0–3	video-recordings (Shneidman & Goldin-Meadow, 2012; Padilla-Iglesias et al., 2021)	word comprehension and production (Shneidman & Goldin-Meadow, 2012); novel noun learning (Shneidman, personal communication)
	Yélî-speaking villages	0–12	long-form photo-linked audio recordings (Casillas et al., 2021)	looking-while-listening, tablet-based word recognition (Casillas, personal communication)

Morphosyntax

Pye's research on the acquisition of K'iche' (Core Kichean, Mayan) in Guatemala is brimming with findings on K'iche' children's morphosyntactic development. We focus here on one pair of results that highlight Pye's combination of observation and experiments. Observational analyses of conversations with three children between ages two and four years suggested that children's spontaneous use of passive constructions (e.g., *The ball was kicked*, in English) was much higher than found in past work on languages such as English (Pye, 1980; Pye & Quixtan Poz, 1988). This past work had led to the assumption that passive constructions are linguistically and cognitively marked in human learners, but the K'iche' data suggested otherwise. Pye followed up on these observational findings with a picture-pointing experiment demonstrating that, in fact, K'iche children appeared to comprehend passives BETTER than actives by 4–5 years, which went against theoretical reasoning on passives at the time. Notably, the pictures were drawn by the researcher while in the field, highlighting the fact that this approach does not require significant resources. That said, the authors comment on the fact that some of the images were more easily understood than others, leading to variable performance and item effects.

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

Table 2. Further information on the communities sampled in studies in Table 1, sorted alphabetically for ease of reference.

Community	Language (family)	Region, country	Main subsistence source
Awadhi-speaking village	Awadhi (Eastern Hindi, IndoAryan)	Shivgarh, Uttar Pradesh, India	Subsistence farming
K'iche'-speaking village	K'iche' (Core Kichean, Mayan)	Zunil, Guatemala	Wage and subsistence farming
Quechua-speaking communities near a town	South Bolivian Quechua (Quechua)	Cochabamba, Bolivia	Informal economy including trading and livestock
Ngas village	Ngas (Chadic, AfroAsiatic)	Tuwan, Plateau State, Nigeria	Not reported
Sesotho-speaking village and towns in Mokhotlong	Sesotho (Sotho, Bantu)	Mokhotlong, Lesotho	Subsistence farming
Tseltal-speaking community	Tseltal (Tseltalan, Mayan)	Chiapas, Mexico	Subsistence farming, supplemented by wages
Tsimane' villages	Tsimane' (isolate)	Beni, Bolivia	Hunting, gathering, fishing, subsistence farming
Villages in Vadodara	Not reported	Vadodara, Gujarat, India	Farming complemented with wages and self-employment
Yélî-speaking villages	Yélî (isolate)	Rossel Island, PNG	Subsistence farming, fishing
Yucatec Maya-speaking villages	Yucatec (Yucatecan, Mayan)	Yucatan, Mexico	Slash and burn subsistence farming

The work of Demuth and colleagues on children learning Sesotho (Sotho, Bantu) provided additional data on passive construction acquisition (we note that this same dataset has been used elsewhere: e.g., Demuth, 1992; Demuth, Machobane & Moloi, 2003; Demuth, Machobane, Moloi & Odato, 2005; Loukatou et al., 2021; including for similar work on double object constructions). Demuth (1989) describes the collection of audio-recordings, gathered longitudinally over a two-year period, from four children growing up in rural Lesotho. Transcriptions were carried out by the author (a linguist) in collaboration with the children's mothers and grandmothers. Passives were overall uncommon but present in children's Sesotho production, making up ~2% of all utterances by about four years of age, compared to ~6% in adult speech. Productive use of the construction with a variety of verbs led Demuth to conclude that passives were acquired by age four in this population. Experimental tests (Demuth, Moloi & Machobane, 2010) confirmed this conclusion, showing that three-year-old Sesotho learners from town-based preschools could reliably identify pictures corresponding to the correct meaning (e.g., point to "the boy is being fastened by the mother", when viewing pictures of a boy fastening an apron onto a grown woman and vice versa). Children identified picture

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

referents for passive utterances with ~73% accuracy (actives were identified at ~82% accuracy). Adults showed 94% accuracy for both, with the 6% errors being attributed to the visual stimuli not being sufficiently clear (e.g., for the verbs corresponding to the English concept "help" and "see"). The hypothesis that passive construction acquisition is well underway at age three years was cemented by additional experiments eliciting children's production of passives, including for novel verbs, where all children produced the passive construction correctly, with an overall group accuracy above 98%. These groundbreaking studies are prime examples of the value of cross-linguistic and cross-cultural work in developing theories of language learning.

Last but not least, Espinosa Ochoa (2017, 2022) pooled together highly naturalistic existing data with existing and newly collected semi-controlled naturalistic video recordings to characterize the different stages in the acquisition of the Yucatec Maya (Yucatecan, Mayan) deictic system. In this case the observational and experimental data were similar enough to undergo the same set of linguistic analyses rather than understood as making separate contributions to understanding deixis. The total data pool involved long-form existing data from one child (1;08), longitudinal semi-controlled, then naturalistic video data from another (2;10–3;05), and newly-collected semi-controlled video data from two more (2;0 and 2;08). The Yucatec Mayan deixis system is complex, involving demonstratives (e.g., *this*, *that* in English), locative adverbs (e.g., *here*, *there* in English), ostensive evidentials (highlighting the source of information), and modal adverbs (comparing or exemplifying an entity). This range of deictics can pick out entities or locations and may depend on the speaker highlighting a presentational act or attending to whether a topic is currently in focus. At first, children produced deictics with only broad functions, often using a shortened proto-form (*(e)lo'/(e)la'*) that reflects the shared adult forms. Locatives and modal adverbs were rare in this stage, with children more often using the demonstratives and ostensive evidentials, but without clearly distinguishing between them yet. Later on, children began to experiment with more specialized deictic functions, sometimes overgeneralizing to produce ungrammatical or infelicitous uses of the adult-like terms – all part of acquiring a complex deictic system, now elaborated for Yucatec Maya thanks to this work.

Gestures and social cues

Whereas in the previous section we gave three separate examples, in this section we can integrate our review across studies because they have used more directly comparable annotation categories and analytical approaches. In particular, the reviewed work covers: (1) proportion of time spent in triadic interaction and (2) gesture use.

There is wide cultural variability in the proportion of time spent in triadic interaction and coordinated joint attention. Childers, Vaughan, and Burquest (2007) employed video-recordings centered on Ngas (Chadic, Afro-Atlantic) mother-child dyads. In their data, infants aged about 1 year spent 50% of the observation time interacting with others, interacting with the same object as others, or in coordinated joint attention. The proportion of time spent in "complex" joint attention (e.g., interacting with the same object as others, or in coordinated joint attention) increased over age. Salomo and Liszkowski (2013) used scan sampling (live observations with online coding of events by a researcher observing the interaction in real time) and found that Yucatec infants spent about 10% of their time in dyadic interaction with a caregiver and about 7% in

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

triadic joint action involving an object. Brown (2011) describes a marked difference in the frequency with which children engage in verbal interaction with others between two rural communities (45% of 5-minute observation periods for children acquiring Yélî Dnye (isolate) versus 20% for children acquiring Tseltal (Tseltalan, Mayan), based on video-recordings), mainly due to Yélî adults initiating twice as many interactions with children as Tseltal adults did (5.1 per minute as opposed to 2.3 per minute). Finally, Abels (2020), who observed a group of Indian infants in the Gujarat region, analyzes her data based on the number of triadic interaction events per hour (4.6 child-initiated, and 2.5 adult-initiated) rather than proportion of time, rendering comparison with the other studies complicated.

As for gestures, Yucatec Mayan infants observed by Salomo and Liszkowski (2013) produced about 5 gestures per hour at 8 to 11 months, and twice as many as that by 12 to 15 months; adults produced about double the number of gestures in interaction with children at each of those ages. In Abels (2020), Indian 9-month-olds were found to produce about .8 gestures per hour, which is markedly lower than the 5 per hour reported by Salomo and Liszkowski (2013) for similarly aged Yucatec Mayans (see also Callaghan, Moll, Rakoczy, Warneken, Liszkowski, Behne & Tomasello, 2011 including Peruvian and Indian samples). Abels (2020), however, comments on the importance of considering behaviors other than point gestures, observing that rural mothers frequently manipulated the child's body to facilitate joint attention (about 2.9 times per minute).

Are there differences in the emergence of pointing across these cultures? Brown (2011) reports qualitative similarities in the function of deictic pointing based on observations gathered from a total of 5 Yélî Dnye- and Tseltal-acquiring children, but comments that sparseness in the dataset did not allow quantitative analysis (note 14). To overcome this data sparseness – common in more naturalistic (uncontrolled) approaches – many researchers have turned instead to experimental approaches.

For example, Childers et al. (2007) report descriptive analyses of a standardized task aimed at eliciting communicative gestures among Ngas toddlers, which involved the use of novel toys (one of which frightened some of the children). Preliminary data from this same study suggested ages for pointing comprehension: 10 out of the 12 one-year-old children tested followed a researcher's live point (and head turn). An ingenuous experimental set-up allowed Liszkowski, Brown, Callaghan, Takada, and de Vos, (2012) to statistically document cross-cultural stability in the frequency and morphology of pointing production: surprising items were displayed on a vertical, out-of-reach surface and caregiver-infant dyads were invited to freely view the items, eliciting spontaneous points from a third of 7–9-month-olds, three quarters of 10–14-month--olds, and 100% of 15–17-month-olds in a sample that included infants from the same Yucatec, Yélî Dnye, and Tseltal communities whose observational data was just summarized, among others. Infants from the same Yucatec Mayan community were reported in other work to comprehend deictic gestures well (unpublished study mentioned in Liszkowski, 2011). In this study, an experimenter sat between two occluders in front of the infant and pointed to one side; then the infant was allowed to explore the environment (walking or crawling). Infants tended to explore more behind the occlusion that had been pointed to. In sum, this work suggests that there is significant cross-cultural variability in the spontaneous use of gesture in general (and pointing in particular), but that there is also cross-cultural stability in both the production and comprehension of pointing when the context is controlled.

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

Phonology

Three recent studies in rural populations have employed non-word repetition (NWR) as part of their experimental approach to gauge phonological development. NWR is an experimental procedure that measures participants' short-term phonological working memory and processing. In the task, participants listen to non-words – word-like phonological strings that are phonotactically legal but semantically meaningless in their language(s) – and attempt to repeat them. NWR can therefore reflect several aspects of phonological encoding and the perceptual-articulatory loop. In each study that employed NWR, the authors made some reference to children's language experiences (e.g., rates of directed linguistic input), as gleaned from observational research on the same population (and in one case, the same children); however, unlike the work on gesture, the three sets of authors do not operate on the same assumptions, nor reach similar conclusions.

Cristia, Farabolini, Scaff, Havron, and Stieglitz (2020) found lower NWR accuracy among 16 children (and 13 adults) in a rural Tsimane' (isolate) sample than in previous developmental NWR studies. Cristia et al. (2020) suggested that the low NWR accuracy in this population could stem from the fact that infants and young children are seldom directly talked to (Cristia et al., 2019). The authors reason that perhaps NWR indexes a specific way of processing phonological material, using short-hand phonological or sub-lexical representations, which is boosted by higher levels of directed linguistic input via greater pressure to engage in phonological processing. It should be noted, however, that the study differed from many other developmental NWR studies in that participants were not tested in isolation but in a group. The target non-word was also produced by a non-native speaker of Tsimane' during the experiment. Finally, the input estimation to which the experimental data are related was done based on systematic observations rather than recordings, which may yield different observer effects, observed activities, and thus observed estimations (Scaff, Casillas, Stieglitz & Cristia, 2022).

Cristia and Casillas (2022) found that NWR rates among 40 Yélî Dnye-acquiring children were considerably high, and certainly higher than reports for the Tsimane' community just summarized. There were many differences between the two studies, including the fact that Yélî Dnye-acquiring children were presented with a native speaker's production and that they were tested individually. Perhaps the most important factor was the presence of a highly experienced local research assistant, who had been working with the team for over a decade, and who was able to coach the majority of children through the basic demands of the task. As a reminder, linguistic input was the attributed cause of lower NWR accuracy in Cristia et al. (2020), but it is unclear how well the NWR findings from Yélî Dnye-acquiring children support this notion, in part because linguistic input has been measured using very different methods between these two communities (Casillas, Brown & Levinson, 2021; Cristia et al., 2019). However, recent work from Scaff and colleagues (2022) suggests that Tsimane' children may hear similar if not greater rates of child-directed speech as Yélî Dnye-speaking children when a permissive input-counting workflow is used. For a more restrictive input-counting workflow, e.g., excluding non-near or overlapping speech, it is unknown how the rates of input for Yélî Dnye compare because this type of analysis has only been conducted for the Tsimane' data (Scaff et al., 2022). Setting this aside, Cristia and Casillas (2022) also reported that children's NWR accuracy at the item level correlated with the frequency with which the items' phones can be found across typologically diverse languages: non-words with sounds that occur more frequently across languages were repeated more accurately. Importantly, this effect was not explained away by phoneme type or token frequency

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

in Yélî Dnye spontaneous speech, which the authors derived from transcriptions of child-directed and child-overheard input captured via long-form recordings (Casillas et al., 2021; see also Bunce et al., 2022).

The third study using NWR bears on children bilingual in Quechua (Quechua) and Spanish (Romance, Indo-European). This community lives in and around a mid-size town in Bolivia, more market-integrated than the two other communities discussed in this subsection. Data collection thus benefited from the presence of electricity, and internet access at a reasonable distance from the community, allowing the investigator (Cychosz) to collect long-form recordings using the cloud-based LENA System (Gilkerson & Richards, 2020). We acknowledge that the concept of rurality applies more loosely to this community. Still, this work represents a community undergoing market integration influencing families' customs and habits. The most innovative aspects of this research lie in the combination of long-form recordings, experimental elicitations, and efficient methods of annotation of long-form data. This combination allowed Cychosz to contribute several interesting findings to this line of work. For example, in Cychosz (2022b), the long-form recordings were split into 30-second clips, algorithmically filtered for having at least some speech, and then presented to human annotators in a random order. The humans made simple decisions upon hearing the 30-s clips – namely, whether the language used was Quechua, Spanish, or both, and who was speaking (the key child, a female adult, etc). The system used to present these 30-s clips was set up such that additional clips from the same child continued to be presented until the portion of clips that were in Quechua versus Spanish stabilized. This allowed the efficient estimation of Quechua/Spanish use by the child as well as Quechua/Spanish exposure from other speakers (see also Cychosz, Cristia, Bergelson, Casillas, Baudet, Warlaumont, Scaff, Yankowitz & Seidl, 2021a; Cychosz, Villanueva & Weisleder, 2021b), which would have been impossibly time-consuming if one would have wanted to exhaustively annotate the 500 hours of audio recordings found across the 40 participants in this study. Those ratios of bilingual language use were then employed to predict children's accuracy and morphological processing in repeating both words and non-words (Cychosz, 2022a, 2022b).

For word repetition tasks, Cychosz (2022b) used a wug-type task, asking children first to repeat the played-back names of high-frequency nouns shown to them with photographs. The experimenter then placed a plastic bug on top of the photograph and asked "where is the bug?", where the response requires the child to append a suffix to the noun. This design allowed the elicitation of both bare nouns and inflected nouns, enabling the researcher to examine coarticulation at and within morphological boundaries. For NWR tasks, Cychosz (2022a) found that children had higher accuracy for Quechua nonwords than Spanish nonwords, and attributed this finding to the longer modal word length in Quechua that may lead children to develop a proportionately stronger phonological working memory for that language. Additionally, using estimates of the children's bilingual language exposure, Cychosz (2022a) found a positive correlation between NWR accuracy (in Spanish or Quechua) and quantity of child-directed speech, but not overall speech, in the children's environments. For the word repetition tasks too, the children's coarticulation patterns varied as a function of morphological complexity in interaction with bilingual exposure, which was quantified from the long-form recording data. In short, children who spoke more Quechua in their daily lives showed stronger evidence of morphological decomposition in their Quechua word production (Cychosz, 2022b). Interestingly, this was not obvious when using parental report of bilingual exposure, highlighting the relevance of using direct measurements as opposed to indirect (e.g., parental) observations.

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

Lexicon

The quantity of directed, child-contingent input is theorized to predict child vocabulary development. In one of the key early findings on children's verbal input and outcomes, Shneidman and Goldin-Meadow (2012) studied the quantity of speech addressed to toddlers growing up in a Yucatec Mayan village. A very small proportion of the utterances in manually annotated, hour-long video recordings was directed at the target children at an early age (21% at 13 months), though this proportion increased with age (60% among 35 month-olds). Across the sample, much of the directed speech came from other children. The researchers repeated the study, but this time with no video camera and by manually tallying utterances as they were spoken, finding higher input rates overall but similar proportions of child-directed input. The relatively high prevalence of overhearable verbal input, together with prior work suggesting Mayan children excel at learning via observation (see Shneidman & Goldin-Meadow, 2012 for a review), led Shneidman and colleagues to collect children's word comprehension by adapting the Peabody Picture Vocabulary Test (Dunn & Dunn, 2007), and word production by adapting the Expressive One-Word Picture Vocabulary Test (EOWPVT; Brownell & Academic Therapy Publications, 2000). In both cases, photographs were used instead of the hand-drawings used in the previously summarized work. The authors also only used items where the label given for a photo was found to be the same across four pilot adult participants.

Using a score that combined comprehension and production, the researchers asked: what aspects of children's input predicted their vocabulary size? Surprisingly, only rates of child-directed (and not overhearable) speech significantly predicted children's vocabulary. Six years later, the team visited the same village again and repeated the collection of video recordings and manual annotations, this time examining how the market integration that had taken place in the intervening years affected the use of Yucatec as opposed to Spanish (Padilla-Iglesias, Woodward, Goldin-Meadow & Shneidman, 2021). Results suggested that a greater proportion of children's overall input was in Spanish, and that this change was driven by other children's speech not directed to the recorded child. Additionally, the team observed slightly lower levels of child-directed speech than seen six years prior, suggesting that market integration in this particular case did not lead to greater verbal input. To our knowledge, they have not yet reported on how these changes impact predictions linking children's language experiences with their vocabulary development.

Fibla Reixachs (2021)'s Chapter 5 reports on a study combining looking-while-listening measures with long-form recordings collected either in infancy (4–13 months) or toddlerhood (14–25 months) in a sample of ~80 infants growing up in poverty in a rural village of Uttar Pradesh (India) and learning Awadhi (Eastern Hindi, IndoAryan), a large sample enabled by a collaboration with a local NGO. Participants were divided into two groups as a function of maternal education: some mothers had completed six or fewer years of education (at most completing primary school, a group the authors called "low-SES") and the other mothers had completed high school ("high-SES"); mothers with intermediate levels of education were not recruited. Fibla Reixachs (2021) extracted peak input and output language by analyzing long-form recordings with automated software, using as predictors the maximum hourly adult word count, child-adult conversational turn count, and child vocalization count across the 16 hours of each child's recording. Surprisingly, peak hourly counts in all three areas (adult words, conversational turns, and child speech) were significantly lower for the toddlerhood recordings than the infancy recordings. In addition, among low-SES children, she found that infants who heard more

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

child-adult turn counts looked longer at the referred-to image during a looking-while-listening task. However, no reliable correlations were found either for low-SES TODDLERS nor high-SES infants and toddlers. Fibla Reixachs (2021) points out critical issues in the interpretation of both the long-form recording data and the looking-while-listening data. For the long-form recordings, she proposed that toddlers spend time with different people (for example, older children) than infants; and that perhaps with greater mobility, older children leave an adult-centered, language-rich environment, leading to lower input counts for older (mobile) children. As for looking-while-listening, Fibla Reixachs (2021) notes that the psychometrics of the task had not been firmly established for this population; a study summarized in chapter 4 of the same dissertation demonstrated that only children over 41 months of age showed above-chance performance.

Casillas et al. (2020, 2021) established the quantity and composition of input afforded to 10 Yélî Dnye- and 10 Tseltal-acquiring children under 3;0. The team also used the children's spontaneous productions from the long-form recordings to establish that, in broad terms, these infants were linguistically developing along a timeline similar to that documented in prior language acquisition research: canonical syllables emerge within the first year, then an increasing proportion of vocalizations containing recognizable words appears at the start of the second year, with word combinations following soon after. Children in both communities, tested up to age 4;0, exhibit above-chance word recognition in looking-while-listening studies (Casillas, personal communication; sample sizes around N = 20–30 in each study). Independent of this work, Foushee and Srinivasan showed evidence in a different Tseltal community for early recognition of both common nouns and honorifics using a yoked-pairs version of this same looking-while-listening type technique with children ages 5 to 12 months (Foushee & Srinivasan, 2022).

The use of a portable lab (Figure 1) may have been key to successful execution of the looking-while-listening experimental method in these studies. In these cases, the lab was constructed as a tent or cabin-like structure inside of an existing building. This type of setup also allows the experimenter to hide behind the testing structure (e.g., behind the wall that participants face), affording the participant both audio and visual isolation from distraction. In all cases, the accompanying caregiver sat with the child on their lap, in front of the displayed stimuli (i.e., single screen for Casillas and colleagues; two screens for

Figure 1. Portable tent for experiments used by Casillas and colleagues (left) and Foushee and colleagues (right). Left: the tent includes a participant observation video camera (above laptop screen), stereo speakers, heart rate monitor (worn by the child), and left- and right-side target regions for gaze following and/or head-turn experiments. The model participants are from Rossel Island and are shown sitting on the low stool used for testing; during an experiment they would typically be sitting in the center of the tent, facing the screen. Right: the tent is tucked into a corner of the room with significant control over available light and includes separate spaces for the experimenter and the participants.

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

Foushee) while a video camera centered on the child's face allowed the experimenter to monitor the participant's looking behaviors throughout the session. Although we cannot be certain, it is likely that these conditions directly contributed to good data quality – in unpublished work we are aware of, Scaff and colleagues also attempted to measure word comprehension and audio-visual matching using looking-while-listening procedures in the same Tsimane' site described above, but piloting did not show encouraging results and data collection was abandoned. Why? The team had tried to collect the data outdoors, typically by having the mother and child sit on a mat on the ground in the yard or veranda of the participant's or other's house and placing a computer about 60 cm away from the infant. However, infants' direction of gaze on the screen was difficult to reliably annotate due to unreliable lighting, and children appeared to sometimes be distracted by other peoples' visible nearby activities.

Taking a step back, we extract two primary takeaways from this body of work. First, the combination of spontaneous production data with controlled elicitation or experimental measures of comprehension is enormously fruitful in examining rare (or highly specific) phenomena, identifying robust patterns in learning, and testing causal hypotheses (e.g., passives in K'iche' and Sesotho, specific morphological inflections in Yucatec Maya and Quechua). Second, the context of data collection influences the perceived frequency of these phenomena and, consequently, how often we believe children encounter them in their input (e.g., pointing in many of the communities mentioned, directed speech in Yucatec Maya, Tseltal, and Tsimane'; vocabulary knowledge in Yucatec Maya, Tseltal, Awadhi, and many others). Input frequency constrains what we entertain as internal mechanisms for language learning, which means that we must attend closely to how our methods influence our results. The example of input frequency is just one instance of how developmental theory can benefit from models of children's experiences that are grounded in real-world observation. More generally, we view approaches that combine observational and experimental methods as valuable for opening up opportunities for (or indeed, often demanding) some amount of interdisciplinary collaboration – thereby increasing the audience, interpretability, and robustness of the results.

Challenges in the study of language acquisition in rural settings

The preceding summaries highlight the importance of insights gained by studies combining observational and experimental approaches in rural populations for a range of linguistic phenomena. They also shed light on some of the DIFFICULTIES relating to this type of work. We now lay some of these challenges out more clearly, beginning with researcher responsibilities to participating communities.

Researchers have numerous responsibilities to their participating communities, but among them, capacity building remains one of the more complex issues, and one that is often difficult to approach. Capacity building involves researcher efforts to build up a foundation of local community expertise in research to integrate local perspectives in the work. Capacity building can include, for example, helping members of the participating community to develop expertise in language description or behavioral analysis. However, such expertise may not always be a priority for the communities. For instance, the Tsimane' community has ongoing partnerships with anthropologists who have been working with them for over two decades (see, e.g., Gurven, Stieglitz, Trumble, Blackwell, Beheim, Davis, Hooper & Kaplan, 2017). Some of the anthropologists set up a fund to

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

sponsor young Tsimane' adults' university degrees (One Pencil, 2022). These students primarily go on to study agriculture, with medicine coming in second, and only recently one student started a law degree. If these numbers are any indicator, community members are at present more urgently concerned with issues of food, economy, health, and law than theories of language and language acquisition. Thus any research that imposes this capacity-building goal under the guise of participatory research could be viewed as highly selective (and perhaps not highly relevant) from the perspective of community members.

Even when researchers do find community members who are interested in language documentation or linguistics, they must still closely attend to the power dynamics of the researchers' work: researchers may be wealthier than many community members and may look and speak in ways that corroborate local power structures (i.e., speaking the colonizing language). Many of the studies reported in Table 1 rely on native speakers for fundamental elements of research, such as explaining experimental procedures, translating coding categories, adding annotations, and making methodological decisions related to stimuli and experimental design. Many of these roles qualify these native speakers as coauthors, according to current authorship definitions (e.g., CReDiT, Holcombe, 2019). However, in developmental psychology, authorship is instead typically decided in terms of intellectual contributions, such as writing and analyses. Even when the decision is made to include a community member as co-author, one difficulty is ensuring that the co-author can authorize the manuscript at each stage of submission. This practice may be impractical or impossible for rural communities that do not have regular telecommunications access; the fast turnaround times required by academic journals worsen this issue. Researchers, especially those in the US and Canada, inhabit a professional "publish or perish" system, which translates to a desired output of multiple journal articles per year, and grants that fund 2- to 5-year projects. This rhythm for publication and funding does not support the development of strong and sustainable researcher-community bonds.

These issues also affect collaboration across researchers based in different nations. That is, many rural communities reside in low- or middle-income countries, but the researchers who have sufficient funding to undertake the scientific work reside elsewhere, in richer countries. This state of affairs leads to a system where the foreign, better-resourced researchers want to document acquisition in a rural community that is already in a relationship with local, less-well-resourced researchers who are not themselves members of the community of study, resulting in a three-tiered system of foreign researcher, local researcher, and local participant. For instance, one of us (Cristia) published a paper on the Tsimane' (Cristia et al., 2019) without involving any researchers located in Bolivia, despite the fact that doing so could have had a host of benefits, including the dissemination of their results in the local media. From the perspective of foreign researchers, the work of local researchers may be hard to discover if the local researchers publish primarily in a language that is not English or in journals that are infrequently read in the researchers' home countries. Collaboration between foreign and local researchers is often complicated by a history of foreign researchers extracting the data from study communities without involving or recognizing relevant local researchers. As a result, local researchers may actively avoid collaborations with foreigners – for example, by publishing in local media or in the language that is being studied/a local major language, but not in English. These moves strengthen collaboration with other local researchers and keep the data in low supply to those outside of that local network, but also lead to the relatively low visibility of their work to those outside those local circles

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

(Dutra, 2021). This exchange results in foreigners continuing to bypass local researchers, culminating in a negative feedback loop.

As a result of having limited availability of expert collaborators in the particular topics we work on, it is sometimes hard to develop stimuli, paradigms, and even coding categories for observational research that are relevant for the local culture and its values. Instead, foreign researchers impose paradigms and categories that are culturally meaningful from their perspective. For some research questions, linguistic annotation done by foreigners may be sufficient, and we highlight recent work that is turning to novel methods to annotate large amounts of naturalistic linguistic data at scale, including relying on citizen scientists (Cychosz et al., 2021a). However, such approaches cannot replace informed annotation relying on local knowledge. Researchers sometimes turn to the families themselves in order to ensure quality of the annotation. For instance, Demuth (1989) worked side by side with mothers and grandmothers for her Sesotho transcripts. The Tseltal and Yélî Dnye data (Casillas et al., 2020, 2021) were generated with community members that were familiar with all the recorded families, making it possible to identify individual family members in the long-form recordings. These methods of annotation – local expertise and large-scale – can be used in tandem to develop large, naturalistic, and informative datasets.

In our experience, also, some information that is often assumed as trivially easy to collect in certain contexts, like participant date of birth, may require some effort to establish. For infants in the Tsimane', Tseltal, and Yélî Dnye contexts, establishing exact date of birth has sometimes required triangulation between multiple data sources, including querying date information across families and asking the order in which several children in the village had been born. Even official documents, such as birth certificates, may not contain correct information and may instead be based on an official's guess of the child's age.

In many of these rural contexts, it may also be hard to glean information comparable to "typically developing" in urban communities. Urban communities can have easier access to healthcare facilities and screenings, with births more often taking place in well-established hospitals instead of rural clinics or the home. With hospital births, prematurity is noted and children are more likely to undergo medical visits to detect issues with vision and hearing, among other things. In many rural communities, however, access to medical services can be more limited owing to geographical distance or financial hardship (though this can also be the case in some urban samples). In these cases, it is up to the researcher to establish if a child is typically developing, which may be far from trivial. Cychosz (2021), for instance, attempted to measure children's hearing thresholds using the typical procedure whereby one asks the child to raise their hand when they hear tones of varying frequency; but some children would not raise their hand at all, despite no other evidence of hearing loss (i.e., oral speech delay), possibly because the children were afraid to make a mistake. It is also difficult to carry out standardized procedures, such as hearing or vision thresholds, in noisy or distracting environments where it is difficult to determine if lack of an effect is due to higher sensory thresholds or simply other distracting elements.

In some of the work mentioned here, existing ethnographic description was essential to make hypotheses, anticipate required methodological changes, and create experimental stimuli. Meaningful expectations regarding differences in children's language experiences, including local caregiving practices, housing and family arrangements, and daily routines, are invaluable for designing suitable observational and experimental studies in under-researched contexts. However, such information is not always available to

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

developmental psychologists. Even when it is available, it may sometimes not reflect the changing circumstances of the community (e.g., if the ethnography is more than a few years old) – recall the example of Padilla-Iglesias, Shneidman, and colleagues returning to the Yucatec Mayan community just six years after initial data collection to examine how market integration in the intervening years impacted language socialization practices. Rather than attempting to conduct ethnographic work without training or relying on simplified observation scales or surveys, we encourage researchers to reach out and collaborate with experts in ethnographic methods. We have often found that there is not a great deal of ethnographic information that these developmental researchers can build on when beginning work in an understudied population, including information about caregiving behaviors and expectations regarding children's development. Since we know that these details vary a great deal across populations, we cannot simply assume that they are equivalent to populations that are more frequently studied.

A great deal of experimental work involves choosing words and phrases (both visual referent and word form) or generating novel items that do not exist in the language. To this end, researchers need basic linguistic descriptions of the language, including what is grammatical, default word order, as well as frequency information for phonemes, phonotactic patterns, morphemes, words, and phrases. In major languages and cultural groups, databases like Open Subtitles (Lison & Tiedemann, 2016) can serve this role, providing reliable information on lexical statistics such as type or token frequency or co-occurrence patterns in as many as 62 languages. But these statistics primarily reflect adult-directed language and are likely to differ across registers. Those needing age of acquisition data for words can try using the Wordbank database to examine parental reports of word learning from 29 languages using the MacArthur-Bates Communicative Development Inventory and its adaptations (Frank, Braginsky, Yurovsky & Marchman, 2017), though we note that the included languages are primarily used by urban populations and that the 29 languages do not include any of those in the studies reviewed here. Furthermore, multilingualism is the rule, not the exception, for most children. So even if statistics or age of acquisition norms exist for a given language, they could differ greatly depending on the extent of language contact or type of multilingualism in the child's community. Considering that there are many large-scale languages (with millions of speakers) that do not have psychometric instruments for child language testing, it should not come as a surprise that such resources will not be available when working with rural populations, which sometimes speak languages used by only a few thousand people. Creating such instruments entails a long process of piloting in the field, which often results in uncontrolled cross-item variability (e.g., Cristia & Casillas, 2022; Pye & Quixtan Poz, 1988).

Even in the case of better-documented languages, such as Tseltal summarized above, local details and language change are difficult to account for. The nominal plural marker – *etik,* for example, is described as optional (that is, nouns not marked for number can refer to singular entities or pluralities), but is often used obligatorily to mark plurality, possibly due to transfer from Spanish. Similarly, in some, but not all Quechua-II language varieties, the nominal plural marker *-kuna* has been replaced with Spanish *-s*, but the extent of the replacement varies by speaker and Quechua language dominance. However, when there is detailed linguistic information available, language development research can benefit from it, as demonstrated by language-specific and cross-linguistic comparative work on Mayan language acquisition (e.g., Pye, 2017).

For languages without available linguistic descriptions, language development researchers may be tempted to rely on notions from basic linguistics courses, e.g., that

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

there will be a clear distinction between nouns and verbs. But such universals are challenged (Evans & Levinson, 2009). Instead, researchers should work with documentary linguist collaborators to generate a high-quality initial description and learn how to mine reference grammars and grammatical sketches for important information about the language's structure.

Some aspects of work in rural communities lead to smaller sample sizes than would be acceptable in urban settings where, e.g., the current sample size recommendation is over 40 for looking-time based studies with infants under 1;6 (Oakes, 2017). For example, researchers engaging in deep ethnographic, linguistic, and/or longitudinal data collection will likely be unable to track more than 2–4 children's development (e.g., Brown, 2011; Demuth, 1992; Pye, 1980). This kind of deep descriptive work is essential for understanding development in an understudied context, and so expectations about sample size must be balanced against the quality and novel contribution of the resulting analyses. Small population sizes and significant distances between test sites may also hamper data collection – ongoing habituation-based speech perception experiments with Yélî Dnye-learning infants (Casillas, personal communication) require a total of three or four visits to the island to accumulate two samples of 40 infants between 0;10 and 1;2. On each visit the researcher visits multiple villages with her portable lab (~20 kilos) to maximize data collection. Each village visit requires prior negotiation with local leadership and identifying a home in which to host the lab. The distance between villages is such that the researcher must plan ahead to bring a barrel of diesel to the island so that she can travel with the equipment by dinghy. So even with great effort and resources, achieving "typical" sample sizes in small-scale and rural contexts represents an enormous hurdle. On the other hand, low sample sizes mean that it is hard to statistically capture the degree of individual and group variation expected given current theories and findings (Kline, Shamsudheen & Broesch, 2018). It is therefore worthwhile to try and adapt methods for achieving larger samples when possible.

Last but not least, many rural communities live in some degree of connection with a mainstream economic system, in which case there will be contact with other languages. As mentioned above for the case of children acquiring Quechua, the author made a point to measure the degree of individual children's exposure to Quechua and Spanish to model individual differences in the children's phonetic production and phonological processing (Cychosz, 2022a, 2022b). Language contact and multilingualism also matter because caregivers may use both languages and hold various beliefs about each of them. This was salient in Padilla-Iglesias et al. (2021)'s study summarized above, where caregivers reported that Spanish needed to be taught, whereas Yucatec could be learned by exposure.

Desiderata

In this section we briefly discuss specific practices in the fields of developmental psychology and developmental linguistics that would improve the outlook of research on diverse rural contexts, including steps in the manuscript review and publication process, concepts of comparison, and amplifying the voices and contributions of researchers from underrepresented groups (such as variable academic backgrounds, ranks, and domains of interest).

For more work on language development among children in rural contexts to enter the public sphere, we need to change how it undergoes peer review. Often, work conducted in small-scale societies is measured using the same ruler as research in industrialized

https://doi.org/10.1017/S0305000922000617 Published online by Cambridge University Press

societies. But given all the challenges described here, such standards often relegate research conducted in rural or small-scale societies to less frequently read academic venues. So editors and reviewers of work conducted in rural, or similarly challenging understudied settings, need to understand that the level of study-supporting information and resources available is different. At present, we must be more accepting of methodological variation and sample sizes. With appropriate caveats in the interpretation of the research, which do not oversell the conclusion, such work can enter the public eye.

It is also inappropriate, in our view, for editors or reviewers to request "control groups", in which the same method is used across urban and rural sites. When a paper on English-learning, North American infants is published, authors typically assume universal generalizability in their findings, and are not requested to generalize their approach to a "control group" from a rural community. This is despite the fact that early experiences alter even newborns' behavior (Dixon, Tronick, Keefer & Brazelton, 1982) – overrepresented populations, like underrepresented ones, are brought into a specific cultural and linguistic context from the very start, and neither group REQUIRES a cultural comparison group to investigate how enculturation processes begin. As long as authors do not make comparative statements and generalize their theoretical and interpretive scope appropriately, no second or comparative group should be needed.

We have here tried to avoid comparative statements, because although they can seem informative to readers accustomed to Anglophone North American samples, they may lead us to think of the latter as "standard", the "norm", against which all other populations need to be measured (Singh et al., 2021). This leads to "othering", or presenting some populations as different – a mistake of which we have certainly been guilty (Cristia et al., 2019). Thus, we encourage readers of literature on rural populations to avoid viewing this literature through an "othering" lens – even if the authors of the work themselves have made that error.

Finally, we would be excited to see greater promotion of researcher diversity in the field, with greater opportunities for training and career development for members of under-represented groups (e.g., Aravena-Bravo et al., 2023). Although some rural communities may not have individuals who would be willing and able to benefit from these opportunities, many others might. This is saliently underscored by the work of Mayan specialists; for instance, Lourdes de León, at the Centro de Investigaciones y Estudios Superiores en Antropología Social (CIESAS), along with others in the Universidad Nacional Autonoma de Mexico, encourages indigenous people to apply and enroll in Masters and PhD programs at the university. These students are promoted as "actors, subjects, and authors" to "document through the lens of their own language and culture what it means to be born and grow in Mesoamerica and the Andes" (our translation; de León, 2019).

Finally, we could increase researcher diversity in our field by encouraging heterogeneity among collaborators, including those with different academic backgrounds and those at different career levels. For instance, authors and journal editors could (1) promote the inclusion of at least one member from an underrepresented group on research papers; (2) foster multilingualism as lingua franca, and therefore read and accept writings in different languages apart from English (a good starting point would be to publish papers with two or three abstracts written in different languages). Above all we encourage journals to offer free English language copywriting services, as English language skill is a serious bottleneck to increasing researcher diversity. Further, governmental and non-governmental funding should aim to (3) increase collaborations among South-South or Triangular Cooperation, (4) make open science infrastructures (virtual and physical)

more accessible to researchers in countries where internet access and/or data-relevant electronic devices is restricted (for example, by publishing open-access), (5) engage and return to the communities that researchers have worked with (e.g. giving full access to research data, including that which may help with language maintenance/revitalization); and, (6) promote career development for members of underrepresented groups (Aravena-Bravo et al., 2023), among many other initiatives. Many of these aims are part of bigger initiatives like the Helsinki Initiative on Multilingualism in Scholarly Communication (2019), the Open Science Recommendations made by UNESCO (2021), and even in the policy recommendations made by the United Nations development system to deliver on the 2030 Agenda (UN, Secretary-General, 2017).

On the scale of what individual researchers can do, we point again to the work of our colleagues at CIESAS and the Universidad Nacional Autonoma de Mexico, where indigenous researchers are being trained and promoted in the process of knowledge production about their own communities (de León, 2019). Along these same lines, we encourage our colleagues to provide and enhance initiatives that foster cooperation in a fair and equitable manner.

Conclusion

In this paper, we have attempted to summarize a growing body of work combining experimental and observational approaches to studying language learning in young children growing up in rural communities. The work cited here has contributed to our understanding of language development in a handful of cultures and languages, and we hope the current summary helps inspire more such work. To this end, we provide information on the challenges involved in this research, and how the field as a whole can help overcome them.

Acknowledgements. The authors wish to thank the communities who hosted this work and engaged in research on/in Rossel Island, Papua New Guinea (MCa), Tenejapa, Mexico (MCa, RF), southern Bolivia (MCy), and the Bolivian lowlands (CS, AC). PAB also thanks those within the LatiNLAR community for providing thoughtful comments and sharing their valuable expertise on this topic. The authors acknowledge financial support from Agence Nationale de la Recherche (ANR-17-CE28-0007, ANR-16-DATA-0004 ACLEW, ANR-17-EURE-0017), the ERC (ExELang, 101001095), and a J. S. McDonnell Foundation Understanding Human Cognition Scholar award (AC), the Society for Research in Child Development and Center for Effective Global Action at UC Berkeley (RF), and the Oswalt Documenting Endangered Languages Fund, Center for Latin American Studies at UC Berkeley, and NIH-NIDCD award F32DC019539 (MCy). Additional thanks for insights and comments along the way to Linda Abarbanell, Peggy Li, Thomas Castelain, Shanley Allen, Andrea Taverna, Nicola Yuill, Paul Ibbotson, and Katherine Demuth.

Supplementary Materials. To view supplementary material for this article, please visit http://doi.org/10.1017/S0305000922000617.

References

Abels, M. (2020). Triadic interaction and gestural communication: Hierarchical and child-centered interactions of rural and urban Gujarati (Indian) caregivers and 9-month-old infants. *Developmental Psychology*, **56**(10), 1817.

Aravena-Bravo, P., **Garcia, R.**, **Kotera, H.**, **Laranjo, R.**, **Nicolas, R. K.**, and others (2023) Towards Diversifying Early Language Development Research: Lessons from /L+/ 2021. https://osf.io/a3dy8/

Brown, P. (2011). The cultural organization of attention. In *The handbook of language socialization* (pp. 29–55). Wiley-Blackwell.

Brownell, R., & **Academic Therapy Publications**. (2000). *Expressive one-word picture vocabulary test.*

Bunce, J., **Soderstrom, M.**, **Bergelson, E.**, **Rosemberg, C.**, **Stein, A.**, **Migdalek, M.**, & **Casillas, M.** (2022). *A cross-cultural examination of young children's everyday language experiences*. PsyArxiv.

Callaghan, T., **Moll, H.**, **Rakoczy, H.**, **Warneken, F.**, **Liszkowski, U.**, **Behne, T.**, & **Tomasello, M.** (2011). Early social cognition in three cultural contexts. *Monographs of the Society for Research in Child Development*, **76**(2), vii–142. https://doi.org/10.1111/j.1540-5834.2011.00603.x

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.

Childers, J. B., **Vaughan, J.**, & **Burquest, D. A.** (2007). Joint attention and word learning in Ngas-speaking toddlers in Nigeria. *Journal of Child Language*, **34**(2), 199–225.

Cristia, A. (2022). A systematic review suggests marked differences in the prevalence of infant-directed vocalization across groups of populations. *Developmental Science, XX.*

Cristia, A., & **Casillas, M.** (2022). Non-word repetition in children learning Yélî Dnye. *Language Development Research*, **2**(1), 69–104.

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.

Cristia, A., **Farabolini, G.**, **Scaff, C.**, **Havron, N.**, & **Stieglitz, J.** (2020). Infant-directed input and literacy effects on phonological processing: Non-word repetition scores among the Tsimane'. *PloS one*, **15**(9), e0237702.

Cristia, A., **Foushee, R.**, **Aravena-Bravo, P.**, **Cychosz, M.**, **Scaff, C.**, & **Casillas, M.** (2022). Online supplementary materials to Combining observational and experimental approaches to the development of language and communication in rural samples: Opportunities and challenges.

Cychosz, M. (2021). The coarticulation-duration relationship in early Quechua speech. *Journal of Phonetics*, **87**, 101052.

Cychosz, M. (2022a). Language exposure and the development of phonological working memory. Talk given at the 96th annual meeting of the Linguistic Society of America.

Cychosz, M. (2022b). Language exposure predicts children's phonetic patterning: Evidence from language shift. *Language*, **98**(3), XX-XX.

Cychosz, M., **Cristia, A.**, **Bergelson, E.**, **Casillas, M.**, **Baudet, G.**, **Warlaumont, A. S.**, **Scaff, C.**, **Yankowitz, L.**, & **Seidl, A.** (2021a). Vocal development in a large-scale, crosslinguistic corpus. *Developmental Science*, **24**(5), e13090.

Cychosz, M., **Villanueva, A.**, & **Weisleder, A.** (2021b). Efficient estimation of children's language exposure in two bilingual communities. *Journal of Speech, Language and Hearing Research*, **64**(10), 3843–3866.

de León, L. (2019). Nacer y crecer en Mesoamérica y Los Andes: Teorías parentales y prácticas de crianza infantil contemporáneas. México: CIESAS, Centro de Investigaciones y Estudios Superiores en Antropología Social.

Demuth, K. (1989). Maturation and the acquisition of the Sesotho passive. *Language* **65**(1), 56–80.

Demuth, K. (1992). The acquisition of Sesotho. In *The crosslinguistic study of language acquisition* (pp. 557–638). Psychology Press.

Demuth, K., **Machobane, M.**, & **Moloi, F.** (2003). Learning animacy hierarchy effects in Bantu double object applicative constructions. In *Linguistic Typology and Representation of African Languages* (pp. 23–33). Africa World Press.

Demuth, K., **Machobane, M.**, **Moloi, F.**, & **Odato, C.** (2005). Learning animacy hierarchy effects in Sesotho double object applicatives. *Language*, **81**(2), 421–447.

Demuth, K., **Moloi, F.**, & **Machobane, M.** (2010). 3-Year-olds' comprehension, production, and generalization of Sesotho passives. *Cognition*, **115**(2), 238–251.

Dixon, S., **Tronick, E.**, **Keefer, C.**, & **Brazelton, T. B.** (1982). Perinatal circumstances and newborn outcome among the Gusii of Kenya: Assessment of risk. *Infant Behavior and Development*, **5**(1), 11–32.

Dunn, L. M., & **Dunn, D. M.** (2007). *PPVT-4: Peabody Picture Vocabulary Test*. Pearson Assessments.

Dutra, N. B. (2021). Commentary on Apicella, Norenzayan & Henrich (2020): Who is going to run the global laboratory of the future? *Evolution and Human Behavior*, **42**(3), 271–273. https://doi.org/10.1016/j.evolhumbehav.2021.04.003

Espinosa Ochoa, M. R. (2017). Un acercamiento a la construcción de la circunclisis en el maya yucateco infantil. *Estudios de Lingüística Aplicada*, **66**, 41. https://doi.org/10.22201/enallt.01852647p.2017.66.682

Espinosa Ochoa, M. R. (2022). Germination, early development, and creativity in the acquisition of the Yucatec Maya deictic system. *Journal of Child Language*, **49**(1), 24–37. https://doi.org/10.1017/S0305000920000689

Evans, N., & **Levinson, S. C.** (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, **32**(5), 429–448.

Fibla Reixachs, L. (2021). *Relating language input to language processes early in development: using the early language processing task in UK and India* (Doctoral dissertation, University of East Anglia).

Foushee, R., & **Srinivasan, M.** (2022). Evidence for early lexical knowledge in infants primarily exposed to overheard speech. https://osf.io/kcza9/?view_only=64ed45a0da2c429eaf5b9c2f0c6f799f

Frank, M. C., **Braginsky, M.**, **Yurovsky, D.**, & **Marchman, V. A.** (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, **44**(3), 677–694. doi:10.1017/S0305000916000209

Gilkerson, J., & **Richards, J. A.** (2020). *A Guide to Understanding the Design and Purpose of the LENA® System*. LENA Foundation: Boulder, CO. Available from https://www.lena.org/wp-content/uploads/2020/07/LTR-12_How_LENA_Works.pdf.

Gurven, M., **Stieglitz, J.**, **Trumble, B.**, **Blackwell, A. D.**, **Beheim, B.**, **Davis, H.**, **Hooper, P.**, & **Kaplan, H.** (2017). The Tsimane health and life history project: integrating anthropology and biomedicine. *Evolutionary Anthropology: Issues, News, and Reviews*, **26**(2), 54–73. https://doi.org/10.1002/evan.21515.

Holcombe, A. O. (2019). Contributorship, not authorship: Use CRediT to indicate who did what. *Publications*, **7**(3), 48.

Kidd, E., & **Garcia, R.** (2022). How diverse is child language acquisition research?*First Language*, XX.

Kline, M. A., **Shamsudheen, R.**, & **Broesch, T.** (2018). Variation is the universal: Making cultural evolution work in developmental psychology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **373**(1743), 20170059.

Lison, P., & **Tiedemann, J.** (2016). OpenSubtitles2016: Extracting Large Parallel Corpora from Movie and TV Subtitles. In Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016). Database available from https://opus.nlpl.eu/OpenSubtitles-v2018.php

Liszkowski, U. (2011). Three lines in the emergence of prelinguistic communication and social cognition. *Journal of cognitive education and psychology*, **10**(1), 32–43.

Liszkowski, U., **Brown, P.**, **Callaghan, T.**, **Takada, A.**, & **de Vos, C.** (2012). A Prelinguistic Gestural Universal of Human Communication. *Cognitive Science*, **36**(4), 698–713. https://doi.org/10.1111/j.1551-6709.2011.01228.x

Loukatou, G., **Scaff, C.**, **Demuth, K.**, **Cristia, A.**, & **Havron, N.** (2021). Child-directed and overheard input from different speakers in two distinct cultures. *Journal of Child Language*, **49**(6), 1173–1192.

Oakes, L. M. (2017). Sample size, statistical power, and false conclusions in infant looking-time research. *Infancy*, **22**(4), 436–469.

Ochs, E., & **Kremer-Sadl, T.** (2020). Ethical blind spots in ethnographic and developmental approaches to the language gap debate. *Langage et Societe*, (2), 39–67.

O'Neill, D. (2008). Social organization: An overview of how we create and maintain social groups. Online tutorial, available from https://www2.palomar.edu/anthro/status/default.htm

One Pencil (2022). The Tsimane Scholarship Fund, accessible from https://www.onepencilproject.org/tsimane-scholarship-fund

Padilla-Iglesias, C., **Woodward, A. L.**, **Goldin-Meadow, S.**, & **Shneidman, L. A.** (2021). Changing language input following market integration in a Yucatec Mayan community. *PloS one*, **16**(6), e0252926.

Pye, C. (1980). *The acquisition of grammatical morphemes in Quiche Mayan*. PhD Thesis, University of Pittsburgh.

Pye, C. (2017). *The comparative method of language acquisition research*. The University of Chicago Press.

Pye, C., & **Quixtan Poz, P.** (1988). Precocious passives (and antipassives) in Quiché Mayan. *Papers and Reports on Child Language Development*, **27**, 71–80.

Salomo, D., & **Liszkowski, U.** (2013). Sociocultural settings influence the emergence of prelinguistic deictic gestures. *Child Development*, **84**(4), 1296–1307.

Scaff, C., **Casillas, M.**, **Stieglitz, J.**, & **Cristia, A.** (2022). Verbal input among the Tsimane'. *PsyArXiV* preprint. https://doi.org/10.31234/osf.io/mt6nz

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.

Singh, L., **Cristia, A.**, **Karasik, L. B.**, & **Oakes, L.** (2021). *Why are we still so WEIRD? Barriers and bridges towards a diversified science of early development*. PsyArxiv. https://psyarxiv.com/hgukc/

Sugden, N. A., & **Moulson, M. C.** (2015). Recruitment strategies should not be randomly selected: empirically improving recruitment success and diversity in developmental psychology research. *Frontiers in Psychology*, **6**, 523.

UN, Secretary-General (2017). Repositioning the United Nations development system to deliver on the 2030 Agenda: ensuring a better future for all: report of the Secretary-General. https://digitallibrary.un.org/record/1298793?ln=en

United Nations Educational, Scientific and Cultural Organization (2021). *UNESCO Recommendation on Open Science* (Programme and Meeting Document SC-2021/SANS COTE). Retrieved from the United Nations Educational, Scientific, and Cultural Organization website: https://unesdoc.unesco.org/ark:/48223/pf0000379949.locale=en

Cite this article: Cristia A., Foushee R., Aravena-Bravo P., Cychosz M., Scaff C., & Casillas M. (2023). Combining observational and experimental approaches to the development of language and communication in rural samples: Opportunities and challenges. *Journal of Child Language* 1–23, https://doi.org/10.1017/S0305000922000617