Early language experience in a Papuan village

|  |
| --- |
| Marisa Casillas1, Penelope Brown1, & Stephen C. Levinson1 |
| 1 Max Planck Institute for Psycholinguistics |
|  |

# Author note

Correspondence concerning this article should be addressed to Marisa Casillas, P.O. Box 310, 6500 AH Nijmegen, The Netherlands. E-mail: [Marisa.Casillas@mpi.nl](mailto:Marisa.Casillas@mpi.nl)

Abstract

To be completed later.

*Keywords:* Child-directed speech, linguistic input, non-WEIRD, vocal maturity, turn taking, interaction, Papuan

Word count: XXXXX (XXXX not including references)

Early language experience in a Papuan village

# Introduction

In their first five years of life, children hear an extraordinary amount of language in a wide variety of interactional contexts. Tracking the distribution and characteristics of this linguistic input over the day, across age, and between children is a difficult task. Until recently, developmental language science has relied on short video recordings of caregiver-child interaction, at home or in the lab, to get a grasp on what kinds of language children typically hear. This has been a fruitful approach in teasing out individual and group-based differences in interactional style (REFS). However, short recordings are limited in their insight because they represent only a small slice of the child’s language abilities and experiences (REFS).

Improved recording hardware and advances in speech technology have recently allowed us to use daylong recordings to get a peek into children’s broader language landscapes. Daylong recordings are made with a device, usually positioned on the target child’s chest, while that child freely navigates their social environment for most of a waking day (REFS). This style of audio recording has allowed researchers to track children’s verbal language use across a range of activity and interlocutor contexts, yielding more representative and generalizable measures of their language environments (REFS). While daylong recording collections are typically too large for comprehensive transcription and annotation, a combination of automated tools, (REFS) sampling techniques (REFS), and standardized annotation approaches (REFS) can lead to rich, but efficiently-gained glimpses into the at-home language environment. However, properly collecting, processing, and archiving daylong data is not easily achieved and may not be well suited for a range of research questions (REFS). At time of writing, there are few options for capturing visual information across the day (but see REFS), limiting this method primarily to acoustic phenomena (REFS).

Daylong recording methods are still relatively new, and their reliability and predictive value for language development have not yet been fully established. For example, one collection of recordings made in the US Northwest suggests that there is so much variability across activities and days in basic talk characteristics (e.g., how much speech comes from what types of speakers) that researchers need several days of recordings before they can expect their input estimates to stabilize (Anderson & Fausey, in prep). Even if one can achieve a reliable estimate of a language environment measure (e.g., overhearable adult words per hour), how and why that estimate relates to deeper factors shaping the learning situation, including caregiving ideologies and language outcomes, is often indirect at best. Relatedly, meaningful differences between individual children may be minimized when averaging across the entirety of the day’s high and low moments; it may well be that a few key interactions throughout the day provide sharper resolution on individual and group-based differences compared to whole-day averages.

Two recent studies have directly investigated the effect of recording duration on caregiver speech, finding that short recordings display denser, somewhat different input than what is found in longer recordings (REFS). Bergelson and colleagues (REFS) analyzed the contexts of noun use encountered by 44 6- and 7-month-old children in the US in both hour-long at-home videos and daylong recordings. The hour-long video differed from the day-long average in the density of noun input, the utterances in which nouns were embedded (e.g., questions vs. declaratives), and the actual nouns used. When the authors compared the hour-long video to the ‘peak’ hour of talk in the day, they found a similar density of noun input across the recording types, but the other differences remained. This difference in input density also appears in short recordings of varying length with US children (REFS). Importantly, however, the speech characteristics heard by individual children in a sample correlates across recording types (i.e., their relative rank for a given measure is stable; REFS). Based on these findings, one could infer that at-home short recordings are influenced by some (but not all) of the same underlying factors that drive language patterns during daylong recordings (e.g., caregiver ideologies about child development, child responsiveness, household composition).

Studies of children growing up in two indigenous Mayan communities of Southern Mexico (Tseltal and Yucatec Mayan) suggest that short and long recordings may yield substantial differences in how the speech environment is characterized (REFS). Previous studies on these communities have tended to use ethnographic and microanalytic analyses of short interactions to examine the character of children’s speech environments. They have found that caregivers shape infants’ and young children’s worlds such that the children learn to attend to what is going on around them rather than expecting to be the center of attention (REFS). Consistent with this goal, direct talk to infants, particularly from adults, is rare until children themselves begin to elicit responses from others (REFS). Because young children are often cared for by older siblings and cousins, a substantial portion of talk to young children was also expected to come from other children (REFS). Similar observations have been reported for multiple other distinct (but ethnolinguistically related) communities in the region (REFS). Following up on this ethnographic work, Shneidman (REFS) used short videos of interaction to conduct a quantitative, longitudinal study of the speech young Yucatec children heard. She found that interactional patterns aligned well with observations in previous work in that community: infants were rarely spoken to at first, but their language input increased enormously with age, mostly due to an influx of speech from other children (REFS). However, when Casillas and colleagues (REFS) used daylong recordings with a Tseltal Mayan community, where a similar caregiver interactional style has been described previously on the basis of short videos, the pattern of findings diverged from expectations. In brief, they found that infants and young children were indeed spoken to rarely, but that there was no increase in speech input with age and the majority of speech came from adult women, even when children were old enough to independently follow their older siblings and cousins around the house. These divergent results betweeh daylong and short video recordings don’t imply that the latter is wrong, only that it is not representative with respect to the child’s language experience over an entire day.

These findings do raise an important issue faced by developmental psychology as it continues to expand the study of child language to more diverse speech communities: when researchers are not members of the community they are studying, it is difficult to know a priori what is typical, representative, or meaningful in children’s language experience. By observing as much speech as possible in a context as ecologically valid as possible and by sampling, annotating, and analyzing the data on the basis of the most established development measures we have, researchers using daylong recordings might hope to approach this issue without first needing to conduct deep ethnographic studies in the community on caregiving practices and ideologies around language use and language development (REFS). When studying members of our own cultural group, we can bridge between simple, observable behaviors and rich interpretations of, thereby expanding our explanatory model beyond the measures directly analyzed (e.g., why child-directed talk might relate to faster vocabulary development). We cannot hope to gain such enriched understandings cross-culturally without ethnographic work; and in the absence of such work we must accept that there may be a dissociation between how we have traditionally understood an operationalized language behavior (e.g., child-directed speech) and what drives the use and form of that behavior in a given community or interactional context (e.g., pedagogical concerns, entertainment of the caregiver, getting the child to assist). Until there are trained researchers working on this topic who were born and raised as members of these communities (what we should be trying to cultivate for the longer term) this is a quandary we will continue to face.

Pairing ethnographic work with broader-scope studies of children’s language environments may be the most fruitful way to ensure that their speech environments and speech development are captured well enough to propose and test meaningful theories cross-culturally. These two methods have complementary roles to play in exploring the landscape of at-home language, and neither should be taken to reflect the ‘true’ language input for a given child; after all, in the example of Tseltal above, many interactions with infants during the daylong recordings came during moments where visitors using a video camera, or even other community members, would not typically be invited (e.g., after the parent was roused by the child, who was waking from an afternoon sleep). If we want to encourage more work on small-scale and/or understudied language learning contexts, it will be important to continue establishing how different methods of measuring the input impact the conclusions that are likely to be made.

In this study we present analyses of daylong recordings from a small-scale indigenous community in which prior ethnographic work has painted a clear picture of early caregiver-child interaction, different from what we have seen in studies of Mayan caregivers: child-centric, face-to-face interaction from the first days of infancy on Rossel Island, Papua New Guinea (PNG). In what follows we will review the ethnographic work done with this community previously, our methods for following up on these findings with daylong recordings, and a discussion of the differences that arose. This study was completed as part of a larger comparative project focusing on children’s speech environments and linguistic development at two sites: the Tseltal Mayan community mentioned above (Casillas et al., forthcoming) and this Rossel Island community. Therefore all methods for annotation and analysis in this study parallel those reported elsewhere for Tseltal Mayan children’s speech environments (Casillas, Brown, & Levinson, forthcoming).