

Dear Dr. Paradis,

We are pleased to send you our resubmission of “Early language experience in a Papuan community”. We appreciate your, Dr. Wagner, and the reviewers’ comments, and we find that addressing them has indeed improved the flow of the paper substantially. Below we provide a point-by-point summary of the changes we made, with each of our comments preceded by ‘>>’. All edited/added text in the resubmission is also highlighted in orange. We hope that you will also find the manuscript much improved and we look forward to your next round of comments.

Sincerely,  
The authors

#####

Dear Dr. Casillas:

JCL-02-20-023 entitled "Early language experience in a Papuan community" which you submitted to the Journal of Child Language, has been reviewed. The comments of the Action Editor (AE: Dr. Wagner) and the reviewer(s) are included at the bottom of this letter.

This manuscript reports on a carefully conducted study with an understudied population addressing some fundamental questions in cross-cultural language acquisition. As such, the reviewers, Dr. Wagner and I would very much like to see this paper published in this journal. That being said, the comments of Dr. Wagner and the reviewer(s) indicate that some major revisions are needed to this manuscript, mainly in terms of organization and writing. Regarding the supplemental materials, reviewer #2 correctly points out that some of this information needs to be in the main text for readers to follow. However, the supplemental materials are very long, so the choice of what to add to the main text needs to be judicious. Further on this point, you might consider shortening the supplemental materials in any case and post the full version on an open access website that readers can be referred to, e.g., osf.

>> As described more below, we have taken up this call for reorganization and refocusing of the text and ultimately have rewritten the majority of the Introduction and Discussion sections. All the materials for replicating this study and/or approach, including scripts for summarizing the data, running the analysis, and even compiling the manuscript are already stored on a private GitHub repository that we will make public with publication of the paper (it is referred to now as an “omitted URL” in the text). Other scripts, training materials, and manuals referred to throughout the text (all “omitted URL”s) are similarly already posted in public repositories on GitHub or OSF.

Therefore, I would like to invite you to respond to the comments of the AE and the reviewer(s) and revise your manuscript. We expect to receive your manuscript within 180 days.

[...]

Action Editor (Dr. Laura Wagner's) Comments to Author:

Action Editor

Comments to the Author:

Both of the reviewers were very enthusiastic about the core elements of this manuscript – both in terms of the non-WEIRD sample you investigated and in terms of the richness of your methodological approach. However, as Reviewer 1 very clearly lays out, the paper lacks focus and as a result, what seems to be the most important result here (namely, how language acquisition is supported in a non-WEIRD environment) gets a little bit lost. Indeed, the introduction is almost exclusively about the methods (Close vs Panoramic sampling) and the information about the PNG child-rearing environment isn't presented until the method section! Thus, a thorough revision of this paper is needed that takes into consideration the comments of both reviewers.

Reviewer 1 advocates for a complete restructuring of the paper that would make the non-WEIRD dimension of the paper the central focus. The methodological approach would then be dramatically shortened and re-cast as a useful way for investigating non-WEIRD cultures (as well as WEIRD ones). Reviewer 2 notes that it would be useful to give more historical context for how ethnographic investigations of non-WEIRD cultures have posed challenges to current (WEIRD-centric) theorizing.

Moreover, as Reviewer 1 notes, tightening up the focus of the paper will also make it easier to stay focused through the results section, since it will be easier to make clear as you go through to see which results are most critical and how they link to the core questions here.

>> Thank you for highlighting the organizational problems with our paper. We certainly don't mean for the main findings to be buried and we are glad to have the opportunity to right this issue. The AE will now find the Introduction and Discussion both heavily edited and restructured in a way similar to what was recommended by Reviewer 1. This includes abandoning mentions of close vs. panoramic approach trade-offs. We agree that the prior framing wasn't fit right to the study, and it's a topic we will think about addressing in a separate manuscript in the future. We hope the AE finds the paper improved in flow since the original submission.

Reviewer(s)' Comments to Author:

Reviewer: 1

Comments to the Author

This paper reports the results of an interesting study of the language-learning environment of 10 children in a Papuan community from birth to 3 years, and the timing of their early language milestones, using a daylong recording approach.

The research is clearly publishable in the journals: the method is sophisticated and well thought-out; the results are appropriately analysed; and the findings have potentially important implications for the field. However, the current version of the paper is overlong and somewhat lacking in focus – with way too much discussion of the difference between close and panoramic approaches and not enough discussion of what the authors' method and data actually tell us about the language learning process. In my view, the authors need to produce a shorter, more focused version of the paper which concentrates on what I take to be the key aims, namely: 1) To look at the language experience of children in a culture which is child-centric (like WEIRD cultures) but where daily life is organised around subsistence farming (like other non-WEIRD cultures), in order to see which of these factors is most important in determining the overall rate of child-directed speech across the day; 2) To look at children's language experience in peak turn-taking interactions to investigate the extent to which this reflects the child-centric nature of the culture, and 3) To look at the age at which the children reach key milestones in order to draw inferences about the extent to which these are affected by the above factors.

>> We are grateful to Reviewer 1 for not only identifying the organizational problem with our text but also for providing a helpful outline of points to hit in revision. The reviewer will find that, though the paper is not much shorter in its revised form (in part due to requests for more methodological information below), it is much more focused. The new Introduction now focuses on the main issue of variation in input rates and their proposed relation to language development, identifying caregiver ideologies and social-organizational factors as two major sources of variation, then mentioning the culture-economy confound in past work and pegging the current study as a critical third datapoint: a child-centric subsistence farming interactional setting. We then briefly explain how peak turn-taking vs. random baseline clips might help us gain different insights, make our predictions, and then jump into the details of the methods.

#### More specific comments

Introduction: The introduction currently reads like an extended discussion of the relative merits of different kinds of methodological approach, much of which is only tangentially relevant to the study that is being reported – and which, in my view, is rather too kind to the traditional approach, which has almost certainly led to an over-estimation of the amount of child-directed speech that children hear in WEIRD as well as non-WEIRD cultures. The authors need to come up with a shorter and more focused introduction that provides a rationale for the study that they have actually done (i.e. an introduction that focuses on a) whether different rates of child-directed speech reflect differences in child-rearing ideologies or differences in the demands of daily life in different cultures; b) whether different child-rearing ideologies result in different ways of interacting with young children during interactional peaks; and c) whether the kind of low rates of child-directed speech found in non-WEIRD cultures affect the age at which children reach different language milestones – and what implications this has for our understanding of the language learning process).

>> We have rewritten most of the Introduction, as described above, to follow a similar flow of information to what is suggested here.

Method: Although the conclusions they are able to draw are obviously limited by the fact that they are only looking at data from 10 children, by focusing so much on the differences between close and panoramic approaches, the authors actually undersell their method, which by sampling both at random and in a targeted way from daylong recordings, allows them both to come up with a realistic estimate of the amount of child-directed speech and other-directed speech that the children actually hear, but also to look at what happens in peak turn-taking interactions – and hence also provide a more in-depth analysis of the kind of child-directed speech the children hear in these interactions – though, unfortunately, the authors’ analysis of the peak turn-taking segments is pretty superficial – and hence does not allow them to say anything very meaningful about how the community’s child-centric ideology affects the nature of the child-directed speech that the children are hearing (see below).

>> We agree about the superficial nature of our current approach, but generating these 6.25 hours of annotated recording time took hundreds of hours spread over three trips to a remote site. We see the current results as giving a first peek into a hard-won dataset that will grow richer with further visits (i.e., further annotation and ethnographic grounding). That said, in this revised version of the manuscript we have highlighted patterns in who is talking to the child (in addition to how much talk the child is hearing) as a step beyond simple quantity-based analysis.

Results: The results section, like the introduction, would benefit from a more focused rewrite. With respect to the first section, the key finding is that the average amount of child-directed speech is just over 3 minutes per hour, which is similar to that reported in a comparable study of children in a Mayan community (despite the difference in ideology) and between a quarter and a third of what is reported in a comparable study of North American children (suggesting that this difference has more to do with the different demands of daily life in these cultures than with differences in ideology). This point would come across much more clearly if the authors presented a straightforward comparison of these 3 rates rather than just the rates in the two non-WEIRD contexts. I also think the authors muddy the waters by talking loosely about the children in their sample hearing ‘minimal’ child-directed speech. ‘Minimal’ is a loaded term that trades off the widespread assumption that children in WEIRD cultures hear hours and hours of child-directed speech every day – when the reality is that, if one looks at the numbers from studies using a similar daylong recording methodology, North American children appear to hear an average of about 100 minutes of child-directed speech per day and the children in this study hear an average of about 30 minutes per day. Sure, 30 is still a lot less than 100, but, when viewed against the more realistic 100-minute reference point, 30 minutes a day can hardly be described as ‘minimal’. The implication of the study is therefore not that children can learn a language despite hearing only a ‘minimal’ amount of child-directed speech (though it’s possible that they can, e.g. by learning through overhearing), but rather that a relatively small amount of child-directed speech per day may be sufficient to ensure that typically developing children reach developmental milestones according to roughly the same timetable as children who happen to be addressed considerably more often.

>> We have removed the word “minimal” because we agree with the problematic inferences possible in this context, as noted by the Reviewer above. However, we decided

*against* presenting the results as a direct 3-way comparison because the findings to-date on input rates from North American children are simply not directly comparable to the other two datasets at hand here (short, random and manually selected clips here and in Casillas et al. (2019) vs., e.g., in Bergelson et al., 2019: clips selected because of the presence of nearby adult speech). Along with the reorganization of the Introduction, the Discussion now focuses on the Tseltal-Rossel comparison as the critical disambiguating one, because we believe that is the only direct comparison of input rates licensed by existing published data.

With respect to the second section, the analysis of the peak turn-taking interactions seems pretty superficial to me. After all, it is hardly surprising that these interactions contain more child-directed speech than the random samples since they have been selected on precisely this basis. Surely what is required here is an analysis of the kind of speech directed to the child during these interactions (and, if possible, a comparison of what has been reported for Mayan and North American children, to understand how any similarities or differences are related to differences in childrearing ideology versus childrearing context).

>> We completely agree with the Reviewer that content features of the turn-taking speech will be fascinating to look at in future work. For now, as mentioned above, the current data represent a first peek into patterns of input in this community, and we have enriched this very basic measure with highlighted reporting about who is providing the child-directed speech. We also agree that it is obvious that TCDS would be higher in the turn-taking clips. In this revised version of the manuscript we better highlight the main point of this analysis: the change in rate and its relationship to time of day and child age is similar to what was found for Tseltal in Casillas and colleagues' (2019) paper, going against our idea that differences between these two contexts would be more apparent in the turn-taking clips.

With respect to the third section, while it is clearly important to establish that children within this culture reach language milestones at more or less the same age as children in WEIRD cultures, it is also important to recognise that the vocal maturity measures presented by the authors are relatively crude and don't tell us anything about the children's vocabulary range or the productivity or syntactic complexity of their multi-word speech. Although it may come as a surprise to some researchers in the field, I'm not sure it's particularly surprising that the ages at which children start to babble, to produce their first words and to start combining words are relatively insensitive to variation in the amount of child-directed speech that they hear. It would be much more surprising if there was no relation between amount of child-directed speech and more fine-grained measures such as vocabulary size, which the authors don't appear to have for the children in this study.

>> We have now underscored the lack of vocabulary measures in our data, as well as the lack of direct measures of interactional quality, and suggest what kind of care must be taken in future work attempting to extract such measures (see the Limitations section). We have also bolstered our argument that first words and word combinations are not expected to show uniformity across language environments—while the onset of these behaviors may be *less* affected by input quantity than, e.g., vocabulary size, they have

indeed been shown to vary between individuals and to predict later language outcomes, and so the fact that they do not seem delayed is meaningful in its own right.

Discussion: The discussion is more focused than the introduction, but there is still too much emphasis on the difference between close and panoramic approaches, and not enough on interpreting the results reported in the actual study. The authors need to work harder on developing an explicit rationale for their study in the introduction, and then use the discussion to reflect back on the questions that the study was designed to investigate. Note that, in suggesting this, I'm not encouraging the authors to engage in HARKing (i.e. Hypothesising After the Results are Known), since they do seem to have some a priori hypotheses that they present at the end of the introduction (albeit hypotheses that actually turn out to be wrong). What I'm saying is that they need to organize their introduction and discussion around these hypotheses rather than around a general (and, frankly, not particularly insightful) discussion of the strengths and weaknesses of close versus panoramic approaches.

>> We thank the Reviewer again for their guidance in reorganizing the main arguments of our paper. As mentioned above, the close-vs-panoramic contrast is now all but deleted from the manuscript and replaced with a comparative argument based on the possible effects of culture vs. economy on children's daylong input rate.

Reviewer: 2

Comments to the Author  
202 JCL Rossell

This paper reports on a monumental descriptive study of the language experience and early language development of children in a Papuan community which has the aims of evaluating different methods of describing language experience and of bringing a description of children's experience and development to bear on the question of how children learn language, given the input they get. The basic finding reported is that young children in this community experience little child-directed speech but reach basic milestones of language development on a similar time schedule as children in Western cultures who experience much more child directed speech.

The method that produced these findings was day-long recordings collected via a recorder that the child wore along with a miniature camera that also took photos of what was in front of the child every 15 seconds. The data come from the recordings of 10 children, selected from a larger sample that was recorded to be representative of this population in terms of maternal education, to span the age range from 0 to 3 years, and to have equal numbers of male and female children. Clips of the day-long recording were selected for transcription as follows: 9 2.5 minute clips were selected at random, 5 1-minute clips were manually selected to represent peak turn-taking activity, 5 1-minute clips were selected to represent peak vocal activity (on the part of the child?), and 1 5-minute clip described as an "expansion of the best one-minute clip. The manner in which these clips were selected is available online, but more needs to be said here. The reader



doesn't know enough about the data to proceed without going to the online material, and that shouldn't be.

>> We have now clarified that, indeed, the peak vocal activity is by the target child in this case. We have also added a paragraph describing the process for manual clip selection.

Each utterance in these clips produced by the target child was coded as non-canonical babble, canonical babble, single word utterance, multi-word utterance, or unsure. Again, the reader who wants to know more is referred to online material. We should at least know how many utterances were coded and the interrater reliability of the coding.

>> We have now added two paragraphs describing the procedure for annotating vocal maturity and addressee. We also now explicitly mention that we cannot conduct interrater reliability on these data due to the fact that only one pair of people (the first author and a native research assistant, working together) could produce these annotations across the three years of field visits. We do, however, now point to a manuscript under review in which these same vocalizations were annotated by crowdsourced non-experts as canonical or non-canonical (note: not any lexical utterances). Though this is only a partial and indirect measure of reliability, the annotations for canonical babbling they produced were similar to our own annotations.

The randomly selected clips provide the basis for baseline description, including effects of child age, time of day, household size, and number of speakers on the rate of child-directed and other-directed speech. The turn-taking clips provided the basis for describing the speech environment during peak interactions. The children's linguistic development is described based on cross-sectional measures of the proportion of less and more mature utterance types at each age.

The significant effects in the randomly selected clips indicate regular patterns in the time of day when children hear the most child-directed speech, that most child-directed speech came from adult speakers, that older children heard more child-directed speech from other children than did younger children. In essence, these analyses provide estimates of how much time per hour the child is talked to, how much time per hour the child hears other-directed speech, and these analyses identify the circumstances that influence who will talk to the child and how much. A finding that is important for the study's larger purpose is that the baseline rate of child-directed speech is low while the rate of other-directed speech is high, compared to Western samples.

The analyses of the clips of peak interaction times yield a different picture—one in which there is much more child-directed speech. This appears to be a central finding of this study—that even in a culture and context where the average rates of child-directed speech are low, there are bursts of dense child-directed speech. Another finding is that even where adults view children as conversational partners—as they do in this community—the circumstances of everyday life shape the timing and overall frequency of their talk to children.

The method and results summarized so far make an important contribution to the field of child language by documenting in greater detail than anyone has to this point the nature of children's linguistic experience in an unstudied community and some of the factors that shape that experience. Indeed, the introduction to the paper presents the research as an investigation into the difference in results that come from day-long recordings vs. recordings of shorter, more structured interactions. This is an important and interesting argument, but in reading it one feels as though one has walked in on the middle of a conversation and missed some important background. The paper introduces the terms "Panoramic method" and "close study." I tend to be against the introduction of new terms where existing words will suffice. At a minimum the contrasting approaches need to be better defined—is ethnography part of both, for example?

>> Reviewer 2 here echoes some of what we also heard from Reviewer 1 and the AE. We have taken on this advice seriously, rewriting most of the Introduction and Discussion. That said, we do hope to publish our thoughts about panoramic vs. close study methods in the future some time, and the question about ethnography is one we will keep in mind.

The next contribution this paper aims to make is to consider how language acquisition occurs, given such circumstances. The paper takes the descriptive, cross sectional findings with respect to age at which the milestones of canonical babbling, first words, and so on appear as evidence that language acquisition proceeds in a typical fashion, raising the question of how the children do it. The paper offers the argument that the bursts of high-density input provide the data for learning and the quiet periods offer the opportunity for consolidation. (It is not quite clear if consolidation is meant in the traditional learning and memory sense or also includes off-line processing of stored, unanalyzed input.) A serious treatment of this hypothesis needs to make this clear and then also tackle how such a proposal fits with other proposed models of language acquisition. Hypothesis testing v. building networks approaches, and the now-or-never bottleneck argument (multiple papers and a book by N. Chater and C. Mortensen).

>> Part of the newly reorganized Discussion is a section on implications of the present findings, which now includes a paragraph specifically dedicated to proposed mechanisms and models of language learning. A survey of the literature revealed no single model that best fits the circumstances of learning on Rossel Island so we focus instead on what kinds of models would be most suitable in this context, giving primary attention to prediction-driven models, an approach also highlighted in Christiansen and Chater's work. We also mention what kinds of model features are unlikely to account for these findings, including models that rely heavily on pedagogical cues, frequent contingent responses, and voluminous child-directed speech. We have also added a sentence clarifying the nature of the consolidation proposed.

In considering the theoretical implications of the finding, the paper makes good and appropriate arguments about the need for more research and, in particular, for studies of what child might learn from overheard speech. There is an important limitation of the study that is not adequately acknowledged: the language development measures are the age at which basic milestones are reached, and the measure of input is quantity. I think most researchers in the field would agree at this point that the quality of input matters more than quantity. Second, the many findings in the literature on effects of variability in



the quantity and quality of input on children's language development do not claim effects on the timing of major milestones. Rather it is the size of vocabulary, the rate of vocabulary growth, and the grammatical complexity of speech that are the associated outcome variables. So maybe input doesn't matter for reaching milestones and does matter for other sorts of individual differences. How will we ever know?

>> As mentioned in response to a similar comment from Reviewer 1, we now highlight these issues in the Limitations section of our paper. Our view is that we will know the answer to the question raised here by Reviewer 2 by measuring all of these characteristics of children's language development and more *rather* than focusing exclusively on vocabulary. In the future, we are potentially interested in measures of vocabulary and measures of the content that is conveyed in child-directed speech here. But adapting these measures such that they can be reliably and validly collected in sites different from those they were designed for takes significant time; developing a framework along these lines will likely take several years of iterated testing paired with ethnographic work. That said, we believe the current work opens up greater possibilities in that direction.

Relatedly, this paper would benefit from a bit more of a historical perspective. The challenge to theories of language acquisition from ethnographic descriptions of societies in which children hear little child-directed speech goes back to the 1970s, as some of the citations in this paper acknowledge. What this paper does not address, but perhaps should, is how this new sort of effort might enable new progress on an old argument.

>> With our thanks to Reviewer 2 for this supportive statement, we want to avoid overlooking other recent work pushing in this same direction. We have therefore only briefly acknowledged this historical point in the Introduction and Discussion.

The paper is very well written. There are a few sentences missing words that I'm sure will be caught in revision. The major difficulty for the reader is that the paper does not stand alone very well. It would be helpful to orient readers to the issues more in the introduction and to make the methods sufficiently clear so that readers can follow along without reading other cited background papers or looking up supplemental material.

>> We have done a careful read for missing words and typos in this new version; any further notes about typos spotted would be much appreciated on the next round of review.