23 June 2019

Dear Dr. Saylor,

We hereby resubmit our manuscript entitled, "Early language experience in a Tseltal Mayan village". Thanks to your and the reviewers' insightful comments, we believe that this second version of the paper is much improved, particularly with respect to how we frame the implications and limitations of the findings.

The main text of the manuscript is now 8903 words long (not including references), with 1 table and 6 figures (2 explanatory and 4 demonstrating the findings). The Supplementary Materials, which, as before, include detailed information about the statistical analyses, is also now reduced to 12 figures and 15 tables including: full model outcomes, conventional sanity checks of model fit, and full information for a set of parallel, alternative gaussian models to the negative binomials used in the main text.

We confirm that all authors have agreed to this new version of the manuscript for resubmission. In what follows, we address the individual points raised by yourself and by the reviewers. Please find our in-line responses below, marked with ">>" at the start of each one. Track Changes was not possible for revision as we are using Rmarkdown to write this paper. Instead we have highlighted the changes to our main manuscript as follows:

Green text = edits in direct response to Reviewer/Editor comment (see below for details)
Turquoise text = minor stylistic edits and updates to references in the text
Orange text = updated numerical information
| = indicates that some text following this spot was deleted

Please note that the results in this revision are very slightly different than the previous version. We had to remove the two-way interaction of target child age and household size from each of the statistical models due to model non-convergence. What happened? In making this revision we detected a bug in the software interface between the paper-writing library (papaja in Rmarkdown) and the statistical library (glmmTMB) we are using in the RStudio console: model non-convergence warnings simply don't show up when the code is run from the report document. With respect to the current paper, the most notable change is that the decrease in other-directed speech with child age is now significant (this is the same direction as in the previous submission). All other result changes involve minor (small and same-direction) edits to the effect size, estimated error, and related statistics or removal of a previous age * household size interaction (found previously in the non-converging model). **In a nutshell, there were previously undetected model convergence issues and we have now resolved them**. Please accept our apologies—the software bug is not our own, but we failed to notice it early on. We will immediately report this issue to the developers of the software so other authors don't have the same issue while using this report-generation pipeline.

Thank you very much again for your consideration, and please do let us know if you need any other information.

Sincerely, The Authors

```
##############################################################
# Action letter from Dr. Saylor
##############################################################
```

Your manuscript has been reviewed by two consultants who are highly knowledgeable in the area of this research. Their reviews are enclosed for your information.

As you will see from the comments below, both reviewers see considerable merit in your paper. For example, Reviewer 1 notes that the study is both "fascinating" and "well-conceptualized" and details several additional strengths of the data-analytic approach. Reviewer 2 is similarly pleased with both the writing and subject matter of the paper. I echo these positive sentiments. At the same time, both reviewers had serious concerns with the framing of your study and your analysis of how the data from your sample may align with data from children in Western cultures. On the basis of these thoughtful and knowledgeable reviews, I have decided that I must reject this manuscript for publication in Child Development. However, I join the reviewers in thinking that it may be possible to respond to the issues raised in the enclosed reviews by a revision. If you decide to revise the manuscript, I would like you to carefully address each of the reviewer's comments, even those I do not highlight below.
>> We very much appreciate the opportunity to revise the manuscript. We hope that you will find that we've satisfactorily addressed all the comments.

Reviewer 1 offers a very nice, extensive analysis of issues with your framing that I hope will offer some guidance to you as your work to reframe the purpose and scope of your study in the introduction. I also encourage you to revisit how you interpret data on language outcomes given Reviewer 1's concerns about the use of gross versus fine grained measures of children's language abilities.
>> Reviewer 1's comments were very helpful in re-shaping our framing. We describe below (in Reviewer 1's section) the precise changes we have made.

Additionally, an important issue raised by Reviewer 1 concerns the size of the sample reported in the manuscript. Given that these data are drawn from a sample that includes recordings from 100 children, it is not at all clear why you only report language samples from 10 children (of varied ages). It is not clear what is gained from this approach. I agree with Reviewer 1 that a better approach would be focus on children who are similar in age.
>> We're sorry that the reasons for our sampling strategy were not clear in the first submission. We have made our decision process more explicit in the manuscript (found under "Data Selection and Annotation" paragraph 1, page 13). Although the whole corpus indeed includes raw recordings from over 100 children (~50 from this Mayan site and ~50 from one in Papua New Guinea), we have only been able to add 10 hours of transcription and annotation to the 500 hours of Tseltal—these 10 hours are the data on which the current paper is based. Why only 10 hours? Annotation proceeds at a rate of approximately 50 minutes of work for every 1 minute of audio. This ratio is so large because the process involves making detailed linguistic transcriptions of highly noisy naturalistic data with help from assistants who have limited literacy in Tseltal due to the fact that reading and writing in school takes place in Spanish after ~age 10. We therefore limited our goal to what we could achieve over the course of three 6-week visits to the village over the course of three years. These 10 hours represent a drop in the bucket with respect to the Tseltal corpus as a whole (~2%), but to the best of our knowledge, this is the largest, broadest, and most detailed corpus of daylong indigenous linguistic input in existence (including excellent work by Cristia and Scaff (on Tsimane; Bolivia) and Shneidman (on Yucatec; Mayan)). We are also still investing in these data: the recordings and annotations are designed to be securely shared and expanded upon in the coming years, and are constructed in parallel with the Papuan data mentioned in the paper. We strongly believe that this combination of corpus features (parallel methods, annotation detail, multiple sampling schemes) is critical for building rich, comparative, re-usable non-WEIRD datasets, even though this process does also lead to early limitations in the speed and breadth of available annotated data. Also,

some of the estimates in the current paper are bound to change as we continue to visit these communities in the coming years. For this reason, in the paper we refer readers to a website where they can see and interact with the latest descriptive results. As a whole, we think that this paper gives a valuable glimpse into Tseltal children's early language environments and early language development; the findings are of course preliminary, but they are hard won. If our field required *significantly* more data before publication than what we have presented here, very few scholars would be motivated to engage in this sort of work long-term.

Finally, both reviewers note that you present a somewhat overwhelming amount of analyses in your results section that does not fit together in a coherent manner. I encourage you to go a bit easier on your reader and work to focus on a core set of findings in your reporting in the manuscript (you may use the supplement for additional, more peripheral findings of interest).
>> Thank you for this feedback. We agree that five dependent variables plus the vocal maturity and burstiness analyses were a lot for a single paper. We have therefore focused the new version of on just three measures: the quantity of target-child-directed speech, the quantity of other-directed speech, and children's vocal maturity. We removed all references to the other findings, which means we deleted a few sub-sections in the Results, Discussion, and Supplementary Materials, indicated with the red-highlighted in the main text. As Reviewer 2 notes, some of these other measures can be taken as somewhat indexical of input 'quality', so we are happy to bundle and present them in a separate paper.

Regrettably, at this time I cannot be certain whether or not a revision will be successful, and thus if you do decide to revise and resubmit, you should do so with the knowledge that I cannot guarantee acceptance.

If you do decide to revise and resubmit, I would appreciate a brief cover letter explaining on a point-by-point basis how you have addressed each of the reviewers' comments, including the pages on which changes have been made. Please also include a track-changes version of your manuscript in addition to the clean version prepared in accord with APA style. In general, please prepare your revision in accordance with the attached Revision Checklist.
>> We confirm that we have prepared the revision in accordance with the Revision Checklist. This cover letter explains our changes in response to reviewer comments on a point-by-point basis, with changes highlighted in the text colors mentioned above (Track Changes was not possible). Please note that in order to keep the manuscript short enough and easy to read I have made some stylistic edits and deletions as well, which I have noted in turquoise in the text.

I would like to have the manuscript returned by September 1, 2019. If you anticipate being delayed considerably beyond that time, I would appreciate a note about when to expect the paper so that we can keep the files open. Please let our office know via email (cdev@srcd.org) within the next 2 weeks if you plan to undertake the revision. If you decide to submit a revision, please submit it electronically to the Child Development Online Submission Site at,http://mc.manuscriptcentral.com/childdev, not directly to me.

If you choose to revise, please note that Child Development has recently adopted a new sociocultural policy requiring authors to provide greater detail regarding their sample's characteristics and context. Full policy requirements and examples can be found at http://www.srcd.org/sites/default/files/documents/sociocultural_policy.pdf
>> A demographic summary of the community is provided in the main text of the paper (found under "Corpus" on pages 10–12). We have now also added the name of the region in which the data were collected to the Abstract, which we hope satisfies the requirement for an abbreviated description of the participant group.

Methodological Recommendations: In an effort to further Child Development's tradition of publishing rigorous research, we have put together a set of methodological recommendations for authors. Please review and ensure your paper adheres to them as fully as possible. The recommendations and more information can be found at
http://srcd.org/sites/default/files/documents/cd_methodological_recommendations.pdf
>> We have reviewed the Methodological Recommendations. We believe that, now with our added extended comments on how and why our sample was constrained, we fulfill all the requirements. Note, too, that detailed training materials for all our annotations and the full data, annotations, and analysis scripts are available at the different links referred to in the text (with author-identifiable links masked for anonymity). This open approach should allow those interested in replicating or re-using our data or analyses to do so with little additional effort.

Whatever course of action you decide to follow, I hope that you find the reviewers' comments useful to you as you continue your work in this area. Thank you for considering Child Development as an outlet for your work.
>> Thanks again!

This is a fascinating and well-conceptualized study that raises important questions and presents very rich data on language inputs to children in a unique cultural community. The findings do a great job of raising awareness of how developmental experiences of many children fall outside those described in commonly studied US and Westernized samples. I am very enthusiastic about this research and there is a lot to admire about the work and effort that went into this paper.

Let me begin by stating what I like about this study. First, our field has too few rich, in-depth examinations of the experiences and learning of children (regardless of culture). The data generated from the excerpts drawn from all-day recordings and the analysis of input by time of day, interlocutor, household size, etc. are fascinating and illustrate the richness of a cultural approach. Second, analyses are well explained; use R; are shared through the OSF repository; and accompanied by excellent graphics (although too many). Third, the authors have thoroughly examined multiple features of language inputs to children, from contingency to time of day influences. Still, concerns arise, at both conceptual and methodological levels, with some of those limitations detracting from the paper's value.

First, I struggled with the framing that if children in other parts of the world have language experiences that diverge from the child-directed and other such experiences in the literature, then current models and theories will need to be revised. The research expectations were all clear and well described, but the last claim/hypothesis at the Intro: "that early vocal development would be on par with Western children" is misleading given the problems with defining vocal development as done here (perhaps, changing this hypothesis and working on the framing and key points will address some of my concerns). In my reading, the framing confounds the idea of what is sufficient in a particular context to learn whatever language works in a context, with how language input plays out in individual differences in ultimate skill and growth. Why would the authors attempt to draw equivalencies with the language development of children in other cultures who experience diverse and dense infant-directed speech? Do the authors truly believe that overhearing a small handful of words a minute is the same as hearing many words in interactive bouts in children's development of language? It is long recognized that children's experiences prepare them for integration into their communities, and the specific demands within those communities, and that language customs and practices are bound with culture (dating back to the seminal work of Scheiffelen and Ochs on situation-centered versus child-centered speech, which the authors do indeed cite). And so, there would be no reason not to expect children to acquire language to a sufficient level that allows them to integrate, talk, and engage quite well with others in their cultural contexts, and to show onsets of some basic language measures such as first words or combining words at reasonable ages. But, that does not also mean that a child exposed to more talk, more diverse talk, grammatically complex sentence structures, child directed speech, and so forth, would NOT achieve higher levels of language skill than a child in a very different language context as studied here if rigorous measures of language were adopted. And so yes, the Mayan children sampled in this study showed patterns of language *emergence* or onset times that crudely align with what is seen in other samples. But, the statement that "Tseltal developed early language on a similar timescale to Western children" is based on these very crude data. That is, the problem with the language "outcomes" being very gross is that the authors then claim that key language achievements show parallel paths (but, again, it is based on the timing of first words or combining words. And note, babbling too shows similar paths in all children across all communities regardless of inputs as many have shown). This is quite different than showing that vocabulary size, for example, reaches similar magnitudes at the same ages in two populations with dramatically different exposures, or that syntactic complexity, nesting of clauses, morpho-syntax, and so forth shows the same levels of attainment regardless of inputs. If the message is not very clearly qualified relative to the measures of interest (and why are in fact vocal onsets of interest?), then the ultimate take-home might be that children everywhere

are capable of attaining the same levels of language skills regardless of the nature of their language exposures.

>> We are grateful to Reviewer 1 for pointing out this possible interpretation of our framing—it's not the reading we intended and we agree that it is important for us to clarify what we mean and to also be more explicit about the limitations of our findings in both the Introduction and the Discussion. We do not desire to make misleading equivalency arguments about Tseltal versus other (Western) child language development. We have gone through the manuscript, from Abstract to Conclusion, and have attentively edited the manuscript as follows in order to avoid promoting true "equivalency" claims:

- **No direct implication of equivalency:** we removed uses of phrases like "on par" and "robust" when referring to Tseltal children's language development. Notes of similarity are now interpreted more explicitly as (a) evidence against obvious language delay—we explicitly now state that, based on these coarse measures there is no obvious evidence for language delay (as there would be with, e.g., <15% canonical syllables at 10 months; Oller et al., 1998)—and (b) evidence for the idea that Tseltal children gain enough linguistic evidence with minimal CDS to produce at least some words and multi-word utterances at more or less the ages we expect based on studies with Western children in CDS-plenty environments: pages Abstract, 6, 10, 23–25, 26–27 28, and 33. Every claim regarding the vocal maturity findings is also accompanied by some mention of limitations of our measure and our sample, as described in the following points.

- **Limitations of vocal maturity:** We have made more explicit that our findings are limited to vocal maturity transitional benchmarks (non-canonical babbling to canonical babbling, canonical babbling to first words, and first words to multi-word utterances) and must absolutely be followed up with richer linguistic analyses (something possible with these data, but outside of the scope of this first report): Abstract, pages 8, 23, 24, 26–27, 28, 32, and 33. In nearly every place where we make reference to our vocal maturity measure, we also now explicitly mark that it is "coarse", "rough", "initial", "basic", and/or "preliminary": Abstract, pages 8, 9, 16, 24, 26, 32, and 33.

- **Why vocal maturity:** We have also laid out our reasoning for why vocal maturity is a reasonable first approximation of children's early vocal development and why we might expect it to change across age given variation in children's linguistic environments: sub-section "Vocal maturity of spontaneous speech" (pages 8–9).

- **Broader implications of findings:** In the previous version we broadly stated that non-delayed linguistic development with little CDS should lead us to reconsider theoretical assumptions already in place. Reviewer 1 is correct that this statement could be misleading. We now specify two specific consequences of these findings, should they be shown to be true: Introduction paragraphs 2 and 3 (pages 3–4)

- **What kids are acquiring:** Where appropriate we now explicitly state that we can only make inferences with respect to the acquisition of Tseltal: Abstract, pages 4, 27–28, and 33. Reviewer 1 is also correct on this matter, and we only wish studies of, e.g., English did the same—we're happy to avoid making that mistake.

- **Time-of-day effects:** We're sorry that the significance of the time-of-day effects was not clear. To better foreshadow and then, later, unpack the time-of-day analyses and their theoretical significance we have now added more information about the course of a typical day for a mother in this community. We then use that description to state clear predictions about *when* children will hear speech throughout the day. Finally, we follow-up on the time-of-day findings using the same set of concepts in the Discussion: pages 5–6, 7, 10, 26, and 30–31.

Instead, therefore, I recommend that the framing state that children acquire the skills they need to function as communicative partners in their unique cultural contexts, and that many cultural contexts show very different paths to language learning than others (again dating back to work such as Schieffelin and Ochs). I would suggest dropping the equivalency argument entirely, or qualifying it hugely everywhere, unless other measures of language are obtained at the same ages across populations of interest in comparison.

Until then, the connection to child language outcomes rests on tenuous grounds.

>> As outlined in more detail above, our new version implements nearly all of these recommendations: (a) specifying that, in general, children acquire the skills appropriate to their community and that (b) our conclusions are pertinent to how children learn Tseltal in this community, and by (c) dropping direct/strong references to equivalency and, (d) when a reader might otherwise infer equivalency, reminding the reader of the strong limitations on the current results. We also (e) add a reference to other, independent comparative analyses of the Tseltal vocal data that are set to provide converging evidence for our claims (page 25). To preserve normal standards of author anonymity we have not listed the reference for this work, but we can share in this response letter that it involves citizen science (i.e., naive coder) annotations of vocal maturity for children under 3;0 in six different languages (including Tseltal) from daylong recordings, this time using the trajectory of Canonical Babbling Ratio (a more standard measure, e.g., Oller, 1988) as the outcome measure. The results of that study, which is in preparation, closely parallel our own findings for Tseltal vocal maturity and give compelling evidence for (very coarse) cross-linguistic similarity in the first three years of linguistic development (though without accompanying input measures, as we have here).

Second, related to the issue of child language (as the "outcomes"), the authors present very contradictory data to what they seem to be strongly arguing. The heading "Robust learning with less child-directed speech" suggests that these children would be showing strong language abilities. But, close inspection shows they are vocalizing very infrequently, and the authors themselves later make this point (in the same section as robust speech is stated) by saying: "The rarity of contingent responses may be due to the fact that children did not vocalize often". They then report low median talk and even those numbers included crying and laughing. The claims around robust learners and low talk generating rich language outcomes flies in the face of what was seen. When data were found to map to those reported by others, the measures were weak (such as saying 1 year+ age children were talking single words..... when there were only three 1-year olds or so in the study – see below point on sample--and saying a single word could mean a child has only one word to their vocabulary).

>> Thanks for these observations. We now avoid using the word "robust" to refer to the language learning patterns—the heading is re-named "Learning Tseltal with little child-directed speech". We also now make it clear that our cross-sectional vocal maturity measure is limited in its significance (though we define better what that scope is and what our predictions are based on, as described above). The rate of vocalization, in our opinion, is different from the basic linguistic competence of children in producing mature vocalizations (e.g., two adult speakers could be equal in language skill but differ greatly in how often they speak). That said, we've added a point in our paper highlighting that the vocalization rate is lower than we might expect based on Western norms (though consistent with SES-based patterns) and that more follow-up work with better matching data are needed (pages 25–26, and 32).

Third, the description of sample size was difficult to follow and the small sample of 10, across a wide age range, was a concern. It was also a bit difficult to follow the rationale for the current sample. The abstract starts by stating the recordings were based on 10 children. Then, the methods say recorded data were from 100 children; then the authors refer to "this" Tseltal corpus of 55 children. From that corpus, 10 children were chosen based on "spread in child age" (but why not all 55? Why this subsample? What's the rationale for the maximal spread, and why not focus on a narrower spread given this is Child Development and age makes an enormous difference in children's language development. I imagine the reason was to ask about the age-related changes in language inputs (one reported line of results) but then that should be stated and explicitly. Furthermore, however, the relevance of the "vocal maturity rating" of babbling, single words, multiple words and so forth will depend enormously on the age of child (a 0-3 year range, makes babbling relevant for a handful of the 0-1 year olds; single words for 1-2 year olds; combinations for 2-3 year olds, perhaps. But, with only 10 kids, that means there are around 3 children for whom each of the language measures or categories are appropriate.

>> We appreciate these comments—we wish we had more data too! Please see our responses to the

Editor above for a longer description of our rationale for the sample we use and please see the expanded explanation of these decisions in "Data Selection and Annotation" (paragraph 1, page 13).

Third, I personally am drawn toward rich descriptive data, and commend the efforts to describe children's language experiences presented here. However, even I struggled with the point to some findings and the main take-home messages, which means the broader readership might as well. For example, what should be made of the description of hours in the day when children were most likely to be exposed to speech? Or, what is to be made of the rates of contingent speech by time of day, etc.? What is the point or what is learnt from these data? There are simply too many numbers and figures to parse out the important points. Some streamlining is in order. One option is to place a lot of the data and figures in a supplementary section, and then to zero in on the main findings.

>> Thanks for this comment. As described above, we have now removed many results, only leaving in the findings on target-child-directed speech, other-directed speech, and vocal maturity. All other dependent variables (like the contingency measures) are left for another paper. To make it more obvious why the time-of-day effects might be interesting, we have added sections in the Introduction, Methods, and Discussion describing prior work along these lines and outlining the specific expectations we had for Tseltal time-of-day effects (pages 5–6, 7, 10, 26, and 30–31). We hope that the reviewer finds this new set of results much more streamlined.

Finally, and continuing the concerns about conclusions around the language skills of these children, the Discussion is filled with statements on children's expert language abilities. Again, I am in agreement these children learn language and can do so from overheard speech etc. (many studies show this is possible), but I am not convinced that their skills rival children in language dense communities where speech is directed to children. The data to support many of the conclusions are simply not there.

>> We agree that our prior version could have possibly (or even likely) misled the reader along these lines. We hope that the changes in framing we've made, outlined in more detail above, keep this implication out of the new version of the paper.

```
####################################################################
# Comments from Reviewer 2
####################################################################
```

This cross-sectional study following 10 Tseltal children is a valuable first contribution to the literature on first language acquisition. The innovative methods used in this study provide a crucial insight into the language-learning environment of a non-WEIRD community. I am excited to see this corpus grow as only tentative conclusions can be drawn from such a small cross sectional sample. I think this manuscript is well written and important. These suggestions are intended to help the author(s) in a future revision:

• The introduction sets the paper up as an investigation of CDS quantity but arguably measures of turn taking/contingency tap the quality of interaction. This could be unpicked more with respect to research focused on this aspect of CDS in the introduction/discussion.
>> We agree. Given these comments and those of the other reviewer and the editor, we have now saved the interaction-based analyses for a more in-depth look at Tseltal input 'quality' in a separate paper.

• Given the documented lack of record keeping in the community, how was infant age estimated?
>> Thanks for asking us to add this information—we are sorry we didn't add it to the first version. It can now be found on page 12.

• I bumped up against some of the terminology used e.g., "linguistic evidence to support first language acquisition" (page 2); measures of vocal maturity. Can these be clarified?
>> We have avoided the phrase "linguistic evidence" altogether and have defined "vocal maturity" and "canonical babbling" at the first point where these words appear in the main text. (page 8). This definition comes within a new sub-section describing in more detail what our hypotheses were regarding vocal maturity—hopefully this sub-section will make the definition and purpose of the vocal maturity measure a bit clearer on pages 8–9.

• Milestones are reached on a comparable timeframe to mid SES North American infants. Is there any literature to suggest how they compare on these metrics to low SES western infants (where there is a vocabulary gap)?
>> We do not have vocabulary information for these children (yet). While it would in theory be potentially interesting to have such information, there are a few limitations to doing so right now, including (a) we lack a completely comparative dataset for doing this analysis properly (e.g., TCDS min/hr) measure + vocabulary measures and (b) we don't yet know if vocabulary size has the same kind of predictive relationship with overall linguistic growth in Tseltal like it does in, e.g., English. As Reviewer 1 points out, we can only really understand competence as relative to the set of linguistic behaviors associated with any particular language community. Therefore, we could *hypothetically* find that vocabulary size in Tseltal indeed correlates with quantity of directed speech (it probably does; Shneidman & Goldin-Meadow 2012), but the importance of this possible finding for understanding how Tseltal children become competent adult speakers within their own environment is less clear. In this new version we have been much more specific about the fact that these findings should be interpreted as Tseltal children getting the information they need in order to begin to use and combine Tseltal words on the timescale comparable to what we might otherwise only expect for children in a frequent-CDS environment. Another way to think about this result (combined with the 6-language comparison I hinted at above) is that languages and language environments might be adapted such that children tend to reach basic vocal maturity milestones around the same points in development; this would **not** imply learning of the same things at the same rate across all learning environments, but it would suggest that there might be other potentially interesting factors at play in the mechanisms driving both language learning and evolution. That conclusion is beyond the scope of the current paper, but gives some impression of why more and higher quality comparative data in the future might provide important breakthroughs.

- I applaud the commitment to open science and the extensive appendices, however currently it is difficult to parse the (many) tables and graphs, can more context be provided here?

>> Thanks for bringing our attention to this! Many of the figures have now been removed because we decided to focus only on target-child-directed speech, other-directed speech, and vocal maturity. For the remaining figures in the main text, we have rephrased nearly all of the captions to clarify how they should be read. We hope that this makes it easier to parse the results in the main text. Regarding the Supplementary Materials (SM): It is very important to us to provide detailed background statistical information in the SM for those readers who are predisposed to digging into the analysis details but don't yet have time to download and run the full set of analysis scripts. We have designed the SM with those particular readers in mind. In this new version we have lightly edited the text for clarity, but still rely on the fact that the reader should have prior knowledge of the tools and concepts we employ.

>> Thanks very much again to Reviewers 1 and 2 and the Editor for a very helpful set of comments! We hope you find the paper significantly improved.