Response to reviewers

Editor comments

1. Descriptive analysis rather than significance testing
2. Effect of age of sibling – does having an older sibling lead to more IDS?  
   *Correlations between amount of input and age difference between the target child and each sibling show that:*- *Infants with older siblings hear more words overall*

* *Infants with younger siblings hear fewer words overall*
* *Looking just at caregiver input, age difference between infant and sibling makes no difference to the number of words heard from the caregiver when the siblings are older, but infants with younger siblings hear less input from caregivers overall (though not a strong correlation)*
* *There is no correlation between amount of input heard from the sibling and age difference between the infant and the sibling when the sibling is older, but when the sibling is younger, there is a weak positive correlation between age difference and input (this doesn’t make sense, needs to be plotted!)*

1. Object-presence vs. object-naming  
   Elika – is there a reason for this?

Reviewer 1

1. Findings from Havron et al. (2019) and others: reassess contribution of paper and note lack of robustness of Havron et al finding.   
   *Changed the text to read [lazily just copied and pasted for now]:   
   “Recent studies have also identified effects for age gap between the target child and their siblings [whereby larger age gaps correlate with lower vocabulary scores; @havron\_effect\_2022; @gurgand\_influence\_2022] and sibling sex [whereby older brothers have a negative effect on vocabulary outcomes, but not older sisters; @havron\_effect\_2019-1; @jakiela\_big\_2020] though neither of these effects are found consistently across datasets; @havron\_effect\_2022 and @gurgand\_influence\_2022 find no effect for sibling sex, whereas @havron\_effect\_2019-1 find no effect for age gap. Some of these differences across studies may relate to insufficient power to detect relatively small effects or simultaneous contributing factors that are difficult to disentangle.”*
2. R1 suggests that analysis of object presence adds something novel to the literature; make that clear.

*Need to edit this, have made a start.*

1. Twins: re-run with both twins excluded  
   *Add this to SI for now (and maybe swap analyses around, depending on which makes most sense?)*
2. Object presence: clarify definition with examples  
   *Changed the text to read:  
   “For example, if the caregiver was pointing at a ball while the saying the word \*ball\*, this was coded as "yes". If the infant was holding (but not looking at) a bottle while the caregiver said \*bottle\*, this would also be coded as "yes". On the otherhand, if the caregiver refers to a dog that is barking in the other room, that would be coded as "no", as it was not present during object labelling.”*
3. Object presence and siblings: is it possible to analyse this measure in terms of whether or not a sibling was present in the room with the infant?  
   *Siblings were present in 70% of recordings for all infants who had a sibling. There is a distinction in % of object presence when the sibling is present compared with when they aren’t. Infants with siblings hear less object presence when the sibling is present in the room. Note that it was not possible to determine how many siblings were present in a given session. Impressionistically, when the three SibGroups are compared, there is a gradient decrease in % of object presence, across the board, and also in both contexts where the sibling is and isn’t present. That is to say that no sibling presence still shows a gradient decline in % of object presence across the three sibling groups.*
4. Effect of age: look at changes to input measures over time  
   *From graph inspection, between-group trends appear to be consistent over time for both measures.*
5. Would it be possible to run a mediation analysis between number of siblings, language input and vocabulary?  
   *I have run a (preliminary) mediation analysis between vocab size ~ Sib group + input. No mediation effect found and no correlation between input and vocab size. However, we might want to test this differently if we think it’s worth pursuing further.*
6. Bias in reporting: parents with more children may be less attuned to the target child’s vocab knowledge – *correlate with recorded vocab measures?*
7. Minor comments

Reviewer 2

1. Results: Present a more descriptive analysis draw on effect sizes more than significance tests and uses visualizations to demonstrate these findings (as is already the case).
2. Overstatement of findings: input an vocabulary are never analysed together so it isn’t possible to draw conclusions about a relationship between these. Hypotheses aren’t actually fully tested in the analysis as gradient effects are not observed (check this).  
   A bit confused here: how have we not tested the gradience of productive vocabulary across sibling number? We have tested it, it is just not fully borne out in the findings. Maybe make this clearer in the Discussion?
3. Literature: include literature on overheard speech in the lit review as an example of how there are different ways to learn from different kinds of input  
   In progress – want to firm up the analyses before working any further on this.
4. Target child: make clear that we only analysed speech directed at the target child  
   *This is already made pretty clear in the text but I have added to the “Derived input measures” section:*  
   “*As mentioned above, only speech produced in the infant's immediate surroundings (i.e. speech that would have been clearly heard by the target infant) was coded.”*
5. Input measures: be clear on what we mean about household input and object presence  
   *Done – see response to R1.*
6. I also wonder how big the expected / reported differences are in light of overall individual differences reported in relation to input. This literature also makes it almost obvious that the effect is graded, why would it not be? I would be curious why this would be any different and would hope the authors discuss this in more detail   
   I don’t really understand this comment. So correlations between vocab size and input, or something else?
7. Bias in reporting: parents with more children may be less attuned to the target child’s vocab knowledge  
   *See comment above*
8. Minor comments