Parents fine-tune their speech to children's vocabulary knowledge

XXXXX[1], XXXXX[1], & XXXXX[1,2]

[1] XXXXX

[2] XXXXX

Author Note

All data and code for these analyses are available at https://osf.io/vkug8/.

Correspondence concerning this article should be addressed to XXXXX, XXXXX. E-mail: XXXXX

Abstract

Young children learn language at an incredible rate. While children come prepared with powerful statistical learning mechanisms, the statistics they encounter are also prepared for them: Children learn from caregivers motivated to communicate with them. How precisely do parents tune their speech to their children's individual language knowledge? To answer this question, we asked parent-child pairs (n=41) to play a reference game in which the parent's goal was to guide their child to select a target animal from a set of three. Parents fine-tuned their referring expressions to their children's knowledge at the lexical level, producing more informative references for animals they thought their children did not know. Further, parents learned about their children's knowledge over the course of the game, and tuned their referring expressions accordingly. Child-directed speech may thus support children's learning not because it is uniformly simplified, but because it is tuned to individual children's language development.

Statement of Relevance

The pace at which children learn language is one of the most impressive feats of early cognitive development. One possible explanation for this rapid pace is that the language caregivers produce is tuned to children's developing linguistic knowledge, maintaining just the right level of complexity to support rapid learning. We present the first experimental evidence of just how precise this tuning is, showing that parents tune not just to children's holistic language development, but their knowledge of individual words. We developed a new method in which we experimentally controlled what parents talked about, but not how they could talk or what they could say, increasing the chance that these results will generalize outside the lab. This work points to the importance of studying the parent-child dyad as a unit instead of focusing on children as isolated learners, both in the domain of language and in social learning more broadly.

Keywords: parent-child interaction; language development; communication

Word count: 749

Parents fine-tune their speech to children's vocabulary knowledge

In a few short years children master their native language. Undoubtedly, a large share of the credit for this feat is due to powerful learning mechanisms that children bring to their input (Kuhl, 2004; Saffran, Aslin, & Newport, 1996). However, a share of the credit may also be due to the linguistic input itself: Individual differences in the quantity and quality of the language children hear are associated with differences in language learning (Hart & Risley, 1995; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Rowe, 2012). Further, these associations are driven by speech directed to children. Differences in overheard speech do not predict differences in language learning, even in communities where child-directed speech is relatively rare (Romeo et al., 2018; Shneidman & Goldin-Meadow, 2012; Weisleder & Fernald, 2013).

Child-directed speech differs from adult-directed speech along many dimensions, most of which are characterized by simplification (Snow, 1972). But, the amount of simplification changes over development, e.g. parents direct longer and more complex utterances to older children (Huttenlocher et al., 2010). Child-directed speech may thus facilitate learning not because it is uniformly simpler, but because it is adaptive, tuned to the right level of complexity for children's ongoing language development (Snow, 1972; Vygotsky, 1978). How precisely tuned is child-directed speech?

One possibility is that tuning is *coarse*: Caregivers could tune the complexity of their speech generally, using a holistic sense of their children's developing linguistic abilities. This accords evidence showing that parents tune their utterance lengths, articulation of vowels, and diversity of clauses to children age (Bernstein Ratner, 1984; Huttenlocher et al., 2010; Moerk, 1976). However, linguistic tuning would be an even more powerful scaffold for learning if parents *fine-tuned* their speech, taking into account not only children's global linguistic development, but their specific knowledge of smaller units of language, such as lexical items (Bruner, 1983). To date, the only evidence for fine-tuning comes from two observational studies, one showing that parents are more likely to provide their child with labels for novel as compared to familiar toys (Masur, 1997), and the second showing that one child's caregivers produce their shortest utterances containing a particular word just before the child first produces that word (Roy, Frank, & Roy, 2009).

Here, we present the first *experimental* evidence for fine-tuning. Children and their parents played a reference game in which the parent's goal was to guide their child to select a target animal from a set of three. Parents tuned the amount of information in their utterances not just to the average difficulty of each animal word, but to their prior estimates of their individual child's knowledge of that animal. Further, parents sensitively adapted over the course of the game, providing more information on subsequent trials when they discovered that their child did not know an animal. Together, these results show that parents leverage their knowledge of their children's language development to fine-tune the linguistic information they provide.

Method

Participants

Toddlers (aged 2-2.5 years) and their parents were recruited from a database of families in the local community or approached on the floor of a local science museum in order to achieve a planned sample of 40 parent-child dyads. Because our method was novel, we chose a sample size that would give us 95% power to detect a medium-sized effect (d=.6) within-subjects and rounded up to the nearest multiple of 10. A total of 48 parent-child pairs were recruited, but data from 7 pairs were dropped from analysis because of failure to complete the experiment as designed. Of the 7 pairs that were dropped, 5 children fussed out, 1 had an older sibling interfering with the study, and 1 was a twin (only the twin that participated first was included). The final sample consisted of 41 children aged 24 mo.; 5 days to 29 mo.; 20 days (M = 26 mo.; 0 days), 21 of whom were girls.

In our recruitment, we made an effort to sample children from a variety of racial and socio-economic groups. Our final sample was roughly representative of the racial composition of the Chicago Area and the US more broadly (56.1% White, 26.8% Black, 7.69% Hispanic). However, our sample was significantly more educated than the broader community (85.4% of mothers had a College or Graduate Degree).

Stimuli

Eighteen animal images were selected from the Rossion and Pourtois (2004) image set, a colorized version of the Snodgrass and Vanderwart (1980) object set. Animals were selected based on estimates of their age of acquisition (AoA) for American English learners. To obtain these estimates, we used two sources of information: parent-report estimates of children's age of acquisition from Wordbank (Frank, Braginsky, Yurovsky, & Marchman, 2017), and retrospective self-report estimates of AoA from adults (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012, see Supplemental Materials for details). The AoA of the selected animals ranged from 15 to 32 months. Half of the animals were chosen to have an early AoA (15-23 months), and the other half were chosen to have a late AoA (25-32 months). Each trial featured three animals, all from either the early AoA or late AoA category. This separation was designed to lower the likelihood that children could use knowledge of early AoA animals to infer the correct target on late AoA trials.

A modified version of the MacArthur-Bates Communicative Development Inventory Short Form (CDI; Fenson et al., 2007), a parent-reported measure of children's vocabulary, was administered before the testing session via an online survey. The selected animal words were added to the standard words, producing an 85 word survey. Two of the animal words–one in the early AoA (pig) and one in the late AoA category (rooster)–were accidentally omitted, so trials for those words were not included in analyses as we could not obtain individual-level estimates of children's knowledge.

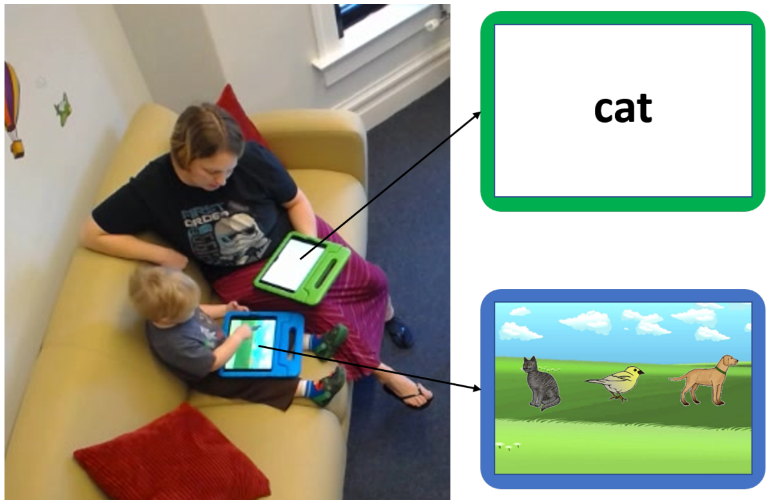

Figure 1. A parent-child dyad playing the reference game. On each trial, the parent's goal was to use language to communicate to their child which animal to choose.

Design and Procedure

Each parent-child pair played an interactive reference game using two iPads (Figure 1). Children began with two warm-up trials in which they tapped on circles that appeared on the iPads. Following these warm-up trials, children and their parents moved on to practice and then experimental trials. On each trial, three images of animals were displayed side by side on the child's screen, and a single word appeared on the parent's screen. Parents were instructed to communicate as they normally would with their child, and to encourage their child to choose the object corresponding to the word on their screen. The child was instructed to listen to their parent for cues. Once the child tapped an animal, the trial ended, and a new trial began. There were a total of 36 experimental trials, such that each animal appeared as the target twice. Trials were randomized for each participant, with the constraint that the same animal could not be the target twice in a row. Practice trials followed the same format as experimental trials, with the exception that images of fruit and vegetables were shown. All sessions were videotaped for transcription and coding.

Data analysis

Our primary quantity of interest was the amount of information that parents provided in each of their utterances. To approximate this, we measured the length of parents' referring expressions–the number of words they produced on each trial before their child selected an animal. Length is an imperfect proxy for information, but it is easy to quantify and theory-agnostic. Because utterance length is highly right-skewed (i.e. most utterances are short), we log-transformed length in all analyses. However, to facilitate interpretability, we show raw utterance length in our figures. Subsequently, utterances were manually coded for the following: (1) Use of an animal's canonical label (e.g., "leopard"), (2) Use of a descriptor (e.g., "spotted"), (3) Use of a comparison (e.g., "like a cat"), (4) Use of a superordinate level category label (e.g., "bird" for peacock), and (5) Use of a subordinate level category label (e.g., "Limelight Larry," a fictional character from a children's book, for peacock). Parent utterances irrelevant to the game (e.g. asking the child to sit down) were not analyzed. Children's utterances were coded when audible, but were not analyzed. Our second source of data was the vocabulary questionnaire that parents filled out prior to participation. Parents indicated whether their child produced each of the 85 words on the survey. In addition to analyzing parents' judgments for the animals in the task, we also computed the total number of words judged to be known for each child as a proxy for total vocabulary.

All of our analyses were done using mixed-effects models. In all cases we began with maximal random effects structures and pruned random effects until the models converged. We removed interaction terms before removing main effects, and opted to keep the most theory-relevant random effects when only a subset of main effects could be kept. For clarity, we present only the key findings and statistics here, but full model details can be found in the Supplemental Materials.

Results

We begin by confirming that our a priori divisions of animals into early and late age of acquisition (AoA) in the study design were reflected in parents' survey judgments, and that children were able to follow parents' references to select the correct target animal on each trial. After this, we show that parents fine-tune their referring expressions, producing more information in their references to animals that they think their individual children do not know. Further, parents update their tuning over the course of the experiment, producing more information on subsequent references to animals they thought their children knew but observed evidence to the contrary (i.e. children made an incorrect selection).

Target animal difficulty

We first confirm that the early AoA animals were more likely to be marked "known" by the parents of children in our studies. As predicted, parents judged that their children knew 94% of the animals in the early AoA category, and 33% of the animals in the late AoA category, which were reliably different from each-other (β = -6.48, $p < .001$, d = -3.57 [-4.48, -2.67]). Parents' judgments for each target word are shown in the Supplemental Materials.

Selection accuracy

On the whole, parents communicated effectively with their children, such that children selected the correct target on 69.05% of trials, reliably greater than would be expected by chance (33%, $\beta = 2.07$, $p < .001$, $d = 1.14$ [0.93, 1.36]). Children were above chance both for animals that parents thought they knew ($M = 75.08\%$, $\beta = 2.61$, $p < .001$, $d = 1.44$ [1.18, 1.70]), and for animals that parents thought their children did not know ($M = 55.19\%$, $\beta = 1.23$, $p < .001$, $d = 0.68$ [0.52, 0.84]). Thus, parents successfully communicated the target referent to children, even when parents thought their children did not know the name for the animal at the start of the game.

Was this accuracy driven by children's knowledge or parents' referential expressions?

To answer this question, we fit a mixed-effects logistic regression predicting children's accuracy on each trial from children's total estimated vocabulary, parent-reported knowledge of the target animal, and the (log) length of parents' expressions. We found that children with bigger vocabularies were more accurate in general (β = 0.40, p = .001, d = 0.22 [0.09, 0.36]), and that children were less accurate for animals whose names parents thought they did not know (β = -1.86, p < .001, d = -1.02 [-1.46, -0.58]) Longer referential expressions were associated with lower accuracy for animals that parents thought their children knew (β = -0.40, p = .007, d = -0.22 [-0.38, -0.06]), but greater accuracy for animals that parents thought their children did not know (β = 0.46, p = .025, d = 0.25 [0.03, 0.47]).

Thus, longer referential expressions were associated with more successful communication for animals that parents thought their children did not know, but were unhelpful for animals that parents thought they did know. We next ask whether parents tuned the lengths of their utterances appropriately, producing longer expressions for animals they believe their children do not know.

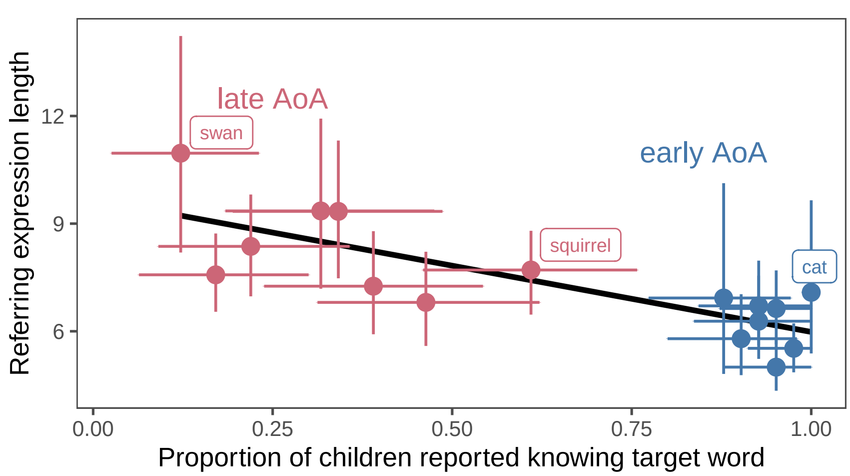

Figure 2. Parents produced longer referring expressions to communicate about animals that children were generally less likely to know.

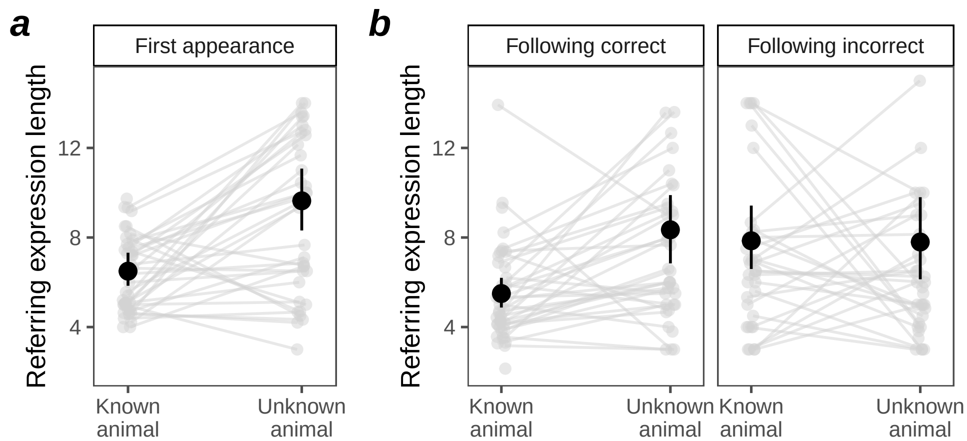

Figure 3. (A) Parents produced longer referring expressions for words that they thought their child did not know on their first apperance. (B) When children selected correctly in response to these expressions, parents continued to produce longer expressions for animals they thought their children did not know (left). However, if parents thought their child knew an animal, and they chose incorrectly, parents produced longer expressions on its second appearance (right). Gray points and lines represent individual participants, colored points and lines show group averaged proportions; error bars show 95% confidence intervals computed by non-parametric bootstrap.

Tuning

If parents calibrate their referring expressions to their children's linguistic knowledge, they should provide more information to children for whom a simple bare noun (e.g. "leopard") would be insufficient to identify the target. Parents did this in a number of ways: with one or more adjectives (e.g. "the spotted, yellow leopard"), with similes (e.g. "the one that's like a cat"), and with allusions to familiar animal exemplars of the category (e.g. "pick Midnight"). In many of these cases, parents would be required to produce more words (see below for further qualitative analyses). Thus, we first analyzed the (log) length of parents' referring expressions as a proxy for informativeness.

When do parents produce longer referring expressions? One possibility is that parents tune at the coarsest level, using more words when speaking to children with smaller vocabularies. This was not the case–the total number of words parents thought their children knew did not reliably affect the length of their referring expressions (β = -0.02, p = .595, d = -0.17 [-0.79, 0.45]). A second possibility is that parents have a sense for how difficult each animal is in general, and tune coarsely to this. Our analyses confirmed this coarse tuning: Parents said reliably fewer words for animals that more children were reported to know (β = -0.17, p = .034, d = -1.19 [-2.26, -0.09]); Figure 3A). Finally, parents could fine-tune their referential expressions to their children's individual knowledge, over and above the average difficulty of each animal. Our analyses supported this conclusion: Parents used reliably fewer words to refer to animals that they thought their individual child knew (β = 0.25, p = .003, d = 0.98 [0.34, 1.61]); Figure 3B). Thus parents fine-tuned the amount of information in their referential expressions, calibrating the amount of information they provided to their children's knowledge, even after accounting for the average difficulty of the target animal.

In addition, because each animal appeared as a target twice, we asked whether parents tuned their referential expressions over successive appearances. We found that parents used fewer words on the second appearance of each animal (β = -0.08, p = .044, d = -1.06 [-2.07, -0.03]), but that the difference in utterance length between animals they thought their children knew versus didn't know was smaller on their second appearance (β = -0.14, p < .001, d = -0.17 [-0.22, -0.12]). Why might that be? One possibility is that parents obtain information from the first appearance of each animal: they may have thought their child knew "leopard," but discovered from their incorrect choice that they did not. If so, they might provide more information the second time around.

To test this prediction, we fit a model predicting the (log) length of parents' referring expressions from appearance type (first, following correct, following incorrect), whether the parent thought their child knew the animal prior to the experiment, and their interaction between appearance type and prior belief. Relative to their utterances on an animal's first appearance, parents produces shorter referring expressions on an animal's second appearance following both correct responses (β = -0.14, p = .036, d = -0.12 [-0.23, -0.01]) and incorrect responses (β = -0.28, $p = < .001$, d = -0.22 [-0.34, -0.11]). As before, parents produced shorter utterances for animals they thought their child knew (β = -0.31, $p = < .001$, d = -0.92 [-1.43, -0.41]). When children were correct on an animal's first appearance, parents referring expressions on its second appearance did not differ in length based on whether they thought their child knew the animal prior to the experiment (β = -0.02, p = .771, d = -0.02 [-0.13, 0.10]). However, when children were incorrect on an animal's first appearance, and parents thought they knew the animal prior to the experiment, they produced reliably longer referring expressions on it's second appearance (β = 0.43, $p = < .001$, d = 0.24 [0.13, 0.35]; Figure 3B).

As we predicted, when parents thought their children knew an animal, but then observed evidence to the contrary, they provided more information in their referring expressions for children to make the correct selection the second time. However, we did not find the opposite pattern: When children were successful for animals that parents thought they did not know, parents did not update their beliefs. Why should parents update their beliefs in one direction but not the other? One likely explanation comes from parents' linguistic tuning itself. Parents' goal in this task is to produce a referential expression that allows their children to select the target animal whether or not they know its canonical label. Consequently, when children select correctly on these trials, parents cannot know whether their child actually knew the animal in question, or whether their referential expression provided information that allowed the child to select the target despite not knowing its canonical label.

Together, these two sets of analyses suggest that parents tune their referring expressions not just coarsely to how much language their children generally know, nor their knowledge about how hard animal words are on average, but finely to their beliefs about their individual children's knowledge of specific lexical items. Further, when interaction allows them to discover that they have incorrect beliefs about their children's knowledge, they update these beliefs in real-time and leverage them on subsequent references to the same lexical item.

Content of referring expressions

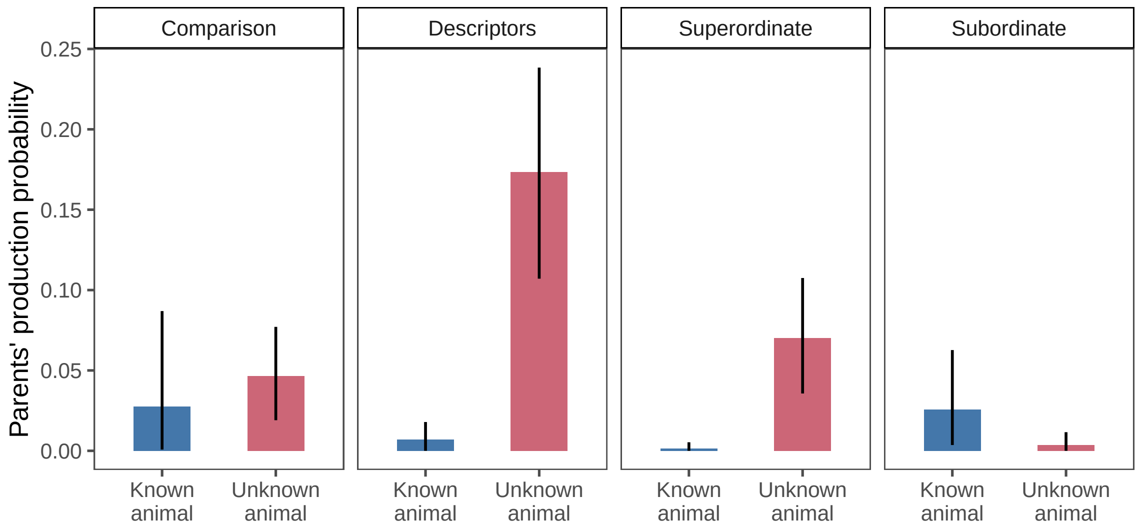

Figure 4. Proportion of trials where parents used comparison, descriptors, superordinate category labels, and subordinate category labels. Each panel shows proportions for andimals that parents thought their children knew and animals parents thought their children did not know.

Parents produced reliably longer referring expressions when trying to communicate about animals that they thought their children did not know. In the analyses presented so far, we used length as a theory-agnostic, quantitative measure of information. *How* did parents successfully refer to animals that their children did not know? As a post-hoc descriptive analysis, we coded five qualitative features of referring expressions: (1) Use of the animal's canonical label (e.g., "leopard"), (2) Use of a descriptor (e.g., "spotted"), (3) Use of a comparison (e.g., "like a cat"), (4) Use of a superordinate level category label (e.g., "bird" for peacock), and (5) Use of a subordinate level category label (e.g., "Limelight Larry" for peacock). Because the rates of usage of each of these kinds of reference varied widely (e.g. canonical labels were used on 94.48% of trials, but subordinates were used on 1.21% of trials), we fit a logistic mixed effects model separately for each reference kind, estimating whether it would be used on each trial from whether the parent thought their child knew the animal. Canonical labels were used on almost all trials, and did not differ in frequency between animals parents thought their children did not know (M = 94.68%) and animals they thought their children knew (M = 93.43%, β = 0.43, p = .216, d = 0.23 [-0.14, 0.61]). All other results are plotted in Figure 4.

Comparisons were used reliably more for animals parents believed their children did not know than for animals they thought their children knew, (β = 2.29, p = .001, d = 1.26 [0.49, 2.04]), as were descriptors (β = 3.09, p < .001, d = 1.71 [1.07, 2.35]) and superordinate category labels (β = 3.01, p = .026, d = 1.66 [0.20, 3.12]). Subordinates were used less for animals parents thought their children did not know than for animals they thought their children knew (β = -2.19, p = .025, d = -1.21 [-2.26, -0.15]). Thus, parents used a variety of strategies to refer to animals that they believed their children did not understand, but the use of descriptors was the most prominent. These descriptors are particularly apt to facilitate children's learning, connecting parents' fine-tuning for reference with their children's language acquisition.

Discussion

Parents have a wealth of knowledge about their children's linguistic development (Fenson et al., 2007). We show that parents draw on this knowledge to tune speech to their children in three ways: (1) they produce longer, more informative referring expressions for later-learned animals, (2) over and above this coarse tuning, parents tune information to their individual children's knowledge of specific animals, and (3) when children do not know an animal that parents thought they did, parents' subsequent references reflect this updated belief. Further, more informative referring expressions were associated with increased likelihood of successful communication: children were more likely to correctly select animals whose names they did not know if their parents produced longer utterances to refer to them. Finally, we found that references to unknown animals were rich with descriptors and comparisons, helping children select the correct animal and potentially serving as a source of learning input.

These data are consistent with a strong form of the linguistic tuning hypothesis, in which parents fine-tune the information in their speech to children's language knowledge at the individual-word level. Why should this happen? Although parents' speech to children is unlikely to reflect a goal to teach, it is nonetheless goal-oriented: parents want to communicate successfully (Bruner, 1983). Fortunately, learning may piggyback on this optimization because of the inherent synergy between communication and learning–it is easier to learn from input that you understand (Yurovsky, 2018). Children bring powerful learning mechanisms to language acquisition, but these mechanisms are supported by an ecological niche designed for their success.

Acknowledgements

Research funded by a James S. McDonnell Foundation Scholar Award.

References

Bernstein Ratner, N. (1984). Patterns of vowel modification in mother–child speech. *Journal of Child Language*.

Bruner, J. (1983). *Child's talk: Learning to use language* (pp. 111–114). Norton.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & others. (2007). *MacArthur-bates communicative development inventories: User's guide and technical manual*. Baltimore, MD: Brookes.

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.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young american children.* Paul H Brookes Publishing.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*(4), 343–365.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*(11), 831–843.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 english words. *Behavior Research Methods*, *44*(4), 978–990.

Masur, E. F. (1997). Maternal labelling of novel and familiar objects: implications for children's development of lexical constraints. *Journal of Child Language*, *24*, 427–439.

Moerk, E. L. (1976). Processes of language teaching and training in the interactions of mother-child dyads. *Child Development*, 1064–1078.

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science*, *29*(5), 700–710.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, *33*, 217–236.

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, *83*(5), 1762–1774.

Roy, B., Frank, M., & Roy, D. (2009). Exploring word learning in a high-density longitudinal corpus. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 31).

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

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. Retrieved from http://arxiv.org/abs/NIHMS150003

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174.

Snow, C. E. (1972). Mothers' Speech to Children Learning Language. *Child Development*, *43*(2), 549–565.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.*

Harvard university press.

Weisleder, A., & Fernald, A. (2013). Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary. *Psychological Science*, *24*(11), 2143–2152.

Yurovsky, D. (2018). A communicative approach to early word learning. *New Ideas in Psychology*, *50*, 73–79.