Parents Fine-tune Their Speech to Children's Vocabularies

Ashley Leung[1], Alexandra Tunkel[1], & Daniel Yurovsky[1,2]

[1] The University of Chicago

[2] Carnegie Mellon University

Author Note

Parts of this work were presented at the Annual Conference of the Cognitive Science Society: Leung et al. (2019). All data and code for these analyses are available at https://osf.io/vkug8/.

Correspondence concerning this article should be addressed to Ashley Leung, 5848 S University Ave, Chicago IL, 60637. E-mail: ashleyleung@uchicago.edu

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 do caregivers adapt their speech in order to support children's comprehension? We asked children and their parents to play a simple reference game in which the parent's goal was to guide their child to select a target animal from a set of three. We show that parents fine-tune their referring expressions to their children's knowledge at the lexical level, producing more informative references for animals that they thought their children did not know. Further, parents learn about their children's knowledge over the course of the game, and tune their referring expressions accordingly. Child-directed speech may thus support children's learning not because it is uniformly simpler than adult-directed speech, but because it is tuned to individual children's language development.

Keywords: parent-child interaction; language development; communication

Word count: 854

Parents Fine-tune Their Speech to Children's Vocabularies

In just a few short years, children develop a striking mastery of their native language. Undoubtedly, a large share of the credit for this feat is due to the powerful learning mechanisms that children bring to bear on their input (Kuhl, 2004; Saffran, Aslin, & Newport, 1996). However, a share of the credit may also be due to the structure of linguistic input itself. In line with this proposal, individual differences in the quantity and quality of the language children hear are associated with individual differences in language learning (Hart & Risley, 1995; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Rowe, 2012). Further, this association is driven by speech children hear in interactions with their caregivers; 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 (Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989; Snow, 1972). But, the amount of simplification changes over development–parents direct longer and more complex utterances to older children (Huttenlocher et al., 2010). Thus, child-directed speech may facilitate learning not because it is uniformly simpler, but because it is adaptive, being 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 developing linguistic abilities. This possibility is consistent with a body of 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 the lengths of three caregivers' utterances containing a particular word are shortest just before their child first produces that word (Roy, Frank, & Roy, 2009).

Here, we present the first *experimental* evidence for fine-tuning. We asked children and their parents 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 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 reference 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 considerable knowledge of their children's language development to fine-tune the 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. 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. The final sample consisted of 41 children aged 24 mo.; 5 days to 29 mo.; 20 days (M = 26 mo.; 0 days), twenty-one 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% White, 27% Black, 8% Hispanic).

However, our sample was significantly more educated than the broader community (85% 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 Supporting Information for details). The AoA of the selected animals ranged from 15 to 32 months. Half of the animals were chosen to have an early age of acquisition (15-23 months), and the other half were chosen to have a late age of acquisition (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 embedded among the 85 in the 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.

Design and Procedure

Each parent-child pair played an interactive reference game using two iPads (Figure 1). Children began with two warm-up trials on which they tapped on circles that appeared on

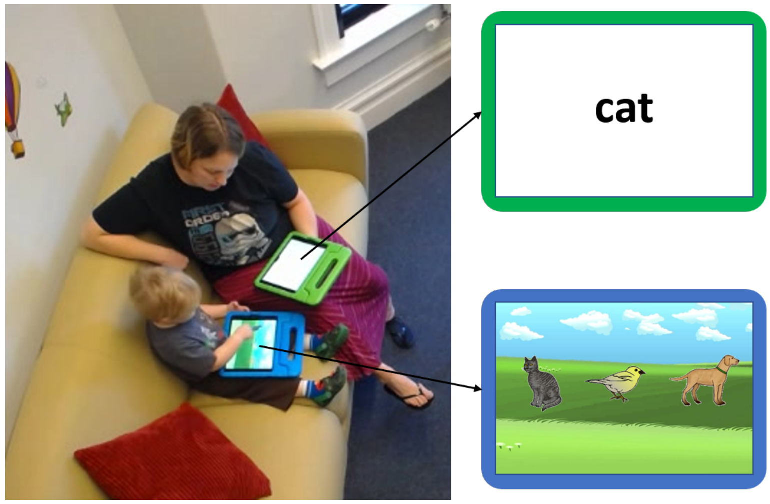

Figure 1. A parent-child dyad playing the reference game. On each trial, the parent's goal was to use language to communicate to their child which animal to choose.

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 utterances 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 subordinate category label (e.g. "Limelight Larry" for peacock), and (5) Use of a basic level category label (e.g. "bird" 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. For clarity, we present only the key findings and statistics here, but full model details can be found in the Supporting Information.

Results

We begin by confirming that our *a priori* divisions of animals into early and late age of acquisition 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 they children knew but observed evidence to the contrary (i.e. children made an incorrect selection).

Target animal difficulty

We first confirm that the animals predicted to be later learned were less likely to be marked "known" by the parents of children in our studies. As predicted, parents judged that their children knew 93% of the animals in the Early AoA category, and 35% of the animals in the Late AoA category, which were reliably different from each-other (β = -8.83, t = -4.18, $p < .001$). Parents' judgments for each target word are shown in the Supporting Information.

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%, $\beta = 2.07$, $t = 10.35$, $p < .001$). Children were above chance both for animals that parents thought they knew ($M = 75.08$, $\beta = 2.61$, $t = 10.93$, $p < .001$), and for animals that parents thought their children did not know ($M = 55.19$, $\beta = 1.23$, $t = 8.35$, $p < .001$). 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, knowledge of the target animal, and the (log) length of parents' expressions. We found that children with bigger vocabularies were more accurate in general ($\beta = 0.03$, $t = 3.19$, $p = .001$), and that children were more less accurate for animals whose names they did not know (β = -1.86, t = -4.55, $p < .001$). Longer referential expressions were associated with less accuracy for known animals (β = -0.40, t = -2.69, $p = .007$), but greater accuracy for unknown animals ($\beta = 0.46$, $t =$

2.24, p = .025).

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

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 length of parents' referring expressions as a proxy for informativeness.

If parents tune their referring expressions to children's knowledge, they should produce more informative–and thus longer–referring expressions when they think their children will need them. To test this hypothesis, we divided every trial of the game into two phases: the time before a child selected an animal, and the time following selection until the start of the next trial. We then fit a mixed effects model predicting the number of words parents produced (log) from phase (before vs. after selection), target appearance (first vs. second), and three potential measures of tuning: (1) The total number of words the parent thought their child knew, (2) the proportion of all children whose parents reported they knew each target animal, and (3) whether each individual parent thought their child knew each individual animal word. We also estimated the interaction of each of these variables with phase. The first two measures of tuning represent what parents may be considering when

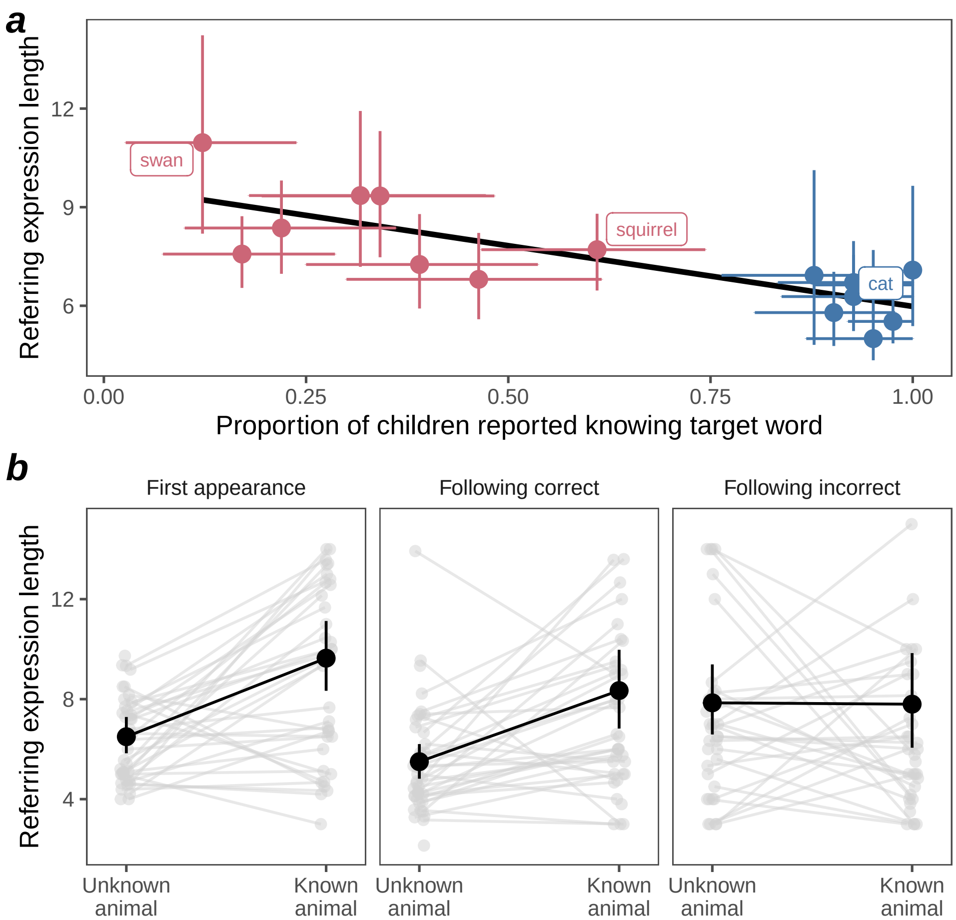

Figure 2. (A) Parents produced longer referring expressions to communicate about animals that children were less lkely to know. Points show group averaged proportions, error bars show 95% confidence intervals computed by non-parametric bootstrap. (B) Parents produced longer referring expressions for words that they thought their child did not know (left panel). When children selected correctly in response to these expressions, parents continued to produce longer expressions for animals they thought their children did not know. However, if children selected incorrectly for animals that parents thought they knew, parents produced longer expressions on their second appearance (right 2 panels). 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.

they coarse-tune their language, whereas the third measure pertains to fine-tuning. The final model included random intercepts and slopes of word-knowledge for subjects, and random intercepts and slopes of appearance for items.

When do parents produce longer referring expressions? In line with our predictions, we found that parents used more words when describing animals that believed their individual chidlren did not know (β = -0.25, t = -3.21, p = .003). They also produced reliably fewer words for animals that more children were reported to know (β = -0.17, t = -2.32, p .034). Thus, as parents were guiding their children to select the correct target animal, they provided more information for animals that were generally known by fewer children (coarse tuning; Figure 2A), but over and above that provided more information for animals that they believed their individual child did not know (fine tuning; Figure 2B). Overall, parents produced reliably fewer words after children selected an animal (β = , t = , p). This is partially due to the fact that a new trial begins almost immediately after children make a selection, thus providing parents very little time to comment on the previous trial. The inclusion of post-selection utterances did not change our results.

Before children selected an animal, parents produced reliably fewer words on the second appearance of each animal (β = -0.22, t = -5.86, $p < .001$), reliably fewer words for animals that more children were reported to know (β = -0.17, t = -2.32, p .034), and reliably more words for animals that they believed their individual child did not know (β = -0.25, t = -3.21, p = .003). Children's total vocabularies did not reliably affect the number of words parents produced (β = 0.00, t = -0.54, p = .595). After children selected an animal, parents produced reliably fewer words (β = , t = , p), but this reduction was smaller on an animal's second appearance (β = , t = , p), smaller for animals known by more children (β = , t = , p), and bigger for children who knew more words (β = , t = , p). The number of words produced after selection did not vary with parents' beliefs about their child's knowledge of that individual animal (β = , t = , p =). Thus, when parents were trying to get their children to select the correct target animal, they provided more information for animals that were generally known by fewer children (coarse tuning; Figure 2A), but over and above that provided more information for animals that they believed their individual child did not know (fine tuning; Figure 2B). In addition, parents produced fewer words after selection for children who knew more words, perhaps because they needed less support and reinforcement.

We found that parents' referring expressions on the second appearance of each animal were affected by both measures of coarse tuning: The child's total vocabulary and the proportion of all children who knew that animal. They were not, however, affected by parents' beliefs about their child's knowledge of that animal. Why not? 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 produce a longer referring expression for the leopard the second time around.

We found that parents selectively produce longer utterances on the second appearance of an animal when they believed their children knew that animal, *and* when they observed an incorrect selection on the first appearance. We fit a model predicting the length of parents' referring expressions on the second appearance of each animal from success on first appearance, whether parents thought their child knew the animal prior to the experiment, and their interaction. The final model included random intercepts and slopes of prior belief by subject and a random intercept for each animal. Parents produced marginally shorter referring expressions when children were incorrect on the first appearance of each animal ($\beta = 0.33$, $t = 4.04$, $p = < .001$), and shorter referring expressions for animals that they believed their child knew ($\beta =$, $t =$, $p =$). However, they produced longer referring expressions following an incorrect response for animals they thought their children knew ($\beta =$, $t =$, p). Thus, 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. Importantly, this same pattern was not found in unknown animals–parents' referring expressions did not differ in length for known and unknown animals when children had incorrectly selected on the first appearance (Figure 2B).

Together, these two sets of analyses suggest that parents tune their referring expressions not just coarsely to their knowledge about how hard individual animal words are, or how much language their children generally know, but also finely to their beliefs about their individuals children's knowledge of specific lexical items. Further, when they 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

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, quantitiative 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 subordinate category label (e.g. "Limelight Larry" for peacock), and (5) Use of a basic level category label (e.g. "bird" for peacock). Because the rates of usage of each of these kinds of reference varied widely (e.g. canonical labels were used on 94.82% of trials, but subordinates were used on 3.66% 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 and random intercepts for subjects and animals. Canonical labels were used on almost all trials, and did not differ in frequency between unknown (M = 95.92%) and known (M = 94.48%) animals (β = -0.10, t = -0.35, p = .724). Comparisons were used reliably more for unknown (M =

7.12%) than for known (M = 5.17%) animals (β = -2.15, t = -2.87, p = .004), as were descriptors (known M = 3.18%, unknown M = 19.37%, β = -3.08, t = -5.31, p < .001). Basic category labels were used marginally more for unknown (M = 8.77%) than known (M = 2.59%) animals (β = -2.29, t = -1.68, p = .092), and subordinates were used marignally less for unknown (M = 2.79%) than for known (M = 5.02%) animals (β = 2.18, t = 2.34, p = .092). Thus, parents used a variety of strategies to refer to animals that 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 they draw on this knowledge to tune their utterances 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, the lengths of parents' utterances are tuned 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. We further found that 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). Our work thus highlights the importance of studying the parent-child dyad as a unit, rather than viewing children as isolated learners. Children bring powerful learning mechanisms to language acquisition, but these mechanisms are supported by an ecological niche designed for their success (West & King, 1987).

Acknowledgements

This research was funded by a James S. McDonnell Foundation Scholar Award to DY.

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.

Nelson, D. G. K., Hirsh-Pasek, K., Jusczyk, P. W., & Cassidy, K. W. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language*, *16*(1), 55–68.

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.

West, M. J., & King, A. P. (1987). Settling nature and nurture into an ontogenetic niche. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *20*(5), 549–562.

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