Parents Calibrate Speech to Their Children's Vocabulary Knowledge

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 code for these analyses are available at \url{https://github.com/ashleychuikay/animalgame}

Correspondence concerning this article should be addressed to Ashley Leung, 5848 S University Ave, Chicago IL, 60642. 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. Do caregivers modify 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 calibrate their referring expressions to their children's language knowledge, 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 calibrate their referring expressions accordingly. These results underscore the importance of understanding the communicative context in which language learning happens.

Keywords: parent-child interaction; language development; communication

Word count: X

Parents Calibrate Speech to Their Children's Vocabulary Knowledge

Intro

In just a few short years, children develop a striking mastery of their native language. Undoubtedly, a large share of the credit for this remarkable feat is due to the powerful learning mechanisms that children bring to bear on their input (DeCasper & Fifer, 1980; Saffran, Aslin, & Newport, 1996). But, part of the credit may also may also be due to the structure of the language input itself. Indeed, individual differences in the quantity and quality of language children hear reliably related to individual differences in language learning (Hart & Risley, 1995; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Rowe, 2012). Further, ambient speech in the child's environment has little predictive power; the child-directed speech that occurs in children's interactions with their caregivers appears to be speech that matters (Romeo et al., 2018; Weisleder & Fernald, 2013). What makes child-directed speech so powerful?

Child-directed speech differs from adult-directed along a number of dimension, the majority of which are characterized by simplification (Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989; Snow, 1977). But, child-directed speech changes over development, with parents' producing longer and more complex utterances as their children grow older (Huttenlocher et al., 2010). Thus, child-directed speech may support learning not because it is simpler, but instead because it changes as children change: Caregivers may tune their speech to just the right level of complexity for children's ongoing language development (Snow, 1972; Vygotsky, 1978). One possibility is that this tuning might happen at a *coarse* level: Parents might calibrate the global complexity of their speech to their estimate of their child's global linguistic development. Alternatively, parents might *fine*-tune their speech, calibrating the way they talk about specific lexical items to what their children know about those same specific items. Fine-tuned speech would be a much more powerful vehicle for learning.

To date, almost all of the evidence of tuning has been found at a coarse level. For instance, the lengths of parents utterances, their articulation of parents' vowels, and the diversity of clauses in parents' speech change as children's speech changes (Bernstein Ratner, 1984; Huttenlocher et al., 2010; Moerk, 1976). The only evidence for fine tuning comes from two observational studies: One showing that parents are more likely to proivde their child with labels for novel as compared to familiar toys (Masur, 1997), and the second showing that the lengths of parents utterances in a high-density longitudinal recording dropped to their shortest just before the target child first produced those words (Roy, Frank, & Roy, 2009).

In this paper, we present the first experimental evidence that parents fine-tune their speech for individual lexical items. Parents played a reference game with their children in which their goal was to get them to pick the correct target animal from a set of three. The length of parents' utterances reflected independent contributions from (1) the difficulty of the target animal word, (2) their global estimate of their child's vocabulary, and (3) their estimate of their child's knowledge of that particular 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.

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 46 parent-child pairs were recruited, but data from six pairs were dropped from analysis due to experimental error or failure to complete the study. 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 broadly representative of the racial composition of the Chicago Area and the US more broadly (56% White). 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 learning. To obtain these estimates, we used two sources of information: Parent-report estimates of children's age of acqusition from Wordbank (Frank, Braginsky, Yurovsky, & Marchman, 2017), and retrospective self-report estimates of age of acquisition from adults (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012, see Supporting Information for details). The age of aquisition 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 low AoA or high AoA category.

A modified version of the MacArthur-Bates Communicative Development Inventory (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 analysis as we could not obtain individual-level estimates of children's knowledge.

Design and Procedure

Each parent-child pair played an interactive game using two iPads. Children were given two warm-up trials to get used to the iPads. The practice and experimental trials began after the warm-up. 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 encourage them to choose the object corresponding to the word on their screen. The child was instructed to listen to their parent for cues. Once an animal was tapped, the trial ended, and a new trial began. There was 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

The data of interest in this study were parent utterances used during the interactive game and parents' responses on the adapted CDI. Transcripts of the videos were analyzed for length of referring expressions. We measured the length of parents' referring utterances as a proxy for amount of information given in each utterance. Subsequently, utterances were manually coded for the following: use of canonical labels, basic category labels, subordinate category labels, descriptors, and comparison to other animals. Parent utterances irrelevant to the iPad game (e.g. asking the child to sit down) were not analyzed. Children's utterances were coded when audible, but were not analyzed.

Results

Word difficulty.

We first confirm that the animals predicted be later learned were less likely to be marked known by the parents of children in our studies. As predicted, animals in the Early AoA category were judged to be understood by 93% of parents, and items in the Late AoA category were judged understood by 35%. We confirmed this difference statistically with a mixed effects logistic regression, predicting success on each trial from a fixed effect of type and a random intercept ans slope of type by subject as well as a ranom intercept for each animal. The Late AoA items were judged known by a significantly smaller proportion of parents (β = -8.83, t = -4.18, p < .001). Parents' judgments for each target word are shown in Figure 1A.

Children's accuracy at selection

On the whole, parents effectively communicated with their children, getting them to select the correct target on 69.05% of trials. To determine whether this was reliably greater than we would expect by chance (33%), we fit a mixed effects logistic regression predicting whether each selection was correct from a fixed intercept, random intercepts for subjects and animals and an offset of log(1/3) so that the intercept estimated difference from chance. The interecept was significantly greater than 0 (β = 2.19, t = 9.06, p < .001), indiciating that children were selecting the correct animal at greater than chance levels. Children were above chance both for animals that parents thought they knew (M = 75.08, β = 2.95, t = 7.36, p < .001), and for animals that parents thought their children did not know (M = 55.19, β = 0.98, t = 2.15, p = .032). Thus, parents successfully communicated the target referent to children, even when parents thought children did not know the name for the animal at the start of the game.

To determine whether variation in parents' utterances impacted children's selections, we fit a mixed effects logistic regression predicting whether children were successful on each trial from the (log) length of parents' referring expression, parents' estimate of whether their child knew the animal, the appearance number of the target animal (first vs. second), the child's (scaled) total vocabulary, and interactions between parents' estimates of whether their children knew the animal and length, apperance, and vocabulary. We began with a maximal random effect structure and removed interactions until the model converged, leading to random intercepts and slopes of whether the child knew the animal for each subject and random intercepts for animals. For known animals, this model showed children were less accurate for longer utterances (β = -1.37, t = -9.49, $p < .001$), animals that parents thought they did not know (β = -3.12, t = -8.97, $p < .001$), and on second appearances (β = -0.20, t = -4.10, $p < .001$). Children with bigger vocabularies were more accurate ($\beta = 1.00$, $t = 4.27$, $p < .001$). For unknown animals, children were more accurate for when parents produced longer utterances ($\beta = 1.24$, $t = 6.22$, $p < .001$), on animals' second appearance ($\beta = 0.29$, $t = 4.13$, $p < .001$), and the effect of having a larger vocabulary was reduced (β = -0.75, t = -3.38, $p = .001$). Thus, longer utterances were associated with more successful referential communication for animals that children did not know, but were unhelpful for animals that they did know. We next ask whether parents calibrated these lengths appropriately.

Testing the tuning hypothesis

If parents calibrate their referential 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. In all of these cases, parents would be required to produce more words (see below for further qualitative analyses). Thus, we analyzed the length of parents' referential expressions as a

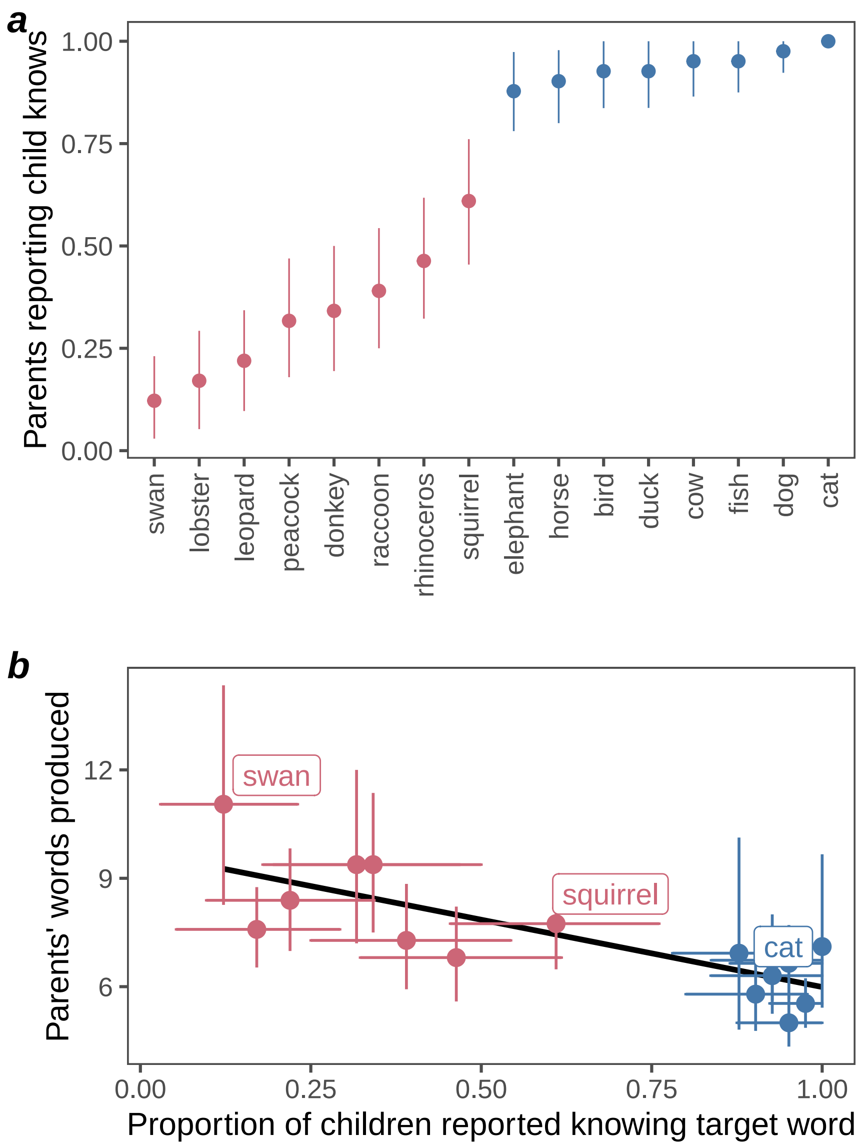

Figure 1. (A) Proportion of parents who reported that their child understood the word for each of our target animals. Colors indicate apriori categorization of words into Early (blue) and Late (red) age of acquisition. (B) Number of words in parents' referential expressions as a function of the proportion of children reported to know the word for target animal. Points show group averaged proportions, error bars show 95% confidence intervals computed by non-parametric bootstrap.

theory-agnostic 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), 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 word. We also estimated the interaction of each of these variables with phase. We began with a maximal random effect structure and removed random effects until the model converged, prioritizing variables of greatest theoretical for subjects and design-relevant variables for items. The final model included random intercepts and slopes of individual-child knowledge estimates for subjects and random intercepts and slopes of appearance for items.

Before children had selected an animal, parents produced reliably fewer words on the second appearance of each animal (β = -0.12, t = -5.72, $p < .001$), reliably fewer words for animals that more children were reported to know (β = -0.19, t = -4.39, $p < .001$), and reliably more words for animals that they believed their individual child did not know (β = 0.16, t = 3.42, p = .001). Children's total vocabularies did not reliably affect the number of words parents produced (β = 0.00, t = -0.90, p = .373). After children had selected an animal, parents produced reliably fewer words (β = -0.48, t = -11.31, $p < .001$), but this reduction was smaller on an animal's second appearance (β = 0.08, t = 7.43, $p < .001$), smaller for animals known by more children (β = 0.25, t = 10.22, $p < .001$), and bigger for children who knew more words (β = 0.00, t = , $p < .001$). The number of words produced after selection did not vary with parents beliefs about their child's knowledge of that individual animal (β = -0.02, t = -1.01, p = .312). 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 1B), but over and above that provided more information for animals that they believed their individual child did not know (fine tuning; 2A)). In addition, parent produced fewer words after selection for children who knew more words, perhaps because they needed less support and reinforcement.

We found that parents referential 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 get 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. To test this hypothesis, we fit a mixed effects model predicted the length of parents' referring expressions on the second appearance of each animal from success on first appearance, phase, (before vs. after selection), whether parents thought their child knew the animal prior to the experiment, and all interactions. We followed the same approach with random effects, beginning with a maximal model and pruning effects until the model converged. The final model included random intercepts and slopes of prior belief by subject and random intercepts and slopes by phase for each animal. Before children had selected a target, parents produced shorter referring expressions when children were incorrect on the first appearance of each animal (β = -0.15, t = -2.17, p = .030), and shorter referring expressions for animals that they believed their child knew (β = -0.25, t = -3.39, p = .001). However, they produced longer referring expressions following an incorrect response for animals they thought their chilren knew (β = 0.41, t = 4.16, $p < .001$). After children had selected a target, parents produced fewer words (β = -0.73, t = -8.41, $p < .001$), but this reduction was smaller for animals that their parents thought their children knew when they were correct on the first appearance (β = 0.28, t = 2.79, p .005), and reliably longer for animals thought their children knew but were incorrect on the first apperance (β = -0.55, t = -3.42, p .001). Thus, when parents thought their children knew an animal, but they observed evidence that they did not, they provided more information in their referential expressions for children to make the correct selection the second time. In fact, parents referential expressions were indistinguishable 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 referential expressions not just coarsely to their knowledge about how hard individual animal words are, or how much language their children generally but know, but also finely to their beliefs about their children's knowledge of individual 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.

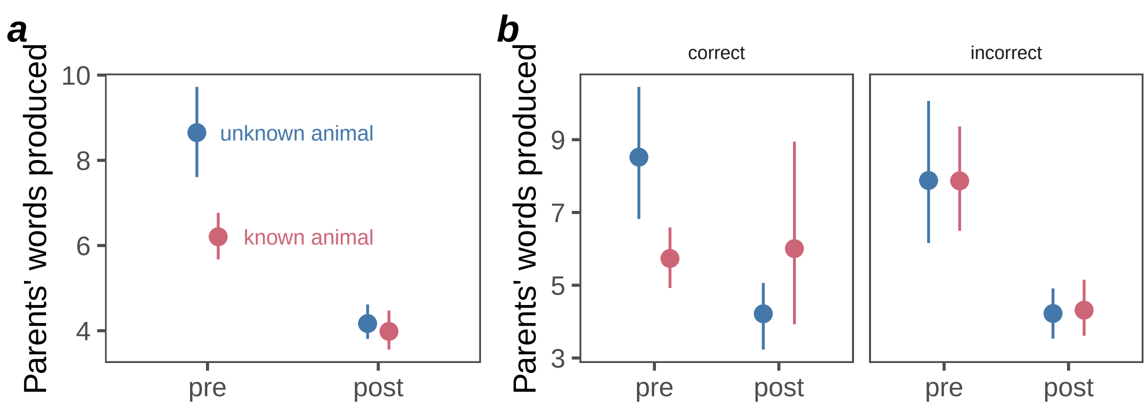

Figure 2. (A) Number of words parents produced before and after their child selected a target animal when they believed their child knew or did not know the animal. (B) LNumber of words parents produced on the second appearance of each animal when the child was correct vs. incorrect on its first appearance. Points show group averaged proportions; error bars show 95% confidence intervals computed by non-parametric bootstrap.

How refering expressions changed

Parents produced reliably longer referring expressions when trying to communicate about animals that they thought their children didn't know. We used length as a theory-agnostic, quantitiative measure of information. But how did parents successfully refer to animals that their children did not know. As a post-hoc descriptive analysis, we coded four 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"), Use of a subordinate category (e.g. "Limelight Larry" for peacock), And use of a superordinate category (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, $p < .001$). Superordinates 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 refer to animals that children did not understand, but the use of descriptors was the most prominent.

Discussion

Parents have a wealth of knowledge about their children's linguistic development (Fenson et al., 2007). Do they draw on this knowledge when they want to communicate? In a referential communication task, we showed that parents references are tuned 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 their utterances are calibrated to their individual children's knowledge of individual animals, and (3) when children do not know an animal that parents thought they did, their subsequent references reflect this update 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 are rich with desriptors and comparisons, helping children to select the correct animal and potentially also serving as a source of learning input.

These data are consistent with a strong form of the linguistic tuning hypothesis, in which parents produce language optimized to children's knowledge at a highly correlated level. Why should this happen? Although parents speech to children is unlikely to reflect a goal to teach, it is nonetheless goal oriented: It is motivated by a desire to communicate successfully (Bruner, 1983). However, communication and learning are inextricably linked–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: both parents and children contribute to the process of language development (Brown, 1977; Hoff-Ginsberg & Shatz, 1982). Focusing on the interactive and communicative nature of language captures a more realistic picture of children's language environments: The input that children receive is not random – it is sensitive to their developmental level.

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*.

Brown, R. (1977). Introduction. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and interaction*. Cambridge, MA.: MIT Press.

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

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, *208*(4448), 1174–1176.

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.

Hoff-Ginsberg, E., & Shatz, M. (1982). Linguistic input and the child's acquisition of language. *Psychological Bulletin*, *92*(1), 3–26.

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.

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.

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.

Snow, C. E. (1977). The development of conversation between mothers and babies. *Journal of Child Language*, *4*(1), 1–22.

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.