1

5

Parents Fine-tune Their Speech to Children's Vocabularies

Ashley Leung¹, Alexandra Tunkel¹, & 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
- ⁷ Society: Leung et al. (2019). All code for these analyses are available at
- 8 https://github.com/ashleychuikay/animalgame.
- ⁹ 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 12 powerful statistical learning mechanisms, the statistics they encounter are also prepared for 13 them: Children learn from caregivers motivated to communicate with them. Do caregivers 14 fine-tune their speech in order to support children's comprehension? We asked children and 15 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 17 referring expressions to their children's knowledge at the lexical-item level, producing more 18 informative references for animals that they thought their children did not know. Further, 19 parents learn about their children's knowledge over the course of the game, and tune their 20 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

26

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 remarkable feat is due to the powerful

learning mechanisms that children bring to bear on their input (DeCasper & Fifer, 1980;

Saffran, Aslin, & Newport, 1996). However, part of the credit may also may also be due to

the structure of the language input itself. In line with this hypothesis, individual differences

in the quantity and quality of language children hear are reliably related to individual

differences in language learning (Hart & Risley, 1995; Huttenlocher, Waterfall, Vasilyeva,

Vevea, & Hedges, 2010; Rowe, 2012). Further, ambient speech in children's environments has

little predictive power; the child-directed speech in children's interactions with their

caregivers is what matters most (Romeo et al., 2018; Weisleder & Fernald, 2013). What

makes child-directed speech so powerful?

Child-directed speech differs from adult-directed speech 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 also changes over development, with
parents producing longer and more complex utterances as children develop (Huttenlocher et
al., 2010). Child-directed speech may scaffold learning not because it is simpler, but because
it is adaptive—Caregivers may tune their speech to the right level of complexity for children's
ongoing language development (Snow, 1972; Vygotsky, 1978). One possibility is that parents
coarse-tune their speech, calibrating global complexity to children's global linguistic
development. Alternatively, parents could fine-tune their speech, calibrating the complexity
of language containing specific lexical items to children knowledge of those same lexical items.
Fine-tuned speech would be a particulary powerful vehicle for language learning, and thus an
important part of the explanation for rapid children's rapid language learning (Bruner, 1983).

Parents have been shown to coarse-tune several aspects of child-directed speech; the lengths of parents utterances, the articulation of parents' vowels, and the diversity of clauses

in parents' speech change over children's language development. (Bernstein Ratner, 1984;

Huttenlocher et al., 2010; Moerk, 1976). However, 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 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).

Here, we present the first experimental evidence for fine-tuning. We asked parents and their children to play a reference game in which the parent's goal was to guide their child to select the correct 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 in they provide.

68 Method

69 Participants

76

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 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 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).

31 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 low AoA or high AoA
category. This separation was designed to lower the likelihood that children could use
knowledge of low AoA animals to infer the correct target on high AoA trials.

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 began 102 with two warm-up trials on which they tapped on circles that appeared on the iPads. 103 Following these warmup trials, children and their parents moved onto practice and then 104 experimental trials. On each trial, three images of animals were displayed side by side on the 105 child's screen, and a single word appeared on the parent's screen. Parents were instructed to 106 communicate as they normally would with their child, and to encourage their child to choose 107 the object corresponding to the word on their screen. The child was instructed to listen to 108 their parent for cues. Once the child tapped an animal, the trial ended, and a new trial 109 began. There were a total of 36 experimental trials, such that each animal appeared as the 110 target twice. Trials were randomized for each participant, with the constraint that the same 111 animal could not be the target twice in a row. Practice trials followed the same format as 112 experimental trials, with the exception that images of fruit and vegetables were shown. All 113 sessions were videotaped for transcription and coding. 114

115 Data analysis

116

117

118

119

125

Ou 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.

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 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 producd each of the 85 words on the survey. In addition to analyzying 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.

130 Results

We begin by confirming that our apriori divisions of animals into low and high 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 present several tests of the fine-tuning hypothesis.

135 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, 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 and slope of type by subject as well as a random intercept for each animal. The Late AoA items were judged known by a significantly smaller proportion of parents ($\beta = -8.83$, t = -4.18, p < .001). Parents' judgments for each target word are shown in Figure 1A.

5 Selection accuracy

On the whole, parents communicated effectively 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 intercept was significantly greater than 0 ($\beta = 2.19$, 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, $\beta = 2.95$, t = 7.36, p < .001), and for animals that parents thought their children did not know (M = 55.19, $\beta = 0.98$, 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, 158 we fit a mixed effects logistic regression predicting whether children were successful on each 159 trial from the (log) length of parents' referring expression, parents' estimate of whether their 160 child knew the animal, the appearance number of the target animal (first vs. second), the 161 child's (scaled) total vocabulary, and interactions between parents' estimates of whether 162 their children knew the animal and length, apperance, and vocabulary. We began with a 163 maximal random effect structure and removed interactions until the model converged, 164 leading to random intercepts and slopes of whether the child knew the animal for each 165 subject and random intercepts for animals. For known animals, this model showed children 166 were less accurate for longer utterances ($\beta = -1.37$, t = -9.49, p < .001), animals that parents 167 thought they did not know ($\beta = -3.12$, t = -8.97, p < .001), and on second appearances ($\beta =$ 168 -0.20, t = -4.10, p < .001). Children with bigger vocabularies were more accurate ($\beta =$ 1.00, 169 t = 4.27, p < .001). For unknown animals, children were more accurate for when parents 170 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 $(\beta =$ 172 -0.75, t = -3.38, p = .001). Thus, longer utterances were associated with more successful 173 referential communication for animals that children did not know, but were unhelpful for 174 animals that they did know. We next ask whether parents tuned the lengths of their 175 utterances appropriately, producing longer utterances for unknown animals. 176

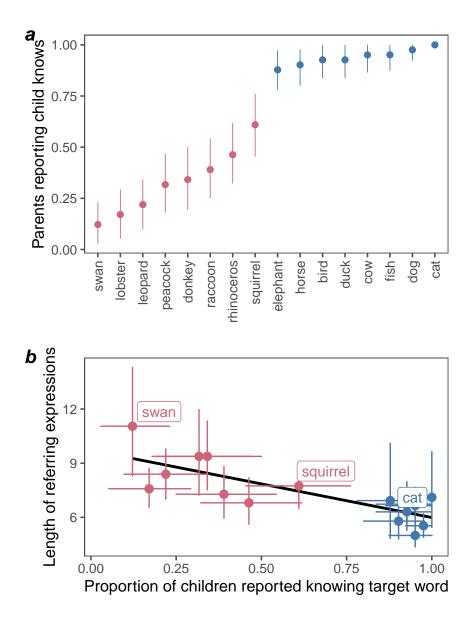


Figure 1. (A) Proportion of parents who reported that their child understood the word for each of our target animals. Colors indicate apriori categorization of words into Early (blue) and Late (red) age of acquisition. (B) Number of words in parents' referential expressions as a function of the proportion 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.

177 Tuning

If parents calibrate their referential expressions to their children's linguistic knowledge, 178 they should provide more information to children for whom a simple bare noun (e.g. 179 "leopard") would be insufficient to identify the target. Parents did this in a number of ways: 180 With one or more adjectives (e.g., "the spotted, yellow leopard"), with similes (e.g., "the one 181 that's like a cat"), and with allusions to familiar animal exemplars of the category. In all of 182 these cases, parents would be required to produce more words (see below for further 183 qualitative analyses). Thus, we analyzed the length of parents' referential expressions as a 184 theory-agnostic proxy for informativeness. 185

If parents tune their referring expressions to children's knowledge, they should produce 186 more informative—and thus longer—referring expressions when they think their children will 187 need them. To test this hypothesis, we divided every trial of the game into two phases: The 188 time before a child selected an animal, and the time following selection until the start of the 189 next trial. We then fit a mixed effects model predicting the number of words parents 190 produced (log), phase (before vs. after selection), target appearance (first vs. second), and 191 three potential measures of tuning: (1) The total number of words the parent thought their 192 child knew, (2) the proportion of all children whose parents reported they knew each target 193 animal, and (3) whether each individual parent thought their child knew each individual 194 word. We also estimated the interaction of each of these variables with phase. We began with 195 a maximal random effect structure and removed random effects until the model converged, 196 prioritizing variables of greatest theoretical relevance for subjects and design-relevant 197 variables for items. The final model included random intercepts and slopes of individual-child 198 knowledge estimates for subjects and random intercepts and slopes of appearance for items. 199

Before children selected an animal, parents produced reliably fewer words on the second appearance of each animal ($\beta = -0.12$, t = -5.72, p < .001), reliably fewer words for animals that more children were reported to know ($\beta = -0.19$, t = -4.39, p < .001), and

reliably more words for animals that they believed their individual child did not know (β = 203 0.16, t = 3.42, p = .001). Children's total vocabularies did not reliably affect the number of 204 words parents produced ($\beta = 0.00, t = -0.90, p = .373$). After children selected an animal, 205 parents produced reliably fewer words ($\beta = -0.48$, t = -11.31, p < .001), but this reduction 206 was smaller on an animal's second appearance ($\beta = 0.08, t = 7.43, p < .001$), smaller for 207 animals known by more children ($\beta = 0.25$, t = 10.22, p < .001), and bigger for children who 208 knew more words ($\beta = 0.00, t = -7.89, p < .001$). The number of words produced after 209 selection did not vary with parents beliefs about their child's knowledge of that individual 210 animal ($\beta = -0.02$, t = -1.01, p = .312). Thus, when parents were trying to get their children 211 to select the correct target animal, they provided more information for animals that were 212 generally known by fewer children (coarse tuning; Figure 1B), but over and above that 213 provided more information for animals that they believed their individual child did not know (fine tuning; Figure 2). In addition, parents produced fewer words after selection for children 215 who knew more words, perhaps because they needed less support and reinforcement. 216

We found that parents' referential expressions on the second appearance of each animal 217 were affected by both measures of coarse tuning: The child's total vocabulary and the 218 proportion of all children who knew that animal. They were not, however, affected by 219 parents' beliefs about their child's knowledge of that animal. Why not? One possibility is 220 that parents get information from the first appearance of each animal: They may have 221 thought their child knew "leopard," but discovered from their incorrect choice that they did 222 not. If so, they might produce a longer referring expression for the leopard the second time 223 around. To test this hypothesis, we fit a mixed effects model predicting the length of parents' referring expressions on the second appearance of each animal from success on first 225 appearance, whether parents thought their child knew the animal prior to the experiment, 226 their interaction. We followed the same approach with random effects, beginning with a 227 maximal model and pruning effects until the model converged. The final model included 228 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 230 incorrect on the first appearance of each animal ($\beta = -0.13$, t = -1.77, p = .077), and shorter 231 referring expressions for animals that they believed their child knew ($\beta = -0.29$, t = -3.40, 232 p = .001). However, they produced longer referring expressions following an incorrect 233 response for animals they thought their chilren knew ($\beta = 0.39$, t = 3.87, p < .001). Thus, 234 when parents thought their children knew an animal, but then observed evidence to the 235 contrary, they provided more information in their referential expressions for children to make 236 the correct selection the second time. In fact, parents' referential expressions were 237 indistinguishable in length for known and unknown animals when children had incorrectly 238 selected on the first appearance (Figure 2). 239

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 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.

246 Content of referring expressions

Parents produced reliably longer referring expressions when trying to communicate 247 about animals that they thought their children did not know. In the analyses presented so 248 far, we used length as a theory-agnostic, quantitiative measure of information. How did 249 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 252 a comparison (e.g. "like a cat"), (4) Use of a subordinate category label (e.g. "Limelight 253 Larry" for peacock), and (5) Use of a basic level category label (e.g. "bird" for peacock). 254 Because the rates of usage of each of these kinds of reference varied widely (e.g. canonical 255

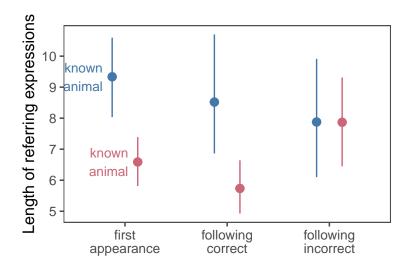


Figure 2. Length of parents' referring expressions on the first appearance each animal, and on the second appearance following correct vs. incorrect selections on the first appearance. When parents thought their child knew an animal, but they were incorrect on its first appearance, they provided more information on the second appearance. Points show group averaged proportions; error bars show 95% confidence intervals computed by non-parametric bootstrap.

labels were used on 94.82% of trials, but subordinates were used on 3.66% of trials), we fit a 256 logistic mixed effects model separately for each reference kind estimating whether it would 257 be used on each trial from whether the parent thought their child knew the animal and 258 random intercepts for subjects and animals. Canonical labels were used on almost all trials, 259 and did not differ in frequency between unknown (M = 95.92%) and known (M = 94.48%)260 animals ($\beta = -0.10$, t = -0.35, p = .724). Comparisons were used reliably more for unknown 261 (M = 7.12%) than for known (M = 5.17%) animals $(\beta = -2.15, t = -2.87, p = .004)$, as were 262 descriptors (known M = 3.18%, unknown M = 19.37%, $\beta = -3.08$, t = -5.31, p < .001). Basic category labels were used marginally more for unknown (M = 8.77%) than known (M = 2.59%) animals ($\beta = -2.29$, t = -1.68, p = .092), and subordinates were used 265 marignally less for unknown (M = 2.79%) than for known (M = 5.02%) animals ($\beta = 2.18$, 266 t = 2.34, p = .092). Thus, parents used a variety of strategies refer to animals that children 267 did not understand, but the use of descriptors was the most prominent. These descriptors 268

are particularly apt to facilitate children's learning, connecting parents' fine-tuning for reference with their children's language acquisition.

271 Discussion

Parents have a wealth of knowledge about their children's linguistic development 272 (Fenson et al., 2007). Do they draw on this knowledge when they want to communicate? In 273 a referential communication task, we showed that parents' references are tuned to their 274 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' 276 utterances are calibrated to their individual children's knowledge of individual animals, and 277 (3) when children do not know an animal that parents thought they did, parents' subsequent 278 references reflect this updated belief. We further found that more informative referring 279 expressions were associated with increased likelihood of successful communication: Children 280 were more likely to correctly select animals whose names they did not know if their parents 281 produced longer utterances to refer to them. Finally, we found that references to unknown 282 animals are rich with descriptors and comparisons, helping children select the correct animal 283 and potentially also serving as a source of learning input. 284

These data are consistent with a strong form of the linguistic tuning hypothesis, in 285 which parents fine-tune the information in their speech to children's language knowledge at 286 the individual-word level. Why should this happen? Although parents speech to children is 287 unlikely to reflect a goal to teach, it is nonetheless goal-oriented: Parents want to 288 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 292 isolated learners. Children bring powerful learning mechanisms to language acquisition, but 293 these mechanisms are supported by an ecological niche designed for their success (West &

295 King, 1987).

296

Acknowledgements

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

298 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.
- 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.
- 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.
- 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.