Exploring the Developmental Feedback Loop: Word Learning in Neural Networks and Toddlers

Clare E. Sims (Clare.Holtpatrick@Colorado.Edu)
Department of Psychology & Neuroscience, 345 UCB
Boulder, CO 80309-0345 USA

Savannah M. Schilling (Savannah.Schilling@ Colorado.Edu)
Department of Electrical, Computer & Energy Engineering, 425 UCB
Boulder, CO 80309-0425 USA

Eliana Colunga (Eliana.Colunga@Colorado.Edu)
Department of Psychology & Neuroscience, 345 UCB
Boulder, CO 80309-0345 USA

Abstract

Early word learning may be supported by a developmental feedback loop: the kind of words a child learns early on support the development of attentional biases, which in turn facilitate further word learning. In neural network simulations and a longitudinal study of toddlers we investigated how the emergence of an attentional bias to shape in word learning impacts vocabulary growth with respect to different kinds of words. If this relationship is causal, we should see that the emergence of a shape bias leads to an increase in the rate of learning of shape-based words relative to other kinds of words. The networks supported this prediction, showing an acceleration of shape- compared to material-based word learning. However, in toddlers, shape- and material-based words were learned similarly around the shape bias emergence. Implications are discussed for the developmental feedback loop account and causal relationships between attentional bias development and vocabulary growth.

Keywords: Word learning; shape bias; neural networks; longitudinal study.

Introduction

Children are skilled word learners, in part because of constraints on the range of things they consider when inferring the referent of a new word. These constraints, sometimes referred to as biases, operate by helping children direct attention, resulting in sensitivity to what information matters most in the context of learning different kinds of words. Although there is debate on the origin of these attentional biases (e.g., Samuelson & Bloom, 2008), evidence from children and from networks suggest that children can learn biases based on the kinds of nouns they acquire early on in their vocabularies (e.g., Colunga & Smith, 2005; Gerhshkoff-Stowe & Smith, 2004). This account entails a developmental feedback loop: the early nouns that children learn give rise to attentional biases, and those biases in turn guide further word learning and impact the structure of children's growing vocabularies. In the current paper we use data from neural networks and toddlers to investigate the latter part of this loop, focusing on how different types of words are learned right around the pivotal moment of word learning bias development.

Children's Early Vocabularies

Words are an important building block in language and cognitive development. Children make the process of word learning look deceptively simple, typically acquiring their first word around the age of 1 year and experiencing a spike in vocabulary development around 18 months of age (Goldfield & Reznick, 1990). Some researchers have observed that this vocabulary spike does not tend to occur until a child has at least 50 words in his or her vocabulary (Lucariello, 1987). Other work shows that the vocabulary spike is not only a function of the number of words a child knows, but also depends on the kinds of words that children learn. For example, Goldfield and Reznick (1990) observed that children exhibiting a vocabulary spike tended to add many words for objects (i.e., nouns) to their vocabularies. Children who did not show this dramatic increase in vocabulary size were steadily adding various types of words instead. This result suggests that while vocabulary size may be one key factor in children's language development, the specific kinds of words that children learn also play a role.

More recent research has investigated the question of why learning nouns may help accelerate subsequent vocabulary growth. One reason is that many different nouns have a basic property in common: they tend to refer to categories of things that are alike in shape. For example, a child will hear the word "ball" used to label a variety of spherical objects. Over time, children may pick up on the general pattern that shape is an important feature when talking about things in the world, and this insight in turn facilitates further word learning. Support for this account comes from a longitudinal study of young children's vocabularies (Gershkoff-Stowe and Smith, 2004). Over a three month period, 17-month-old children had a greater increase in object label nouns compared to other types of words. Over this same time period, children attended more to shape, over and above other features, when generalizing a newly learned word to novel objects. This result suggests that as children learn certain kinds of words, they also learn reliable patterns or constraints about how those words are used in the world. In

this paper we focus on one word learning constraint in particular: the shape bias.

The Shape Bias in Early Word Learning

The shape bias is the tendency for children to generalize newly learned nouns to other objects based on similarity in shape. This is typically tested in novel noun generalization (NNG) tasks (Landau, Smith, & Jones, 1988). In this type of task a child may be taught a novel name for a novel solid object. A shape bias is shown when the child extends that name to other objects matching the original in shape, even if the shape match differs from the original in texture, color, or size. Children show a reliable shape bias by 2 years of age (Samuelson & Smith, 1999).

There is evidence that the emergence of the shape bias can guide children in learning new words. For instance, in one study 17-month-olds were trained on shape-based categories of words, effectively teaching them the shape bias (Smith et. al, 2002). Not only did these children develop a shape bias earlier than the control group, they also showed accelerated growth in overall vocabulary. This suggests that there is an interaction between the development of word learning biases, particularly the shape bias, and vocabulary growth. This finding is one piece of evidence for a developmental feedback loop between vocabulary development and word learning constraints.

A Developmental Feedback Loop

Smith and colleagues (2002) showed that teaching children the shape bias can promote vocabulary growth, but what about the other way around? Many of the previously mentioned studies show a correlation between these two factors. However, rather than word learning biases simply causing vocabulary growth, perhaps these are coupled phenomena that reciprocally influence each other. Previous modeling research suggests this. For instance, Colunga and Sims (2012) trained neural networks with early- and late-talker vocabulary structures as input and then tested for the development of word learning biases. Results showed that networks given late-talker vocabulary input produced different biases than networks with early-talker input. This shows that given only the structure of a child's vocabulary, the network can develop quantitatively different biases, suggesting that vocabulary growth affects bias development. These findings, combined with the experiments of Smith and colleagues (2002), indicate that vocabulary structure and word learning biases may be part of a development feedback loop in which both processes affect one another. Here we investigate the dynamics of this loop in both neural networks and in children.

In prior work, we explored dynamics of and interactions between the shape bias and other word learning biases over time. Neural networks were trained on a typical 30-month-old child's vocabulary structure, then the bias dynamics were observed. We found that as the shape bias emerged, the development of other word learning biases diminished, suggesting a shift in attention as the shape bias is learned (Schilling, Sims & Colunga, 2012). These results were replicated in behavioral data from a longitudinal study of 18- to 30-month-old children (Sims, Schilling, & Colunga, 2012).

In this paper, we look at the same emergence window, but this time concentrate on how different kinds of words are learned before, during and after shape bias emergence. That is, we focus on the other piece of the developmental feedback loop: how vocabulary structure changes as word learning biases develop. What kinds of words do networks and children learn right around the pivotal point of shape bias emergence? To test this, we used network models and vocabulary data from a longitudinal behavioral study.

Approach and Overview

Our approach is to train a network on a typical early child vocabulary in order to observe learning over time that is similar to children's vocabulary development. We use a neural network to model the process of word learning, which differs from some other approaches to modeling word learning. For example, Bayesian networks extract generalities in order to produce a structured system representative of the real world (e.g., a mapping of a child's word representations; see Xu & Tenebaum, 2007). Our networks instead begin with structured representations as input and produce attentional biases. In order to investigate the developmental feedback loop, we are interested in the process: *how* the network forms these attentional biases from vocabulary structure input. We tested both networks and children on novel word generalization (using a virtual NNG task with the networks and a lab NNG task with children; see Sims et al., 2012 for details) over multiple points of vocabulary development. From this data, we identified the point in word learning at which the shape bias emerged for each individual network and child. Finally, we looked at the kinds of words that networks and children learned around their respective emergence points.

Network Simulations

Method

Our network was implemented in the Emergent software package (O'Reilly, Munakata, Frank, & Hazy, 2012), to model word learning. The network was given input structured like that of a typical 30-month-old's vocabulary. Throughout learning, we tracked what kinds of words the network successfully learned and tested for attentional biases. By analyzing the word learning biases the network developed and how they affected vocabulary structure, we were able to focus on the developmental feedback loop between attentional biases and vocabulary growth over time.

Network Dynamics

Our models use the Leabra algorithm (Local, Error-driven and Associative, Biologically Realistic Algorithm), which combines Hebbian and error-driven learning (O'Reilly et al., 2012). The Hebbian, self-organizing learning uses longer time-scale statistics about the environment and is useful for

extracting generalities. However, this type of learning is not as good at compensating for specific, complicated patterns. Therefore, we use error-driven learning, which actively utilizes differences between expectations and outcomes. The total weight change in the network is the sum of that of the error-driven learning and that of the Hebbian learning.

The network uses a function called the eXtended Contrastive Attractor Learning (XCAL) rule. This function uses sending and receiving activity input and has a floating threshold that regulates changes in weights over learning. This function is used for both the Hebbian and error-driven learning with different inputs to the function. Inputs affect threshold changes and therefore different inputs elicit different weight change dynamics.

The error-driven weight changes are updated based on the short-term average connection activity ($<xy>_s$) and the medium-time scale average connection activity ($<xy>_m$).

$$\Delta w_{error} = f_{xcal}(< xy >_s, < xy >_m) = f_{xcal}(x_s y_s, x_m y_m)$$

Where $<xy>_m$ represents the emerging expectation about a current situation and $<xy>_s$ reflects the actual outcome and therefore the result of the received error information.

The Hebbian weight changes are based on the short-term connection activity (xy_s) and long-term average activity of the receiving unit ($<y>_l$).

$$\Delta w_{Hebbian} = f_{xcal}(xy_s, x < y >_l) = f_{xcal}(xy_s, xy_l)$$

Based on $<y>_l$, the threshold for weight change is adjusted, making the weights more likely to change in the direction given by xy_s. This creates the structure of generalization for the Hebbian learning mechanism. The combination of these two types of learning mechanisms allows for a balance of feed forward information to form categories and back propagation to allow adjustments based on errors. For more details on network dynamics, see O'Reilly et al. (2012).

Network Architecture

The architecture is adapted from Colunga & Smith (2005) and is implemented as shown in Figure 1. Words are represented discretely and are input on the Word Layer. Features are represented as distributed patterns over several dimensions on the Perceptual Layer. For example, the shape and material of an object (e.g., the roundness of a particular ball and its yellow rubbery material) are represented by an activation pattern along the Perceptual Layer, with 12 units for shape and 12 units for material. Solidity is represented locally; one unit stands for Solid and another for Non-Solid. Finally, there is a 25 unit Hidden Layer that is connected to all the other layers and to itself. The Hidden Layer serves as the bridge between the Word Layer (the sending units) and the Perceptual Layer (the receiving units) and it is where learning occurs. Learning progresses as internal representations, or weights, update and form representations which connect the other two layers.

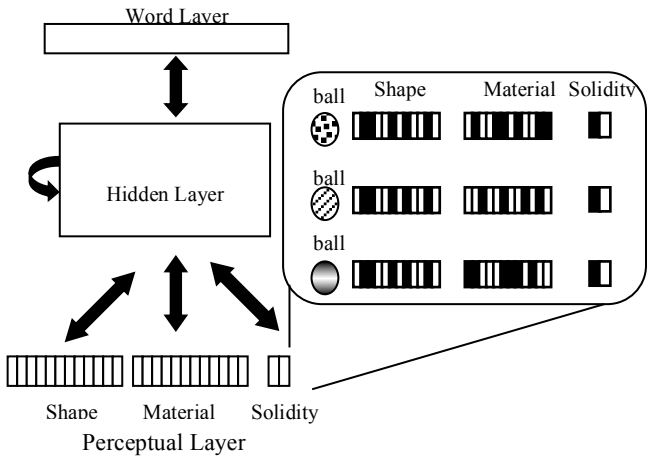

Figure 1. Network architecture and example input patterns.

Vocabulary Structure: Network Input Patterns

The input patterns used to train the network capture the structure of a child's vocabulary and are based on those used in Colunga and Smith (2005). They consisted of 100 noun representations, divided into 6 categories, with a structure analogous to the vocabulary of a typical 30-month-old child (Fenson et. al, 1993). Categories were divided by both solidity (solid or non-solid) and characteristic feature (shape, material, or both), based on adult judgments. From these, the structure of a typical early vocabulary can be expressed as proportions of each category. Therefore, the network learning the entire set of training patterns represents a child learning a typical vocabulary. See Table 1 for the 6 categories and proportions used in the study.

These input patterns have a correlational structure such that a network learning them should produce a shape bias for solids (and indeed this was first shown by Colunga & Smith, 2005). This means that learning in the network arises from the structure of the input patterns themselves. The purpose of the network, then, is not to help us discover structure in the input, but to observe the process of leveraging this structure over the course of word learning.

Table 1: Noun category percentages and example words.

	Shape	Material	Both
Solid	52%  *ball*	10%  *chalk*	12%  *penny*
Non-Solid	4%  *bubble*	16%  *milk*	6%  *jeans*

Network Training and Testing

Over the course of training, the network formed biases based on the structure of the vocabulary input. On each trial of training, a word was paired with a pattern of features representing the features of the noun category. For example,

a word for a solid item characterized by shape (like a ball) should be used to label things that are like each other in shape but differ from each other in material. To simulate this pattern, we randomly selected an input vector to represent, for example, ball shape. On individual training trials, we paired that shape pattern with the label ball and a randomly selected material pattern. Therefore over multiple training trials, a word for a solid item characterized by shape would be represented by the same shape but different material patterns (see Figure 1). We did this for each of the 100 nouns in the training set.

We tested 10 runs of the network at multiple points throughout word learning. Weights and words learned from each of the 6 categories of interest were recorded at thirteen discrete checkpoints during the course of each training run. For example, the network was tested at 5 words learned, 10 words learned, and so on. The endpoint of learning was at 500 epochs of training, which was around when the network learned 75 words. For more information on network testing procedures, see Schilling et al. (2012).

Rationale and Predictions

We focused our analyses of early child vocabulary composition, particularly shape-based and material-based words, on the period of time during which each network developed a shape bias in the context of solid objects. This approach may offer further insights into the relationship between attentional shifts in word learning and the course of vocabulary acquisition. As skilled attention to shape in the context of solid objects emerges, the networks should more easily learn shape-based words. Also, increased attention to shape may facilitate the learning of shape-based words over and above the learning of material-based words. This would be seen in a relatively lower rate of learning for material-based compared to shape-based words.

Results

The first question is how the networks learned shape-based words over the time window in which the shape bias emerged. The dependent measure was proportion of shape-based words learned at a given time point relative to the total number of shape-based words in the input vocabulary. Proportions of shape-based words learned were submitted to a linear regression with time point (before, at, or after shape bias emergence) as the predictor variable. Shape-based word learning increased significantly over time, $b = 0.06$, $t(28) = 7.70$, $p < .001$. The networks showed significant increases in proportions of shape-based words learned from before shape bias emergence ($M = .01$, $SD = .02$) to the point of emergence ($M = .06$, $SD = .03$), and from emergence to the following time point ($M = .14$, $SD = .06$; $t(9) > 4.60$, $p \leq .001$, Cohen's $d > 1.45$ for both paired comparisons). That is, the networks' learning of shape-based words increased over time, and showed a particularly large increase following the emergence of the shape bias.

The next question is whether the networks' learning of shape-based words differed from learning of material-based words over the same time period. Proportions of material-based words learned were similarly computed relative to the total number of material-based words in the input vocabulary. Proportions of words learned were submitted to a multiple regression including time, word type, and the interaction between the two. Overall, these variables explained a significant proportion of the variance in the networks' word learning, $R^2 = .69$, $F(3, 56) = 40.90$, $p < .001$. Consistent with the result above and the fact that the networks continually learned new words over time, time was a significant predictor of word learning overall, $b = .03$, $t(56) = 4.80$, $p < .001$. The networks showed increases in learning both shape- and material-based words over the time window of interest (see Figure 2). Word type was also a significant predictor of learning, in that the networks on average learned a greater proportion of shape-based than material-based words, $b = .03$, $t(56) = 4.03$, $p < .001$. Additionally, the interaction between time point and word type predicted learning, $b = .03$, $t(56) = 3.07$, $p < .01$. As can be seen in Figure 2, there was a steeper increase in the trajectory of learning for shape- compared to material-based words over the time window of interest.

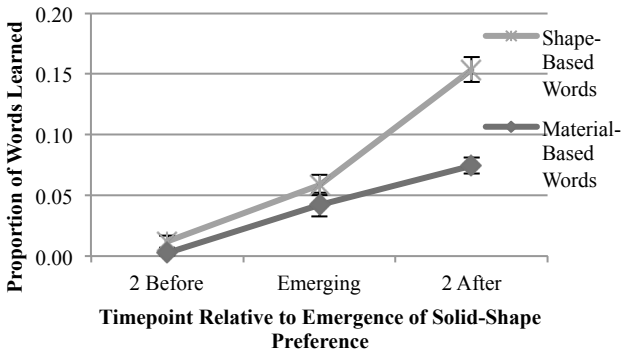

Figure 2. Shape- and material-based word learning in the network simulations over time.

Discussion

The results of the network simulations show that the emergence of the shape bias coincided with changes in vocabulary acquisition for different kinds of words. Before the emergence of the shape bias, the networks steadily increased the amount of both shape- and material-based words in their vocabularies to an equal extent. However, after the emergence of the shape bias, learning of shape-based words outpaced learning of material-based words. This result adds support to a developmental feedback loop account of word learning. Adding to previous work showing that networks can learn attentional biases from the vocabulary input of a typical toddler (Colunga & Smith, 2005; Schilling et al., 2012), the current study shows that in these same kinds of networks, attentional biases in turn influence the trajectories of subsequent vocabulary learning. Next we tested our network predictions in a behavioral study of toddlers.

Behavioral Study

Rationale and Predictions

To test the predictions of the network simulation we conducted a similar analysis on a longitudinal sample of toddlers. To explore vocabulary learning over time we looked at a parent-filled, standardized vocabulary inventory (the MacArthur-Bates Communicative Development Inventory [MCDI]; Fenson et al., 2007) that had been collected every month for a year for each child in the sample. As in the networks, we centered our analyses of child vocabulary development on the time at which each child first showed a shape bias for solid objects.

The network simulations predict that the emergence of the shape bias for solids leads to a change in the course of subsequent vocabulary learning. Specifically, this change was seen in the trajectory of shape-based relative to material-based word learning. If this prediction bears out in children, we should see a similar pattern in the toddler data.

Method

Participants

Nineteen children were recruited from the Boulder, CO area. Two children were excluded from the current analyses because they knew greater than 80% of the nouns in the MCDI at the beginning of the time window of interest. The final sample analyzed here included 17 children (M_{age} = 22.09 mo., SD = 2.69 mo., 9 girls).

Progression of Word Learning

Children participated in 12 monthly visits over the course of one year. At each visit children were tested for attentional biases in novel word learning. There were three different stimuli sets, each consisting of an exemplar, and five test items matching the exemplar in shape, material and/or color. The children saw a single set in each visit and thus rotated through the three sets every three months. We calculated the point of emergence of the shape bias as in Sims et al. (2012), for each individual child, as defined by the child having shown a persistent shape bias during three consecutive visits, that is, for all three stimuli sets.

Vocabulary development was measured longitudinally through parent-completed, monthly MCDI checklists of words their child knew at the time of each visit to the lab. We focused our analyses on children's noun learning over the time period of interest. At the beginning of the analysis windows, children had on average 108 nouns (SD = 84 nouns) from the MCDI in their vocabularies. To explore shape- and material-based word learning separately, we used the vocabulary structure classifications collected by Colunga and Smith (2005; see Table 1), combining solid and non-solid shape-based nouns, and solid and non-solid material-based nouns to get our two categories of interest.

Results

As in the network simulation analyses, the first question we investigated in the behavioral data was whether children's learning of shape-based words increased over the window during which each child developed a shape bias. The dependent measure was children's proportions of shape-based words learned at a given time point relative to each child's total number of shape-based words attained at the end of the study. These proportions were submitted to a linear regression with time point as the predictor. Shape-based word learning increased significantly over the time window of interest, b=0.13, $t(43) = 2.26$, $p = .03$. Post-hoc paired comparisons showed increases in words learned from before shape bias emergence (M = .50, SD = .34) to the point of emergence (M = .64, SD = .32), and to the time point after emergence (M = .76, SD = .23; $t(14) > 3$, $p < .01$, $d > .90$ for both).

Next we compared children's learning of shape-based words to their learning of material-based words over the same time window. Proportions of words learned were submitted to a multiple regression including time, word type, and the interaction between the two. Overall, these predictors explained a significant proportion of the variance in children's word learning, $R^2 = .13$, $F(3, 86) = 4.32$, $p < .01$. Time point was a significant predictor, showing that the proportions of both shape- and material-based words increased over the time window of interest, $b = .13$, $t(86) = 2.18$, $p = .03$. Word type was a marginally significant predictor of learning, $b = .12$, $t(86) = 1.86$, $p = .07$. As can be seen in Figure 3, the proportion of shape-based words learned (M = .63, SD = .32) tended to be higher than the proportion of material-based words learned (M = .51, SD = .34) across all time points. The interaction between time point and word type was not a significant predictor of learning. That is, children's learning of shape- and material-based words followed the same trajectory.

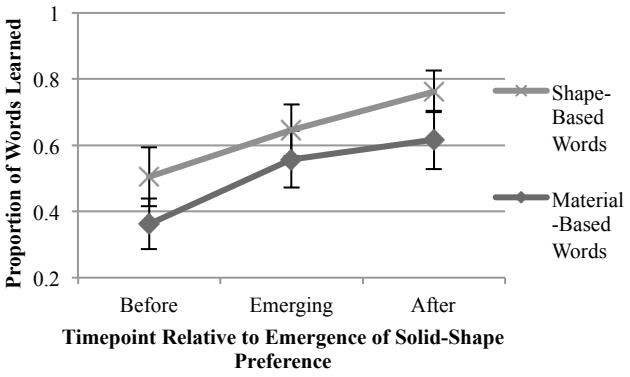

Figure 3. Shape- and material-based word learning in the behavioral study of toddlers over time.

Discussion

These results show that there is an increase in the number of shape based words that children learn as the shape bias emerges. This result is consistent with the networks and supports one piece of the developmental feedback loop in children. However, unlike the network simulations, the increases in children's word learning did not show a marked acceleration for shape-based words. Although proportionally more shape-based words were learned

compared to material-based words, the trajectory of learning for these two types of words did not differ significantly within this time window.

General Discussion

In the current studies we found that vocabulary learning around the emergence of the shape bias supported the developmental feedback loop account in our network simulations, but toddlers showed a different pattern. Adding to our previous work with these word learning networks, the current simulations contribute evidence for the effects of attentional biases on subsequent vocabulary learning. The behavioral data show ambiguous results in relation to the developmental feedback loop. There are several possible reasons for this pattern that will inform future research.

Methodological constraints could have contributed to these intriguing results. For example, the networks' performance in the generalization task is much more consistent than the children's performance. Even though individual networks do vary on the epoch at which they show a shape bias, once it emerges, it stays. This is not the case for children. To deal with this, we used a stringent criterion to define the time of emergence for the children by making sure that the preference for shape was present during three consecutive visits, for three different stimuli sets. Probably because of this criterion, the points of bias emergence that we identified tended to occur when children had on average over 100 nouns, with high variation between individuals. This suggests that we may have identified shape bias emergence relatively late in vocabulary development for some of the children in the sample, at least when compared with the criterion used in Gershkoff-Stowe & Smith (2004). A related possibility is that our network shows bias emergence and subsequent vocabulary changes at a relatively earlier (or "younger") point than the toddlers in our sample. If this is the case, vocabulary changes in the network may be easier to detect because it has progressed less far in the proportion of words learned and thus can statistically show greater growth compared to the toddlers. On the other hand, when the shape bias emerges in toddlers, they already know over half of the words in the MCDI, and thus have relatively less room for growth. Nonetheless, we would still expect to see differences in how toddlers learn shape- and material-based words, yet these interactions are either not present or not being captured by our current measures. In future analyses we plan to explore other measures such as rate of vocabulary growth that may better equate learning in the network and in toddlers. We also plan to look at dynamic attention to shape as a continuous measure over the entire trajectory of learning. Perhaps the emergence of the shape bias puts into motion long-term, rather than immediate, changes in the trajectory of vocabulary growth.

More interesting than methodological explanations are the theoretical implications of this finding. The behavioral results, along with our previous work (Sims et al., 2012), suggest that vocabulary growth precedes bias development, but the causality of this relationship may not go the other way. Perhaps once children have a consistent shape bias for solids, and those words come easy, they begin to focus on other kinds of words that do not conform to their expectations of naming categories organized by shape. Further work is necessary to see if that is the case. Overall, our novel approach using neural networks allows us to explore not just a causal effect of biases on vocabulary, but the dynamic feedback relationship between the two, the very relationship that builds the developmental trajectory.

References

Colunga, E., & Sims, C. E. (2012). Early-talker and late-talker toddlers and networks show different word learning biases. *Proc. of the 34th Ann. Conf. of the Cog. Sci. Society*. pp. 246-251. Austin, TX: Cog. Sci Society.

Colunga, E., & Smith, L. B. (2005). From the lexicon to expectations about kinds: A role for associative learning. *Psych. Review, 112,* 347-382.

Fenson, L., Dale, P., Reznick, J. S., Thal, D., Bates, E., Hartung, J., et al. (1993). The MacArthur-Bates Communicative Development Inventory: User's guide and technical manual. San Diego, CA: Singular Pub.

Gershkoff-Stowe, L., & Smith, L. B. (2004). Shape and the first hundred nouns. *Child Development, 75,* 1098-1114.

Goldfield, B. A., & Reznick, S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language, 17,* 171-183.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cog. Dev, 3,* 299-321.

Lucariello, J. (1987). Concept formation and its relation to word learning and use in the second year. *Journal of Child Language, 14,* 309-332.

O'Reilly, R. C., Munakata, Y., Frank, M. J., & Hazy, T. E. (2012). Computational Cognitive Neuroscience.

Samuelson, L. K., & Bloom, P. (2008). Special section: What counts as an explanation of development: The shape bias as a case study. *Developmental Science, 11,* 183-184.

Samuelson, L. K., & Smith, L. B. (1999). Early noun vocabularies: Do ontology, category structure and syntax correspond? *Cognition, 73,* 1-33.

Schilling, S.M., Sims, C.E., & Colunga, E. (2012). Taking development seriously: Modeling the interactions in emergence of different word learning biases. *Proc. of the 34th Ann. Conf. of the Cog. Sci. Society*. pp. 246-251. Austin, TX: Cog. Sci Society.

Sims, C. E., Schilling, S. M., & Colunga, E. (2012). Interactions in the development of skilled word learning in neural networks and toddlers. *Proc. of the ICDL/EpiRob.* pp. 1-6. San Diego, CA.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. K. (2002). Creating a shape bias creates rapid word learners. *Psych. Sci., 13,* 13-19.

Xu, F. & Tenebaum, J. B. (2007). Word Learning as Bayesian Inference. *Psych. Review, 114*, 245-272.