Statistical Regularities in Vocabulary Guide Language Acquisition in Connectionist Models and 15–20-Month-Olds

Larissa K. Samuelson
University of Iowa

This research tested the hypothesis that young children's bias to generalize names for solid objects by shape is the product of statistical regularities among nouns in the early productive vocabulary. Data from a 4-layer Hopfield network suggested that the statistical regularities in the early noun vocabulary are strong enough to create a shape bias, and that the shape bias is overgeneralized to nonsolid stimuli. A 2nd simulation suggested that this overgeneralization is due to the dominance of names for shape-based categories in the early noun vocabulary. Two subsequent longitudinal experiments tested whether it is possible to create word learning biases in children. Children 15–20 months old were given intensive naming experiences with 12 noun categories typical of the types of categories children learn to name early. The children developed a precocious shape bias that was overgeneralized to naming nonsolid substances; they also showed accelerated vocabulary development. Children taught an atypical set of nouns or no new nouns did not develop a shape bias and did not show accelerated vocabulary development.

Learning word meanings in a first language should be a difficult task. The number and range of possible meanings to be learned are immense and, by some arguments, indeterminate from the typical information provided to young children (Quine, 1960). Indeed, very young children have great difficulty learning their first words. In the early stages of word learning, children need to hear instances named many times before they produce a word on their own, and they often make mistakes when using their first words (Clark, 1973; Mervis, 1987; Mervis, Mervis, Johnson, & Bertrand, 1992). This picture has changed, however, by the time children reach their second birthdays. Indeed, studies suggest that between 18 and 30 months of age, the typical child's productive vocabulary increases tenfold (Fenson et al., 1994) and that children are especially smart noun learners. They are so smart that they seem to learn the whole category to which a novel name applies from hearing a single instance named (Imai & Gentner, 1997; Landau, Smith, & Jones, 1988; Markman, 1989; Samuelson & Smith, 2000; Soja, Carey, & Spelke, 1991; Waxman & Hall, 1993). This is particularly remark-

able given the many different kinds of entities children learn to name—objects, substances, people, animals, and places. How do they do it?

Theoretical answers to this question suggest that the task of learning nouns is made easier by biases or constraints that reduce the problem of finding the correct word–referent mapping to a solvable size (Landau et al., 1988; Markman, 1989, 1992; Soja et al., 1991). And indeed, when children are asked to learn words in laboratory settings, they appear biased to map novel words to unnamed rather than named objects, to whole objects rather than salient parts, and to taxonomically rather than thematically organized categories (Markman, 1992). They also reliably interpret novel count nouns as referring to categories of similar things (Katz, Baker, & Macnamara, 1974; Markman, 1992; Waxman & Booth, 2001), proper nouns as referring to individual entities (Hall, Lee, & Belanger, 2001; Macnamara & Reyes, 1994), and novel verbs as referring to actions rather than objects (Brown, 1957; Tomasello & Akhtar, 1995). Other research suggests that children are even biased to attend to specific perceptual features of novel objects—their shape or material substance, for example—when learning novel names.

Much of what we know about these attentional biases comes from experiments using novel noun generalization tasks (Brown, 1957; Imai, Gentner, & Uchida, 1994; Katz et al., 1974; Landau et al., 1988; Soja et al., 1991). In these tasks, a young child is presented with a novel object that is named (e.g., "This is a dax"). The child is then presented with novel test objects that match the exemplar in particular ways—for example, in shape only, material only, or shape and color but not material—and is asked which of these test objects can be called by the same name as the exemplar. Because novel objects and only minimal input—a novel noun applied to a single instance—are used, the novel noun generalization task provides insight into the child's general expectations about what novel nouns refer to.

Numerous studies have shown that when the exemplar object in

this task is made of a solid, rigid material such as wood or hardened clay, children think novel nouns refer to categories organized by similarity in shape. That is, they generalize the novel names to other objects that match the exemplar in shape. This "shape bias" has been demonstrated in many laboratories, with stimuli ranging from real, three-dimensional objects specially constructed for the experiment (Imai & Gentner, 1997; Landau et al., 1988; Soja et al., 1991) to pictures of familiar objects (Imai et al., 1994). It is important to note, however, that children do not always attend to shape in the novel noun generalization task. For example, when the exemplar has eyes or shoes, children generalize novel nouns only to test objects that match in both shape and texture (Jones & Smith, 1998; Jones, Smith, & Landau, 1991), and when a functional property is demonstrated, children generalize names to test objects that do the same thing as the exemplar even if they look very different (Kemler Nelson, Russell, Duke, & Jones, 2000). Particularly critical to the present research are additional findings that when the exemplar object is made from a nonsolid substance such as hair gel or face cream, older children are more likely to generalize the novel name to test objects made from the same material as the exemplar (Dickinson, 1988; Soja, 1992; Soja et al., 1991). This "material bias" has also been studied in numerous laboratories with varying stimuli, but it does not appear to be as robust as the shape bias (Samuelson & Smith, 1999).

Recent research has also linked the development of these biases to the development of the early productive noun vocabulary. A growing body of data suggests that attentional biases such as the shape bias change with development and that as language learning progresses, children's attentional biases become more task and context specific (see Smith, 1995, for a review). Smith and colleagues have shown that children younger than 24 months of age do not reliably generalize novel names by shape similarity in the novel noun generalization task (Gershkoff-Stowe & Smith, 2001). However, even younger children (17-month-olds) given extensive naming experience with categories of artificial objects tightly organized by similarity in shape developed a precocious shape bias. These results suggest that learning names for categories organized by shape may teach children to attend to shape when learning names for novel objects (Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002).

The research presented here takes a closer look at the relation between the nouns children know and the development of word learning biases by focusing on the development of two particular attentional biases: the shape bias for learning names of solid objects and the material bias for learning names of nonsolid substances. These two biases lend themselves to the investigation of developmental origins and mechanisms for three reasons. First, children's attention to shape and material in the context of solid and nonsolid stimuli has been studied by multiple researchers in multiple contexts. Thus, we have a relatively clear picture of the task, stimuli, and language contexts that are relevant to the development of these attentional biases. Second, solidity, shape, and material are object properties likely to be critical to the organization of even young children's categories. For example, solid things, but not nonsolid things, will make loud noises when one bangs them together, and material substance is critical to determining which white goopy things one can eat (Cool Whip) and which ones one can use to stick things together (glue). Third, a recent study of a corpus of early-learned nouns shows clear links

between the structure of the early productive noun vocabulary and the development of the shape and material biases (Samuelson & Smith, 1999).

Naming Solids and Nonsolids: The Shape and Material Biases

Children's differing generalizations for names of solid objects and nonsolid substances have been taken to indicate an early understanding of an ontological distinction between solid objects and nonsolid substances (Soja, 1992; Soja et al., 1991). A closer look at the data, however, suggests that children may really only understand how categories of solid objects are organized and named. All previous studies with children 24 months of age and older have found that children generalize names for solid objects by shape at levels well above chance. Data on children's generalizations of names for nonsolid substances are mixed. In some studies, 2- and 2½-year-olds (Soja et al., 1991) and 3-, 4-, and 5-year-olds (Dickinson, 1988) generalized names for nonsolid substances by material at better than chance levels. In other studies, children between 17 and 33 months of age did not systematically generalize names for nonsolid substances (Samuelson & Smith, 1999). In still other studies, younger but not older children demonstrated a material bias (Imai & Gentner, 1997), and sometimes older children did not generalize novel names for nonsolids by material (Subrahmanyam, Landau, & Gelman, 1999).[1] Children's generalizations of names for nonsolid stimuli are also more likely to be influenced by the syntactic frame in which the name is presented. When a mass noun syntactic frame (e.g., "some zup") was used to name solid objects, children still generalized names by shape at levels well above chance (Imai, 2000; Subrahmanyam et al., 1999). Presenting a novel name for a nonsolid substance in a count noun syntactic frame (e.g., "a zup") caused children's generalizations to fall to chance levels (Imai, 2000; Subrahmanyam et al., 1999).

Children's more robust knowledge about how categories of solid things are organized and named is also reflected in their early productive noun vocabulary. Samuelson and Smith (1999) examined the statistical structure of a corpus of 312 nouns commonly found in the productive vocabulary of children between 16 and 30 months of age. They asked adult native speakers of English to judge (a) whether nouns in the corpus referred to categories of solid objects or nonsolid substances, (b) things similar in shape, material, or color, and (c) whether each noun was a count noun or a mass noun. These dimensions were chosen because they are the ones typically tested in novel word generalization tasks with children, and they seem to control children's novel noun generalizations in specific contexts (see Jones & Smith, 1993, for a review). Samuelson and Smith found that the majority of the nouns children learn early are count nouns, names for solid things, or names for things in categories well organized by shape. These classifications are also tightly correlated: Many of the nouns

---

[1] The discrepancies in the results of these studies may be due, in part, to differences in the methods and stimuli used (stimuli presented as single piles with definite shapes vs. in a few smaller piles with irregular shapes). Regardless, this lack of consistency highlights the robustness of the shape bias, which has been consistently demonstrated in numerous laboratories with varying stimuli.

children learn early are count nouns that name solid things in shape-based categories. In contrast, few of the nouns children learn early are mass nouns, names for nonsolid substances, or names for things in categories well organized by similarity in material substance. And the correlation between these classifications is weak: Young children do not learn many mass nouns that name nonsolid substances in material-based categories.

One possible developmental relation between the pattern of shape and material biases children demonstrate in the novel noun generalization task and the asymmetries in the corpus of early-learned nouns is that the strength of the shape bias and the weakness of the material bias are the developmental product of the biased noun vocabulary. Data relevant to this issue come from a second experiment by Samuelson and Smith (1999). They compared novel noun generalizations with solid and nonsolid stimuli in children with a range of vocabulary sizes. Samuelson and Smith found that the children with the fewest nouns in their productive vocabularies did not distinguish between the solid and nonsolid stimuli—they were equally likely to generalize by shape and material with both kinds of stimuli. Only children who had at least 150 nouns in their productive vocabularies reliably generalized novel names for solid objects to other solid objects by shape. Furthermore, there was no evidence of a systematic material bias throughout the vocabulary range studied (18 to 309 nouns in the productive vocabulary). It appears, then, that the shape bias emerges only after children have already learned many names for solid objects in shape-based categories.

These results suggest an intriguing possibility. Perhaps the correspondence between solid objects, count noun syntax, and categories well organized by shape in the early noun vocabulary is so strong that it teaches children to attend to shape when learning a new name for a novel solid object. By contrast, the lack of systematicity between nonsolid substances, mass noun syntax, and material-based categories gives children at best a weak basis for generalizing names for nonsolid substances. In this way, the shape bias seen in novel noun generalization tasks may be the product of statistical regularities found among the nouns children know. The goal of the present research was to further investigate this statistical regularities hypothesis.

The Statistical Regularities Hypothesis

Considerable evidence suggests that people, including infants and young children, are ready learners of statistical regularities (Harris & Coltheart, 1986; Hasher & Zacks, 1984; Kelly & Martin, 1994; Saffran, Aslin, & Newport, 1996). The current proposal is that statistical regularities in language—among words, categories of objects, and (later) their associated syntax—provide the basis for the development of the language learning biases that aid children in mapping words to referents. The basic idea is that children may learn their very first names for things either by brute force or very weak biases, but the statistical regularities that characterize those first few names will teach children how to learn more names, creating new or stronger biases. Thus, the statistical regularities hypothesis suggests that the development of a shape bias depends on the child abstracting a general principle of how objects are named that is based on the individual nominal categories he or she already knows.

Smith and colleagues (2002) have proposed that the development of the shape bias involves four steps. First, the child needs to learn names for some number of particular objects. Second, the child needs to identify the basis of similarity shared by members of those nominal categories. That is, the child must apprehend that balls are round, cups are cup shaped, and all ice cream is the same stuff. This abstraction of within-category similarity will enable the child to generalize the nouns he or she has learned to novel contexts and instances. This abstraction is not a generalized word learning bias, however. Rather, the development of a word learning bias requires the child to learn that there is some regularity in what features of objects they should attend to when learning *novel* object names. Thus, the third step toward the development of a shape bias requires that the child abstract a second regularity that applies across the nominal categories he or she has already learned. In the case of the shape bias, this means learning that for solid things, in general, names refer to categories of instances similar in shape. Finally, in the fourth step, the child needs to apply this general abstraction to novel entities he or she has never seen before. This is the ability children demonstrate in the novel noun generalization task that experimenters take as evidence of a shape bias. Smith et al. have suggested that this ability will also lead to the more rapid acquisition of object names.

Note that the statistical regularities hypothesis is limited in that it does not say anything about how children learn the initial set of nominal categories from which they abstract the shape bias. Rather, the statistical regularities hypothesis concerns how children go from knowing some set of individual nouns to knowing how to learn new nouns, and it suggests that this knowledge comes from a vocabulary of individual names that (mostly) name things in the same way. The present article focuses on the third and fourth steps in the development of the shape bias—the abstraction of an attentional bias from the statistical regularities in the child's vocabulary and the application of that bias to completely novel things. In the present research, the statistical regularities hypothesis is tested in two ways: (a) by simulating in a connectionist network the acquisition of a shape bias in the context of learning names for solid things and (b) through experimental studies with children that specifically sought to promote the development of attentional biases by manipulating the proximal mechanism—the words children know.

The Present Simulations and Experiments

Simulation 1 investigated the abstraction of a shape bias from a vocabulary of individual nouns by asking whether the statistical regularities inherent in the early noun corpus would be sufficient to create the pattern of word learning biases seen in children. That is, would the statistical regularities be strong enough to create a shape bias in the context of solid objects and too weak to create a material bias in the context of nonsolid substances? A recurrent network was taught a noun "vocabulary" that contained the same statistical regularities found in young children's productive vocabularies: It was taught individual names for specific instances such that at a global level, the noun vocabulary it learned reflected the same statistical structure found in the child's vocabulary. After it had learned this vocabulary, the network was tested in an analog of the novel noun generalization task: It was shown novel solid and nonsolid objects to see what properties of those objects it attended

to. The results suggested that the statistical regularities *were* strong enough to support the development of a shape bias. In fact, they were so strong that the network overgeneralized the shape bias to nonsolid stimuli. This overgeneralization was a novel and unexpected result. Consequently, Simulation 2 investigated which aspect of the statistical regularities produced the overgeneralization.

Two subsequent experiments asked if it would be possible to create word learning biases *in children* by manipulating the same proximal cause manipulated in the connectionist simulations. In Experiment 1, very young children were given intensive exposure to a noun vocabulary statistically biased toward shape or material categories in an attempt to promote the development of a precocious shape or material bias. The children's productive vocabulary growth was tracked, and at the end of the 9-week longitudinal study, the children were tested in a novel noun generalization task with both solid and nonsolid stimuli. In Experiment 2, the productive vocabulary growth of a third group of children was tracked. These children also visited the laboratory 9 weeks in a row and participated in a novel noun generalization test, but they did not receive intensive exposure to a special noun vocabulary. Data from these two experiments support the statistical regularities hypothesis and suggest how the abstraction of the shape bias is related to the larger developmental chain of events by which children become smart word learners.

Simulation 1

The statistical regularities hypothesis suggests that the pattern of word learning biases seen in children is a result of statistical regularities found in their early productive vocabularies. A connectionist network provides an ideal test of the statistical regularities hypothesis because, in the simplest terms, what connectionist nets do is pick up on statistical regularities in their input and use those regularities as a basis for generalization to new instances. Indeed, a previous simulation by Gasser and Smith (reported in Smith, 1995) suggested the feasibility of modeling the development of the shape bias with connectionist models. Gasser and Smith taught a three-layer, feed-forward network names for categories of objects, most of which were organized by shape similarity. After learning the names, the network was presented with novel test objects that either matched on shape or another dimension, and the similarity of the hidden layer representations across presentations was measured. Gasser and Smith found that the network's internal representations of shape-matching test objects were more similar than its representations of test objects that matched on other dimensions, which suggested that the network had developed a shape bias. The model presented here builds on this previous simulation by incorporating the statistical regularities in the early noun corpus studied by Samuelson and Smith (1999).

The goal of this simulation was to replicate in a connectionist network the child's transition from specific knowledge of individual noun categories to more abstract knowledge of a global similarity among those noun categories. Such a model can then inform an understanding of what aspects of the statistical regularities promote the development of abstract attentional biases. This will be the case, however, only if the model accurately reflects the critical aspects of the learning situation presented to children. Thus, the following sections present an overview of the model along with the rationale for the modeling approach taken.

Overview of Model and Rationale

What the Model Is Not

Before addressing the specifics of the model, it is useful to consider what the model is *not*. This is not a model of how children learn words. The model is trained to produce a vocabulary that, in global statistical terms, matches the vocabulary produced by young children, but no effort is made to model the early trajectory of word learning seen in children (e.g., the initially slow effortful learning followed by a more rapid "vocabulary spurt"). Likewise, because the network is trained to know the same words children produce, not all the words they comprehend, this model does not address the comprehension–production distinction typically seen in children. Thus, if one assumes that the ability to comprehend a noun depends on some lexical knowledge of the nominal category, the network's knowledge at the point it is tested for a shape bias is an underestimation of the child's knowledge to the extent that the network does not represent any of the many categories children comprehend but do not produce. This issue notwithstanding, the model focuses on statistical regularities in the productive vocabulary for four reasons. First, there are currently no documented reliable measures of a child's total comprehensive vocabulary (Tomasello & Mervis, 1994). Second, there are no complete theories of the relation between comprehension and production. Third, there is no way of knowing the structure of the vocabulary representation children acquire. And fourth, the statistical regularities in the productive vocabulary have been sufficiently characterized in the literature.

This model also does not address the issue of why children learn a vocabulary with this particular structure. That is, why—out of all the words children hear—are these the particular ones they learn? This issue might be addressed by studying the input provided to children—all the things parents name for children, all the other words parents say to children, plus all the words children overhear parents say. However, even given a complete corpus of all the words ever heard by a given child, at this point there is no understanding of why a child learns some of those words and not others, no way to predict which words will be the known words, and, thus, no way to know which words would contribute to the development of word learning biases. Further, the statistical regularities hypothesis suggests that it is the *known* words that mechanistically contribute to the development of word learning biases. Thus, this model concentrates on the relation between the known vocabulary (indexed by productive measures) and the development of word learning biases.

What the Model Is

By necessity, the representation of a complex learning process in a computational model requires formalizations and/or simplifications of the learning situation, the input to the model, and the output created by the model. All models, including connectionist ones, are limited as psychological theories of development by the specifics of these simplifications. The particular model presented here entails simplifications of the vocabulary input to children and of the novel noun generalization task used to test children's word learning biases. How each of these aspects of the learning situation are represented in the present model and the particular assumptions embodied in these formalizations are discussed in turn.

The simulation presented here tested the statistical regularities hypothesis by examining whether a four-layer Hopfield network taught a "vocabulary" with the same statistical composition as a young child's productive vocabulary would demonstrate a word learning bias. The network had three "input" layers: the lexicon layer, the referent layer, and the syntax layer. These three input layers were connected bi-directionally through a hidden layer. Particular words were represented by the activation of a single unit in the word layer. Word referents were represented as a pattern of activation that spanned the units in the referent layer and corresponded to the shape, material, solidity, and "other" features of the named object. The syntax associated with a naming instance was represented by the activation of one of three units in that layer corresponding to either a count noun, a mass noun, or an "other" (nonnoun) syntactic frame.

The network was taught either a vocabulary designed to match the noun vocabulary young children learn to produce—one biased toward count nouns that name solid objects in categories well organized by similarity in shape—or a control vocabulary designed to be completely unbiased. Training consisted of repeatedly presenting each of the 22 words in a vocabulary to the network along with its corresponding referent and the associated syntax and allowing the network to incorporate the information through a Contrastive Hebbian Learning algorithm (Movellan, 1991). The network was said to "know" a word when it was able to produce the correct word in response to presentation of only an object (a novel member of the named category). Thus, the measure of whether the network knew a word was whether it could produce that word, not whether it comprehended the word (which would be demonstrated if the network produced the representation of an object in response to a presentation of the word). After the network had learned all the words in the vocabulary, it was tested on an analog of the novel noun generalization task used with children.

Assumptions and Simplifications

The division of input to the model into separate representations of words, referents, and syntax embodies two important assumptions concerning the word learning situation. Specifically, the presentation of words as activation of discrete units on the lexicon layer and as separate from syntax assumes both that the words are segmented from the speech stream and that the word and syntax are independent sources of information that are only associated to the extent supported by the statistical regularities of the vocabulary. Those assumptions seem reasonable given that young children's novel noun generalizations are influenced by the syntactic context in which a noun is presented (Hall et al., 2001; Katz et al., 1974; Waxman & Markow, 1995).

The representation of within-category structure used in this model somewhat simplifies the category learning situation presented to children. The referent of a word was presented as a distributed pattern of activation on the referent layer. Individual instances of the same word were represented by holding constant the pattern of activation across a subset of the units in the referent layer while randomly varying the activation across the remaining units. Which parts of the activation pattern were held constant depended on the organization of the referent category and whether the word referred to solid or nonsolid things. For example, for each presentation of the word *ball*, the network got the same pattern of

activation across 15 units in the referent layer—12 of which were designated as representing the round shape of balls and 3 as representing the fact that balls are solid objects—and a different random pattern of activation across the remaining 24 units. This corresponds to a child seeing many different things called *ball*, all of which share a common shape but many of which are different colors and are made from different materials. Likewise, a word that referred to a nonsolid substance in a category organized by similarity in material was presented to the network by holding the pattern of activation across a different set of 12 material units and 3 solidity units constant and varying the activation across the remaining units. In this way, the network received both different types of nouns (those defined by similarity in shape and those defined by similarity in material) as well as different instances of those nouns.

Note that this means that for every noun the network learned, category instances were exactly the same on some perceptual dimension. This is, of course, not always the situation presented to children, because many of the nouns children learn refer to instances that are not identical on any perceptual dimension (e.g., not all chairs are exactly the same shape and not all lotion is exactly the same stuff). Nevertheless, this is a reasonable simplification because the model was not designed to be a model of how children learn the initial vocabulary from which they generalize a shape bias. That is, the question at hand is not whether children can learn to label all things that are exactly the same shape by the same name. Rather, the goal of this simulation was to understand the higher order abstraction made by children—that shape is a defining characteristic of categories of solid things in general. The formation of this higher order association between solidity and category organization requires that the network learn that when the solid object unit is active, the pattern of activation across a different set of units in the same layer is what determines which word unit should be activated. Thus, this abstraction is based on the similarity across categories of solid things, not similarity within a given category.

Because the network has no output corresponding to a choice of one object over another (the response typically given by children in novel noun generalization tasks), a reasonably equivalent task was devised. This task was based on the following conceptualization of the novel noun generalization task: When a young child chooses, for example, a shape-match test object over a material-match test object, she does so because she perceives the shape-match test object to be more similar to the exemplar than is the material-match object. In this way, children's choices in the novel noun generalization task reflect the perceived similarity between the exemplar and test objects in the context of a naming event. Thus, the network version of the novel noun generalization test examined the similarity between the network's internal representations of novel test objects that matched in shape and novel test objects that matched in material.

More specifically, the network was presented with several instances of novel objects that matched in shape and several instances of novel objects that matched in material. The similarity between the patterns of activation produced on the hidden layer by presentations of each of these novel objects was computed. These similarity scores were then used to compute the probability of a shape choice versus a material choice in the novel noun generalization task. Thus, like the test used with children, the novel noun

generalization test given to the network used novel objects and minimal input to investigate the network's general expectations about the structure of noun categories. If the statistical regularities inherent in the early noun vocabulary are sufficient to create the word learning biases children demonstrate in novel noun generalization tasks, then networks trained with a vocabulary that matches the natural statistics found in young children's vocabularies should demonstrate a shape bias—but not a material bias—in this novel noun generalization task. In contrast, networks trained on a balanced vocabulary should show chance performance with both solid and nonsolid stimuli.

Method

Architecture

The network is pictured in Figure 1. In this figure, circles represent individual units, boxes represent layers of units, and arrows between boxes indicate full innerconnectivity between units in connected layers (arrows connecting a layer to itself indicate that all units in that layer are connected to all other units in that layer). All units in the network could assume activations between −.95 and .95 (see the Appendix for activation, input, and weight change functions). Naming events were presented to the network via three "input" layers: a referent layer of 39 units, a syntax layer of 3 units, and a lexicon layer of 22 units. These three input layers were connected bi-directionally through a hidden layer that contained 20 units. These bi-directional connections between the input layers and the hidden layer allowed information in the network to flow in more than one direction (i.e., from a word to a referent or from a referent to a word).

Training Vocabularies

Two training "vocabularies" were presented to the network—the natural statistics vocabulary and the balanced statistics vocabulary. Both vocabularies contained 22 words. Of these, 14 words (65%) were designated as nouns and 8 (35%) were designated as "other" words. The percentages used to determine the size of the noun segment and other segments of the

Figure 1. The network used in the simulations. Circles represent individual units, boxes represent layers of units, and arrows between boxes indicate full innerconnectivity between units in connected layers. Arrows connecting a layer to itself indicate that all units in that layer were connected to all other units in that layer.

vocabularies were based on the mean proportions of nouns and other words found in the productive vocabularies of 4 children studied by Nelson (1973).

The noun subsection of the natural statistics vocabulary was designed to match, as closely as possible, the overall statistical regularities of the early noun corpus as reported by Samuelson and Smith (1999). The numbers of names for solid objects, nonsolid substances, and things ambiguous in solidity in the natural statistics vocabulary were determined by the percentages of each kind of noun found in the early noun corpus. Likewise, the numbers of names for categories organized by similarity in shape, by similarity in material, or ambiguously organized and the numbers of count nouns, mass nouns, and nouns ambiguous in syntax were based on the percentages of each classification found in the early noun corpus. Table 1 presents these percentages from Samuelson and Smith and the resulting proportions of each kind of noun used in the natural statistics vocabulary.

Individual nouns in the vocabulary were created by combining solidity, category organization, and syntax such that the percentages of each solidity–syntax, solidity–category organization, syntax–solidity, and syntax–category organization combination matched, as closely as possible, the percentages reported by Samuelson and Smith (1999; see Table 1). Both the percentage of nouns that were count nouns naming solid objects in categories well organized by similarity in shape (36%) and the percentage of mass nouns naming nonsolid substances in categories well organized by similarity in material substance (7%) were very close to the percentages found by Samuelson and Smith (35% and 2%, respectively).

In contrast, the noun subsection of the balanced statistics vocabulary was composed of approximately equal numbers of each solidity classification, category organization, and syntactic category (see Table 2), and each noun presented a different combination of solidity, category organization, and syntax. Thus, there were approximately equal numbers of each solidity–syntax, solidity–category organization, syntax–solidity, and syntax–category organization combination.

Individual words were presented to the network as patterns of activation on the three input layers (referent, syntax, and lexicon). The solidity of a given word was represented by a pattern of activation across a subset of three units in the referent layer. This is illustrated in Figure 1 by the Solidity label next to the row of three units in the referent layer. A solid object was represented by full positive activation (.95) on the first of these three units and full negative activation (−.95) on the other two units. Likewise, nonsolidity was represented by full positive activation on the second unit and full negative activation on the other two units. Ambiguous solidity was represented by full positive activation of the third unit and full negative activation of the other two units. Because it is not clear what solidity input nonnouns such as verbs, adjectives, and prepositions provide, all three solidity units were given full negative activation for all other words.[2] The category organization of a given word was represented by a unique random pattern of full positive (.95) or full negative (−.95) activation across 12 of the units in the referent layer. This subset of the activation pattern on the referent layer was the same for each individual presentation of a given word. This is illustrated in Figure 1 by the Shape, Material, and Other labels next to rows of different sets of 12 units in the referent layer. Ambiguously organized noun categories were represented by holding constant unique random patterns on both the shape and material subsets of the referent layer across individual presentations of a word. The syntax associated with a given word was represented by full positive activation on the corresponding unit and full negative activation on the other two units. Ambiguous noun syntax was represented by full positive
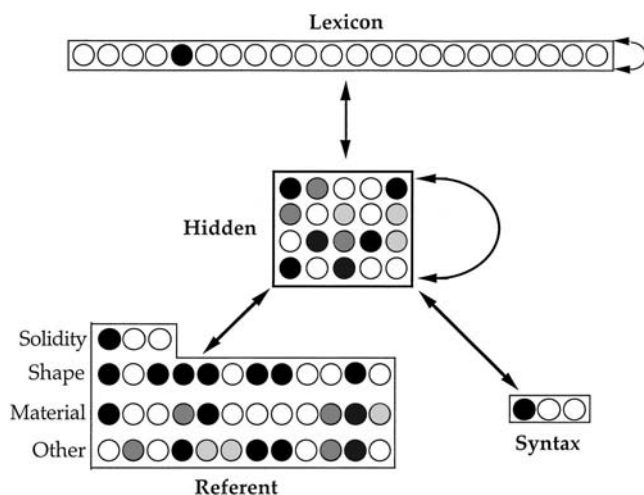
---

[2] As a check that this representation of solidity for other words did not bias the network's performance, another simulation was run with the solidity of other words represented as random activation across the three solidity units. The performance of this simulation was identical to the performance of the simulations reported in the text.

Table 1

Percentages of the 312 Nouns Studied by Samuelson and Smith (1999) and the 14 Nouns in the Natural Statistics Vocabulary From Each Solidity Classification, Syntactic Category, and Category Organization and the Proportion of These Nouns That Were Also Names for Solid Objects, Nonsolid Substances, or Things Ambiguous in Solidity and That Were Count Nouns, Mass Nouns, and Nouns of Ambiguous Syntax

Vocabulary, solidity classification, and syntactic category	Solidity classification			Syntactic category			Category organization		
	Solid	Nonsolid	Ambiguous	Count noun	Mass noun	Ambiguous	Shape based	Material based	Ambiguous
Percentage in Samuelson & Smith (1999) corpus (N = 312)	63	4	32	74	10	16	48	16	39
Percentage in natural statistics vocabulary (N = 14)	64	7	29	64	14	21	50	29	21
Solidity classification									
Solid				.79 (.85)	.00 (.02)	.22 (.13)	.79 (.61)	.00 (.11)	.22 (.31)
Nonsolid				.00 (.00)	1.00 (.79)	.00 (.21)	.00 (.00)	1.00 (.51)	.00 (.39)
Ambiguous				.50 (.65)	.25 (.16)	.25 (.19)	.00 (.30)	.75 (.15)	.25 (.50)
Syntactic category									
Count noun	.79 (.71)	.00 (.00)	.22 (.29)				.56 (.59)	.11 (.06)	.33 (.38)
Mass noun	.00 (.13)	.50 (.35)	.50 (.52)				.00 (.00)	1.00 (.58)	.00 (.42)
Ambiguous	.67 (.56)	.00 (.06)	.33 (.38)				.67 (.28)	.00 (.43)	.33 (.39)

Note. Numbers in parentheses are the proportions of the same correlations found in Samuelson and Smith's (1999) study of the early noun corpus.

activation on both the count noun and mass noun units and full negative activation on the other syntax unit. This representation of ambiguity reflects the fact that most of the nouns judged as ambiguous by adults in Samuelson and Smith's (1999) study were those that could be used with both a count noun or mass noun syntactic frame (e.g., "I'd like four Cokes, please" and "Can I please have some Coke?").

Training

Networks were trained with Contrastive Hebbian Learning (Movellan, 1991). In this learning algorithm, the connection weight between two units is updated by a proportion of the difference between activations in clamped and free-running phases. In the clamped phase, the patterns of activation corresponding to a word, a referent, and the associated syntax are held constant on the lexicon, referent, and syntax layers of the network, and the activations of units in the hidden layer are allowed to "settle" (i.e., they are adjusted according to Equations A1 and A2 in the Appendix). This phase

corresponds to the network being told that a word (the pattern on the lexicon layer), a referent (the pattern on the referent layer), and a syntactic context are associated. The activations from this "instructional" phase are recorded to serve as the correct pattern for the next, free-running phase. In the free-running phase, the pattern of activation across the lexicon layer is unclamped. The network is allowed to settle again, and the activations are again recorded. This corresponds to the network being presented the referent without the word or syntax. The difference in activations from the clamped and free-running phases is computed, and this "error" is then multiplied by a constant learning rate and used to adjust the connection weights (see Equation A3 in the Appendix).

This complete sequence was repeated for each of the 22 words in the training vocabulary. Each presentation of the complete set of 22 words defined an epoch of training. Note that the use of an error-driven learning algorithm is sometimes taken to imply the necessity of detecting a difference between the current representation of an input and the "correct" representation in the course of word learning. However, given that the goal

Table 2

Numbers of the 14 Nouns in the Balanced Statistics Vocabulary From Each Solidity Classification, Syntactic Category, and Category Organization and the Proportion of These Nouns That Were Also Names for Solid Objects, Nonsolid Substances, or Things Ambiguous in Solidity and That Were Count Nouns, Mass Nouns, and Nouns of Ambiguous Syntax

Number in vocabulary, solidity classification, and syntactic category	Solidity classification			Syntactic category			Category organization		
	Solid	Nonsolid	Ambiguous	Count noun	Mass noun	Ambiguous	Shape based	Material based	Ambiguous
Number in vocabulary	5	5	4	5	5	4	5	5	4
Solidity classification									
Solid				.40	.40	.20	.40	.40	.20
Nonsolid				.40	.40	.20	.40	.40	.20
Ambiguous				.25	.25	.50	.25	.25	.50
Syntactic category									
Count noun	.40	.40	.20				.40	.20	.40
Mass noun	.40	.40	.20				.40	.20	.40
Ambiguous	.25	.25	.50				.25	.75	.00

in the present simulations was only to analyze the statistical regularities in the vocabulary known by the network after training, the decision to use this particular word learning algorithm was based primarily on the network architecture (e.g., the use of a standard backpropagation learning algorithm would require specification of the direction of activation flow and designation of input and output layers).

This learning algorithm was used to teach 10 networks the natural statistics vocabulary and 10 networks the balanced statistics vocabulary. Each network had a different randomly generated set of initial connection weights. To determine when the network had learned the vocabulary, it was tested every four epochs by presenting a pattern of activation across the referent layer corresponding to each word and measuring activation produced on the lexicon layer (no syntax information was presented or tested). The network was said to "know" a word when it correctly generated maximum activation at the unit in the lexicon layer corresponding to the presented referent. Networks were trained on the vocabularies until they reliably produced all the nouns (32 epochs). There were no differences in acquisition for networks taught the natural statistics or balanced statistics vocabularies.

Testing

After a network had learned the training vocabulary, it was tested in a novel noun generalization task. Like the novel noun generalization task used previously with children, the network was tested with both a solid stimulus set and a nonsolid stimulus set and was not given any information about syntax. The solid set consisted of 10 solid + shape categories and 10 solid + material categories. Each solid + shape category was defined by four instances. For all four instances, the pattern of activation across the referent layer matched on 13 of the 39 units (the units corresponding to the shape and solidity subsets used to define shape-based categories and solid objects in training) and had a random (and thus different) pattern of activation across the rest of the units in the referent layer. Likewise, each solid + material category was defined by four instances that had the same pattern of activation on the solidity and material subsets of the referent layer but a random (and different) pattern of activation across the rest of the units in the referent dimension.

To find the similarity among instances within a category, the network was presented with each of the four category instances, and the vector of activation across the hidden layer produced by each instance was recorded. Next, the euclidean distance between each possible pair of these four vectors was calculated (i.e., v1&v2, v1&v3, v1&v4, v2&v3, v2&v4, v3&v4). The average of these six distances was used as the euclidean distance measure for each category. Smaller euclidean distances indicated that patterns of hidden layer activation were relatively similar. These category distances were then averaged across the 10 solid + shape categories, and across the 10 solid + material categories, creating the two distance measures for the solid test set (one solid + shape and one solid + material).

Next, these distance measures (*D*), were converted to normalized similarities with the following equation:

$$Normalized\ Sim = 1 - \frac{D}{Max * S} \qquad (1)$$

where *Max* was the maximum distance for the hidden layer size and *S* was a constant scale factor (.22). The constant scale factor was used when computing the similarity scores because the euclidean distances were relatively small. It is important to note that the same scale factor was used for all networks. The two normalized similarity scores for a test set were submitted to Luce's choice rule to produce a probability of a shape choice and a probability of a material choice for the solid test set. The same procedure was repeated for the 10 nonsolid + shape and 10 nonsolid + material categories to produce the probability of a shape choice and the probability of a material choice for the nonsolid test set.

Results

The results of the novel noun generalization test at Epoch 32 are pictured in the left panel of Figure 2. The figure shows the mean
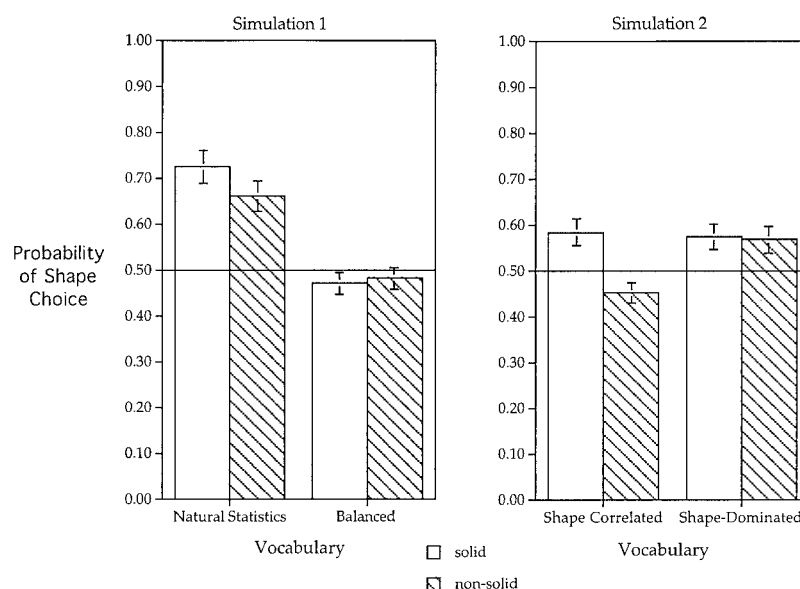

Figure 2. Results of Simulations 1 and 2. The mean probabilities of a shape choice for solid and nonsolid stimulus sets at Epoch 32 for networks trained on the natural statistics and balanced statistics vocabularies (left panel) and on the shape-correlated and shape-dominated vocabularies (right panel) are shown. Vertical lines depict standard errors of the means.

probability of a shape choice with the solid and nonsolid test sets for networks trained with the natural statistics and the balanced statistics vocabularies. A training vocabulary (natural vs. balanced) by test set (solid vs. nonsolid) analysis of variance (ANOVA) revealed a significant main effect of training vocabulary, $F(1, 18) = 134.25$, $p < .001$, and a significant Vocabulary $\times$ Test Set interaction, $F(1, 18) = 8.50$, $p < .01$. The main effect of test set was marginally significant, $F(1, 18) = 4.13$, $p < .07$. As can be seen in Figure 2, only networks taught a vocabulary that closely matched the statistical regularities found in young children's vocabularies demonstrated a shape bias. The graph also suggests that the probability of a shape choice was greater for the solid test set than for the nonsolid test set for networks trained on the natural statistics vocabulary. This was confirmed by a dependent samples t test, $t(9) = 2.78$, $p < .05$. Nonetheless, it is clear in Figure 2 that networks trained on the natural statistics vocabulary overgeneralized the shape bias to nonsolid stimuli. This was confirmed by t tests comparing the number of shape choices with the solid and nonsolid test sets with chance: $t(9) = 8.57$, $p < .001$ and $t(9) = 14.45$, $p < .001$, respectively. Thus, networks taught the natural statistics vocabulary generalized names for both solid and nonsolid stimuli on the basis of shape similarity.

One possible cause for this overgeneralization is the lack of syntax cues in the novel noun generalization test. To test this possibility, the networks were tested in a novel noun generalization task that included syntax. The results are presented in Table 3. These data were analyzed in a Training Vocabulary (natural vs. balanced) $\times$ Test Set (solid vs. nonsolid) $\times$ Syntax (included vs. not included) ANOVA. Given that the primary goal of this analysis was to investigate potential syntax effects, only significant syntax effects are reported. The ANOVA revealed a Vocabulary $\times$ Syntax interaction, $F(1, 36) = 6.26$, $p < .05$. As can be seen in Table 3, the addition of syntax to the novel noun generalization task reduced the proportion of shape responses made by networks trained with the natural statistics vocabulary. However, the addition of syntax did not eliminate the overgeneralization of shape responding to the nonsolid test set. T tests confirmed that the numbers of shape choices with the solid and nonsolid test sets were both significantly different from chance: $t(9) = 7.94$, $p < .001$ and $t(9) = 6.27$, $p < .001$, respectively. Therefore, networks taught a vocabulary similar to that of the young child generalized names for both novel solid and nonsolid stimuli by shape similarity regardless of whether the task included syntax information.

Table 3
Proportions of Shape Responses for Networks Trained With the Natural Statistics or Balanced Statistics Vocabularies in the Novel Noun Generalization Task Without Syntax and With Syntax Added

	Natural statistics vocabulary		Balanced statistics vocabulary	
Syntax	Solid	Nonsolid	Solid	Nonsolid
Without syntax	.72	.66	.47	.48
With syntax	.70	.57	.49	.50

Discussion

The results of Simulation 1 suggest that the statistical regularities inherent in the noun subsection of young children's productive vocabularies are strong enough to support the development of a shape bias but not strong enough to support the development of a material bias. Further, the statistical regularities are so strong that networks taught the natural statistics vocabulary overgeneralized the shape bias to the nonsolid stimulus set. This unexpected overgeneralization does not appear to be based on the lack of syntax in the novel noun generalization task. Simulation 2 investigated two other aspects of the statistical regularities that may have contributed to the overgeneralization made by the network.

Simulation 2

The results of Simulation 1 suggest not only that statistical regularities in young children's vocabularies are sufficient to support the development of a shape bias but that they can create an overgeneralized shape bias. There are two aspects of these statistical regularities that could be responsible for the network's strong bias to generalize novel names for novel stimuli by shape. In their study of the early noun corpus, Samuelson and Smith (1999) found that young children's noun vocabularies were dominated by names of solid objects, categories organized by similarity in shape, and count nouns and that these classifications were highly correlated—many of the early nouns were count nouns that named solid objects in categories well organized by similarity in shape. In contrast, few of the nouns in the early productive vocabulary were names for nonsolid substances, categories organized by similarity in material substance, or mass nouns. Further, these classifications were not highly correlated—there were very few mass nouns that named nonsolid substances in material-based categories. Thus, there is more support in the early noun corpus for the development of a shape bias than for the development of a material bias, both in terms of the correlation between solidity and category organization by shape and the sheer dominance of names for shape-based categories. Either of these aspects of the vocabulary could have led the network to overgeneralize the shape bias. The current simulation asked whether it was the high frequency of names for shape-based categories or the tight correlation between solidity and category organization that led the network to develop a shape bias that was overgeneralized to nonsolid stimuli.

As with young children's productive noun vocabularies, the two vocabularies used in Simulation 1 differed in both the numbers of nouns in each kind of solidity classification, category organization, and syntactic category and in the strength of the correlations between solidity, category organization, and syntax. To test the relative contributions of frequency and correlations to the development and overgeneralization of a shape bias, two new vocabularies were constructed (see Table 4). The shape-correlated vocabulary had equal numbers of nouns in each category organization, solidity classification, and syntactic category (like the balanced statistics vocabulary) but tight correlations among category organization, solidity, and syntax (like the natural statistics vocabulary). In contrast, the shape-dominated vocabulary was dominated by names for shape-based categories, names for solid objects, and count nouns (like the natural statistics vocabulary) but had only weak correlations among category organization, solidity, and syntax (like the balanced statistics vocabulary).

Table 4

Relation Between Dominance of Names for Solid Objects, Shape-Based Categories, and Count Nouns and Correlation Between These Noun Classifications in the Vocabularies Used in Simulation 2

Correlation between solidity, syntax, and category organization	Frequency of count nouns, solid objects, and shape-based categories	
	High	Low
High	Natural statistics vocabulary	Shape-correlated vocabulary
Low	Shape-dominated vocabulary	Balanced statistics vocabulary

Networks were taught each of the new vocabularies and were tested in the same novel noun generalization task. If the overgeneralization found in the previous simulation was due to the high frequency of names for shape-based categories, then only the shape-dominated vocabulary should display an overgeneralized shape bias. If, instead, the previous overgeneralization was due to the tight correlations between categories organized by similarity in shape, solid objects, and count nouns, then only the shape-correlated vocabulary should display an overgeneralized shape bias.

Method

Architecture

The same network architecture used in Simulation 1 was used in this simulation.

Training Vocabularies

Two training vocabularies were presented to the network, the shape-correlated vocabulary and the shape-dominated vocabulary. As in Simulation 1, both vocabularies contained 22 words—14 nouns and 8 other words.

The numbers of nouns in the noun subsection of the shape-dominated vocabulary that were from each solidity classification, category organization, and syntactic category were based on the percentages reported in Samuelson and Smith (1999). However, in contrast to the natural statistics vocabulary, solidity, category organization, and syntax were combined in such a way as to avoid a tight correlation between count nouns, names for solid objects, and shape-based categories (see Table 5).

In contrast, the noun subsection of the shape-correlated vocabulary contained equal numbers of nouns from each solidity classification, category organization, and syntactic category. These were combined in such a way as to match, as closely as possible, the percentages of each solidity–syntax, solidity–category organization, syntax–solidity, and syntax–category organization combination reported by Samuelson and Smith (1999). The percentages of each correlation in the vocabulary are given in Table 6 along with the percentages found by Samuelson and Smith. Thus, even though count nouns, names for solid objects, and shape-based categories did not dominate this vocabulary, there was a tight correlation between count nouns, names for solid objects, and shape-based categories. Note that because this vocabulary contained equal numbers of count and mass nouns, names for solid and nonsolid objects, and names for categories organized by shape and by material, there was also a tight correlation between mass nouns, names for nonsolid substances, and material-based categories.

Training and Testing

Training and testing proceeded just as in Simulation 1.

Results

For networks taught the shape-correlated and shape-dominated vocabularies, the mean probabilities of a shape choice on the solid and nonsolid test sets in the novel noun generalization test at Epoch 32 are pictured in the right-hand panel of Figure 2. The data were analyzed with a Vocabulary (shape-dominated vs. shape-correlated) × Test Set (solid vs. nonsolid) ANOVA. This analysis revealed a significant main effect of test set, $F(1, 18) = 18.40, p < .001$, and a significant interaction between test set and vocabulary, $F(1, 18) = 14.98, p < .01$. As can be seen in Figure 2, networks taught the shape-correlated vocabulary generalized novel names

Table 5

Percentages of the 312 Nouns Studied by Samuelson and Smith (1999) and the 14 Nouns in the Shape-Dominated Vocabulary From Each Solidity Classification, Syntactic Category, and Category Organization and the Proportion of These Nouns That Were Also Names for Solid Objects, Nonsolid Substances, or Things Ambiguous in Solidity and That Were Count Nouns, Mass Nouns, and Nouns of Ambiguous Syntax

Vocabulary, solidity classification, and syntactic category	Solidity classification			Syntactic category			Category organization		
	Solid	Nonsolid	Ambiguous	Count noun	Mass noun	Ambiguous	Shape based	Material based	Ambiguous
Percentage in Samuelson & Smith (1999) corpus ($N = 312$)	63	4	32	74	10	16	48	16	39
Percentage in shape-dominated vocabulary ($N = 14$)	64	7	29	50	29	21	64	14	21
Solidity classification									
Solid				.44	.22	.33	.33	.33	.33
Nonsolid				1.00	.00	.00	1.00	.00	.00
Ambiguous				1.00	.00	.00	.75	.25	.00
Syntactic category									
Count noun	.44	.11	.44				.44	.33	.22
Mass noun	1.00	.00	.00				.50	.50	.00
Ambiguous	1.00	.00	.00				.67	.33	.00

Table 6

Numbers of the 14 Nouns in the Shape-Correlated Vocabulary From Each Solidity Classification, Syntactic Category, and Category Organization and the Proportion of These Nouns That Were Also Names for Solid Objects, Nonsolid Substances, or Things Ambiguous in Solidity and That Were Count Nouns, Mass Nouns, and Nouns of Ambiguous Syntax

Number in vocabulary, solidity classification, and syntactic category	Solidity classification			Syntactic category			Category organization		
	Solid	Nonsolid	Ambiguous	Count noun	Mass noun	Ambiguous	Shape based	Material based	Ambiguous
Number in vocabulary	5	5	4	5	5	4	5	5	4
Solidity classification									
Solid				.80 (.85)	.00 (.02)	.20 (.13)	.80 (.61)	.00 (.11)	.20 (.31)
Nonsolid				.00 (.00)	.80 (.79)	.20 (.21)	.00 (.00)	.80 (.51)	.20 (.39)
Ambiguous				.25 (.65)	.25 (.16)	.50 (.19)	.25 (.30)	.25 (.15)	.50 (.50)
Syntactic category									
Count noun	.80 (.71)	.00 (.00)	.20 (.29)				.60 (.59)	.00 (.06)	.40 (.38)
Mass noun	.00 (.13)	.80 (.35)	.20 (.52)				.00 (.00)	.60 (.58)	.40 (.42)
Ambiguous	.25 (.56)	.25 (.06)	.50 (.38)				.25 (.28)	.25 (.43)	.50 (.39)

Note. Numbers in parentheses are the proportions for the same correlations found in Samuelson and Smith's (1999) study of the early noun corpus.

for solid objects by shape. Moreover, because of the equally tight correlation between mass noun syntax, nonsolidity, and category organization by material in this vocabulary, these networks were also more likely to make a material choice when the test objects were nonsolid. A dependent samples t test comparing the performance of networks trained with the shape-correlated vocabulary on the solid and nonsolid test trials found a significant difference across test sets, $t(9) = 2.73$, $p < .05$. However, t tests comparing their performance with chance performance (.5) indicated that the bias to attend to shape with solid test objects was stronger than the bias to attend to material for nonsolid test objects. Specifically, the number of shape choices made by the networks taught the shape-correlated vocabulary with solid test objects differed significantly from chance, $t(9) = 4.87$, $p < .001$, whereas the performance of these networks on the nonsolid trials was only marginally different from chance, $t(9) = -2.01$, $p < .10$.

In contrast to the networks taught the shape-correlated vocabulary, networks taught the shape-dominated vocabulary were more likely to make a shape choice on both the solid and nonsolid test trials. This finding was confirmed by t tests comparing chance performance with the number of shape choices the networks taught the shape-dominated vocabulary made with the solid and nonsolid test sets: $t(9) = 3.41$, $p < .01$ and $t(9) = 3.11$, $p < .05$, respectively. Thus, networks taught a vocabulary dominated by shape-based categories overgeneralized the shape bias to nonsolid stimuli.

Discussion

The results of Simulation 2 suggest that a simple statistical learner taught a vocabulary containing a high frequency of count nouns, names for solid objects, and shape-based categories will develop an overgeneralized shape bias in the novel noun generalization task. The simulation also suggests that a simple statistical learner taught a vocabulary containing strong correlations among solidity, category organization, and syntax will learn to differentiate solid objects and nonsolid substances in the novel noun generalization task.

Taken together, the results of the two simulations support the statistical regularities hypothesis. Networks taught a vocabulary containing the same statistical regularities as the young child's productive vocabulary developed a shape bias. Thus, these networks, like the young child, use their knowledge about how specific noun categories are organized to form a generalized attentional bias that can be applied to novel instances. Further, these simulations suggest that different aspects of the statistical regularities may lead to different applications of the shape bias to novel stimuli. When the early productive vocabulary is dominated by names for categories organized by shape, names for solid objects, and count nouns, the shape bias may be overgeneralized to nonsolid stimuli.[3] In contrast, when the early vocabulary contains tight correlations between count noun syntax, solid objects, and shape-based categories, and equally tight correlations between mass noun syntax, nonsolidity, and material-based categories, both a shape bias for solid stimuli and a material bias for nonsolid stimuli may develop.

Thus, these simulations suggest a strong test of the hypothesis: If the statistical regularities inherent in the noun categories children learn early are strong enough to teach connectionist networks a shape bias, then they may also be strong enough to teach very young children to attend to shape when learning a novel name for a novel solid object. Further, the simulations suggest that the dominance of names for shape-based categories in young children's productive vocabularies should lead to an overgeneralization of the shape bias to nonsolid stimuli. In contrast, it might be possible to teach young children to attend to material substance when generalizing names for novel nonsolid stimuli by teaching them names that highlight the correlation between nonsolidity,

---

[3] Although the dominance of names for shape-based categories is most likely the factor leading the networks to overgeneralize the shape bias to nonsolid stimuli, the fact that the shape-dominated vocabulary used in this simulation contained high numbers of names for solid objects and count nouns means that I cannot rule out the possibility that these other factors played a role in the overgeneralization. Resolution of this issue requires further simulations, which are currently under way.

mass noun syntax, and categories organized by material. These possibilities were tested in the following two experiments.

Experiment 1

The connectionist simulations suggest that the statistical regularities inherent in the child's early noun vocabulary are sufficient to create a word learning bias in a simple, unbiased learner of statistical regularities. This experiment asked whether the same statistical regularities would be sufficient to create word learning biases in children. Recent data from Smith and colleagues (2002) suggest that it should be possible. These researchers taught young children names for four artificial categories organized by shape. Following training, these children demonstrated a precocious shape bias. The present experiments extended these results in four ways: (a) by incorporating the statistical regularities found in the early noun corpus, (b) by teaching children names for real objects, (c) by attempting to teach a material bias, and (d) by testing the prediction from the simulation studies that children will overgeneralize the shape bias to nonsolid stimuli.

Two groups of children between 15 and 21 months of age participated in a 9-week longitudinal study. Both groups of children were given intensive naming experience with 12 real nouns that, according to the LEX database (Dale & Fenson, 1993), children do not usually learn until after 26 months of age. For children in the natural statistics condition, the nouns were selected to represent, on a global statistical level, the kinds of nouns children typically learn early. Specifically, the statistical composition of the set of trained nouns matched that of the corpus of early-learned nouns studied by Samuelson and Smith (1999).

For children in the material-biased condition, all 12 nouns were names for nonsolid substances in categories well organized by similarity in material (e.g., *lotion, glitter,* and *Jell-O*). In Simulation 2, networks taught a vocabulary containing tight correlations both between shape-based categories, solid objects, and count noun syntax and between material-based categories, nonsolid substances, and mass noun syntax differentiated between solid and nonsolid test sets in the novel noun generalization task. This suggests that tight correlations between names for material-based categories, nonsolid substances, and mass noun syntax will be necessary to create a material bias. Given that young children's productive vocabularies naturally present tight correlations between names for solid objects, shape-biased categories, and count noun syntax (Samuelson & Smith, 1999), the set of nouns taught to children in the material-biased condition was heavily biased toward names that are hypothesized to promote a material bias. One issue, of course, is whether 12 nouns are statistically enough to compete with the shape bias (in terms of statistical regularities) naturally found in young children's productive vocabularies. This number of nouns was selected to keep the length of the study reasonable. To maximize the potential effect of material noun training, only children with fewer than 150 nouns in their productive vocabularies (the number of nouns found to be critical to the demonstration of a shape bias in Samuelson and Smith's study) were included in the experiment.

An important issue in this experiment was whether the children would succeed in learning the nouns presented in each condition so that the manipulated variable would be the same as that used in the simulations (words known). Three specific procedures were included in an effort to ensure that children learned the nouns presented in the experiment: (a) Each word was presented at least 20 times in a training session, (b) children were prompted to produce each of the test words (e.g., "Look, here's some pudding! Can you say pudding?"), and (c) at the beginning of each session, parents were informally asked whether their children had said any of the experimental words or been exposed to toys or stimuli similar to those used in the experiment. Informal observations during the experimental sessions and parent reports of activity at home suggested that most of the children did learn most of the nouns. In addition, several versions of comprehension and production tests for the experimental words were pilot tested. In these tests, children were asked to pick a named item from two choice items that had both been presented in previous experimental sessions (e.g., "Get the lotion" with piles of lotion and pudding as test objects). However, these tests did not provide reliable data because children often picked the wrong object as the referent for a word because of their preference for one of the test foils (e.g., they wanted to eat the pudding) even when it was clear they knew the right answer. For this reason, and because it was clear we were already taxing the limits of what these young children would tolerate during an experimental session, these tests were not included in the final version of the experiment.

Both groups of children were given a novel noun generalization test with both solid and nonsolid stimuli at the end of noun training. Because children this young often have difficulty with the novel noun generalization test, the testing procedure was practiced with familiar stimuli at each experimental visit and with novel stimuli during two laboratory visits prior to the final experimental test. The productive vocabulary growth of all the children was measured via parental report on the toddler form of the MacArthur Communicative Development Inventory (MCDI; Fenson et al., 1994) at the beginning and end of the noun training. In addition, a follow-up report of productive vocabulary was obtained 1 month after the novel noun generalization test.

If the shape bias is the product of learning many names for solid objects in shape-based categories, then children in the natural statistics condition should demonstrate a shape bias with the solid stimuli in the novel noun generalization test. Further, the high frequency of these nouns in the early productive vocabulary should cause these children to overgeneralize the shape bias to nonsolid stimuli. In addition, if the perfect correlation presented by the 12 mass nouns for nonsolids in material-based categories of the material-biased condition is enough to compete with the dominance of count nouns naming solid objects in shape-based categories inherent in the early noun vocabulary, then children in the material-biased condition should demonstrate a material bias with the nonsolid stimuli in the novel noun generalization task.

Method

Participants

Twenty children, 11 boys and 9 girls between the ages of 15 and 21 months, participated (mean age = 18 months 28 days; range = 15 months 20 days to 21 months 3 days). Children were recruited from the child participants file at Indiana University. All children were learning English as their first and only language. Parents of potential participants were contacted by mail and recruited in a follow-up phone call. Potential participants were screened by asking parents approximately how many

words their children could produce and whether their children used any of the nouns to be taught in the experiment. At the first experimental session, children were matched on productive vocabulary scores from a subsection of the MCDI. Members of each matched pair of children were randomly assigned to either the natural statistics or the material-biased condition such that the mean age, $t(18) = 0.53$, *ns*, and vocabulary, $t(18) = 0.01$, *ns*, across conditions did not differ. Four additional children began but did not complete the experiment: 1 because his family moved out of town, 1 because he fussed and refused to participate in the training sessions, and 2 (1 boy and 1 girl) because their parents decided that a longitudinal study was too time consuming. Data from 1 other child were not included in the analyses because the child with whom she was matched dropped out of the experiment. All children received a small prize at each experimental visit, and the family received copies of experimental videotapes and T-shirts at the completion of the study.

Materials

Conditions differed only in the 12 nouns taught to the children over the course of the longitudinal study. All 24 nouns were nouns not usually learned until after 26 months of age. All noun categories were selected to represent a given solidity or category organization on the basis of the adult judgments reported in Samuelson and Smith (1999). Thus, *bucket* was selected as a shape + solid noun because 85% of adults in Samuelson and Smith's study had agreed that buckets were similar in shape and solid. Likewise, *chalk* was selected as an example of a noun with ambiguous syntax because adults in Samuelson and Smith's study did not agree on whether it was a count noun or a mass noun. The noun category training sets for each condition consisted of three examples of each of the 12 nouns. In both conditions, five of the nouns taught to the children named foods.

Natural statistics nouns. The nouns taught to children in the natural statistics condition were chosen to be representative of the kinds of nouns young children typically learn early. Specifically, the nouns were picked such that the numbers of count nouns, mass nouns, and nouns with ambiguous syntax, of names for solid objects and nonsolid substances, and of names for shape- and material-based categories were statistically proportional to the numbers found in the corpus of early-learned nouns studied by Samuelson and Smith (1999). Thus, many of the nouns (5 out of 12) were count nouns such as *bucket, ladder,* and *pear* that name categories of solid objects well organized by similarity in shape. However, other kinds of nouns were included, for example, count nouns such as *pretzel* that name

solid objects in categories organized by similarity in material, and nouns ambiguous in syntax such as *chalk* that name solid objects that are similar in both shape and material.

The three examples of each noun category used in the experiment were chosen to highlight both the similarities and the differences among instances of the named category. However, because of natural variations among the three instances of each category, no two exemplars matched exactly on any one dimension. For example, the three pears chosen were all solid and highly similar in shape but different in color (green, red, or brown) and size (two were similar in size and one was much smaller). Likewise, for artifact categories such as *ladder, bucket,* and *tongs,* exemplars were chosen to be similar in shape but not exact matches. Thus, all noun categories were as richly structured as possible given the limited number of instances. Table 7 lists the nouns, their solidity, category organization, and syntactic frames, and the three specific examples used in noun training.

Material-biased nouns. In the material-biased condition, children were taught 12 mass nouns that name nonsolid things in categories well organized by similarity in material substance (e.g., *frosting, Jell-O,* and *lotion*). In this condition, the three example items for each category were made from the same material but differed in amount, color, and shape. For instance, the Jell-O was either red, orange, or blue and was presented either as a large pile, as a couple of small piles, or in the shape of a teddy bear, and these shapes and amounts changed as the child ate the Jell-O. Likewise, the bottles and containers holding the exemplars were chosen to vary among the instances of a category. For example, the blue lotion was in a short fat bottle, the green lotion was in a taller, thinner bottle, and the white lotion was in a squeeze tube. Furthermore, the bottles and containers were removed as soon as possible—the substance was poured or squeezed onto a small plate when it was introduced, and the experimenter put the container out of sight. It was this plate of the substance that the child played with and heard named, rather than a bottle or container holding the substance. Thus, among the three instances of each category, no two exemplars matched exactly on shape or color, but all three exemplars were always the same material. The complete list of nouns, as well as the three specific examples used in noun training, is listed in Table 8.

Novel noun generalization stimuli. Eight sets of novel test stimuli were constructed for use in the novel noun generalization task. Four of these sets were made from solid materials such as wood and Styrofoam, and four were made from nonsolid materials such as hair gel and face cream. Each set consisted of an exemplar stimulus and four test objects. In each set, two

Table 7

Nouns, Their Solidity, Category Organization, and Syntax, as Well as the Three Examples Used in Noun Training, for the Natural Statistics Condition of Experiment 1

Noun	Solidity, category organization, and syntax	Example 1	Example 2	Example 3
1. pear	solid + sh + cn	Bartlett (green)	sickle (small, brown)	red
2. pretzel	solid + ma + cn	twist	stick	nugget
3. lollipop	solid + ma + cn	Tootsie Pop	Safe-t pop	Whistle pop
4. gummies	amb + ma + cn	blue sharks	red and green worms	peaches
5. Shredded Wheat	amb + ma + sh + amb	medium sized frosted	large plain biscuit	bite-sized
6. scarf	amb + sh + cn	child's black velvet	sheer blue neck	long pink knit
7. ladder	solid + sh + cn	short, wide red wood	white plastic	pink metal
8. chalk	solid + sh + ma + amb	short green	long white	fat yellow
9. bucket	solid + sh + cn	small red plastic	regular blue	white paper
10. funnel	solid + sh + cn	large blue	long, tiny clear plastic	regular metal
11. tong	solid + sh + cn	large wooden	small metal	medium plastic
12. glue	nonsolid + ma + mn	white glue	blue gel	red glue

Note. amb = ambiguous; sh = shape-based category; ma = material-based category; cn = count noun; mn = mass noun.

Table 8
Nouns, and Three Examples Used in Noun Training, for the Material-Biased Condition of Experiment 1

Noun	Example 1	Example 2	Example 3
1. frosting	chocolate	vanilla	strawberry
2. Jell-O	red	blue	orange
3. jelly	grape	apple	strawberry
4. pudding	chocolate	butterscotch	vanilla
5. oatmeal	regular	blueberry	brown sugar
6. lotion	blue	green	white
7. sand	magenta	green	yellow
8. Play Doh	yellow	red	blue
9. paint	blue	green	yellow
10. wax	red wax sticks	sheet of purple wax	small pieces of green wax
11. glitter	dark blue	green	pink
12. felt	black rectangle	purple shapes	tan shapes

test objects were the same shape as the exemplar but were of different colors and were made from different materials, and two test objects were made from the same material as the exemplar but were different in shape and color. For each child, two of the solid and two of the nonsolid stimulus sets were used to practice the novel noun generalization test at Sessions 3 and 6 of training (one solid and one nonsolid set per session), and a different solid set and nonsolid set were used in the novel noun generalization test at Session 9. The assignment of sets to practice and final novel noun generalization tests was counterbalanced across children. Figure 3 depicts examples of a solid stimulus set and a nonsolid stimulus set. Eight unique nonsense words were created for use in the novel noun generalization task. The pairing of names to stimulus sets was counterbalanced across children.

Twenty unique sets of familiar stimuli were also assembled for use in the familiar noun generalization task. These sets consisted of small toys familiar to most 15-month-olds such as balls, toy cars, and cups. Each set consisted of two identical toys and a third toy that differed in color, shape, and size (for example, two purple plastic eggs and a red wooden block).

Design

Children and their parents visited the lab once a week for 9 consecutive weeks. During all experimental sessions, the child sat across a large table from the experimenter with his or her parent. Every experimental session began with a familiar noun generalization task. This task was used both as a warm-up to the experimental session and as a practice of the novel noun generalization test that provided the main dependent variable in this study. At Sessions 1, 2, 4, 5, 7, and 8 children participated in naturalistic naming sessions. These naming sessions were interspersed with two practice novel noun generalization tests at Sessions 3 and 6. At Session 9 children participated in the novel noun generalization task. The practice novel noun generalization tests were exactly like the final test and did not include any feedback. They were included because pilot testing suggested that both the novel noun generalization task and the use of nonsolid stimuli were disconcerting for children this young.

The productive vocabulary of all the children was measured via parent report on the MCDI (Fenson et al., 1994) at Sessions 1 and 9 and at a

Figure 3. Examples of solid and nonsolid stimulus sets used in the novel noun generalization test of Experiments 1 and 2.

follow-up appointment 1 month after the final experimental session. Because participation in the material-biased condition may have led to an increase in names for nonsolid substances or material-based categories in children's vocabularies—names not well represented on the MCDI (Samuelson & Smith, 1999)—all parents were asked to note any other words their children produced that were not included on the MCDI. Reported vocabulary scores include these additional words.

Procedure

Familiar noun generalization task. On familiar noun generalization trials, the experimenter gave the child one set of practice stimuli to examine. After the child had examined the items, the experimenter retrieved the toys, put one of the matching pair and the nonmatching item on a tray, held up the other matching item, and said, "See this, this is my [name of toy]." She then pushed the tray toward the child, saying, "Can you get your [name of toy]?" If the child picked up or gestured toward the matching toy, she or he was praised heavily. If the child picked the incorrect toy, the experimenter said, "Is that the [name of toy]? No! Get the [name of toy]" until the child picked the correct toy.

Novel noun generalization task. This task was identical to the familiar noun generalization task with the exceptions that novel stimuli were used and there was no reinforcement of the child's responses. The child was given the exemplar, one shape-match test object, and one material-match test object from either a solid or nonsolid stimulus set to touch and examine for about 1 min. The experimenter then placed the test objects on the tray, held up the exemplar and said, for example, "See this, this is my bing." The experimenter then pushed the tray toward the child and while looking at the child, not the stimuli, said, "Can you get your bing?" If the child did not respond, he or she was prompted again. The experimenter said either "thank you," "ok," or "umm" in response to the child's choice and then proceeded to the next trial for that stimulus set. The parent was asked not to refer to the stimuli during this task but to encourage the child to respond. There were four trials for each stimulus set (each shape-match test object paired with each material-match test object) and one solid and one nonsolid stimulus set for a total of eight trials. Children never saw the same stimulus set twice. The order of solid and nonsolid sets for each child was counterbalanced across the practice and final novel noun generalization tests, and the order of stimulus sets was counterbalanced across children.

Neutral syntax (e.g., "This is my wug") was used in both the familiar and novel versions of the noun generalization task in accordance with previous investigations with young children (Imai & Gentner, 1997; Samuelson & Smith, 1999; Soja et al., 1991). Further, the use of a neutral syntactic frame is justified by research suggesting that children's command of count mass syntax is weak before 24 months of age (Gordon, 1985, 1988; Soja et al., 1991) and that the syntactic frame in which the noun is presented has a limited effect on children's novel noun generalizations before 30 months of age (Soja, 1992; Soja et al., 1991).

Naturalistic naming sessions. The naturalistic naming sessions were designed to mimic the natural way parents and children play with and name objects. The child was allowed to play freely with the examples of the noun categories as the experimenter and the parent named them. The three examples of each category were played with as a set for approximately 3 min. each. The experimenter then put these items away and brought out the examples of the next noun category. The experimenter named each noun category at least 20 times for each child during each session and encouraged the child to say each noun at least once. The order of noun category presentation was counterbalanced across children.

Because the 12 names taught in the laboratory were words that children typically learn early, a potential concern was that parents would try to support their child's performance in the lab by giving their child extra experience with items from the named categories at home. Of course, this would only be an issue if there was a systematic difference across conditions in parents' behavior as a result of experimental participation. For instance, parents of children in the material-biased condition might be biased to give their children more Jell-O or jelly to eat or more Play-Doh to play with and thus increase their children's exposure to the experimental names relative to children in the natural statistics condition. However, it did not feel appropriate to dictate to parents what they could or could not feed their children or that some common toys and household items were off-limits. Thus, it was decided to tell all parents that they should not alter their interactions at home in order to conform to the experiment. Thus, at the beginning of each session, parents were asked generally about the child's vocabulary acquisition since the last visit and, more specifically, about experiences with toys or stimuli similar to those used in the experiment. These conversations suggested that some children did get extra exposure to experimental stimuli at home, but these experiences were equally distributed across conditions and were not intensive in nature.

Coding

All experimental sessions were videotaped for later coding of naming instances and novel noun generalization responses. Four coders coded the total number of naming instances for a random sample of 67% of noun training sessions (approximately equally distributed across conditions). A random 20% of the coded sessions were coded by two coders. Agreement between coders in the number of naming instances was greater than 96%. All disagreements were resolved by a review of the videotape. In the natural statistics condition, the mean number of naming instances across the 12 nouns was 40.38 (range = 19–65). In the material-biased condition, the mean number of naming instances across the 12 nouns was 38.66 (range = 16–65). There was no significant difference in the number of naming instances across conditions, $t(296) = 1.55$, *ns*.

Three coders blind to the experimental hypothesis coded the novel noun generalization sessions. Coders indicated which test object the child picked in response to the experimenter's request on each trial. If a child did not make a clear selection, the trial was coded as "no response." Forty percent of the sessions were coded by two coders. Agreement between coders was 88%. This overall percentage was strongly affected by a lack of agreement between coders for 1 child in particular who tended to grab both test stimuli in quick succession. When this session was excluded (leaving 35% of sessions coded by two coders), agreement was 96%. All disagreements were resolved by review of the videotapes.

Results

Novel Noun Generalization

Figure 4 shows the mean proportion of shape choices out of total responses in the novel noun generalization test (Session 9) for children in the natural statistics and material-biased conditions. Data for both the solid and nonsolid stimulus sets are shown. To test the effectiveness of the noun training and to investigate differences in responding to solid and to nonsolid stimulus sets, the data were analyzed using a Condition (natural statistics vs. material-biased) × Test Set (solid vs. nonsolid) repeated measures ANOVA.[4] This analysis revealed a significant main effect of condition, $F(1, 9) = 13.00$, $p < .01$. No other significant main effects or interactions were found.

As can be seen in the figure, children in the natural statistics condition generalized names on the basis of shape similarity for both the solid and nonsolid stimulus sets, whereas children in the material-biased condition did not systematically generalize names

---

[4] Because children were assigned to conditions as matched pairs, the conditions are not independent. Therefore, both factors were treated as within-subject factors.
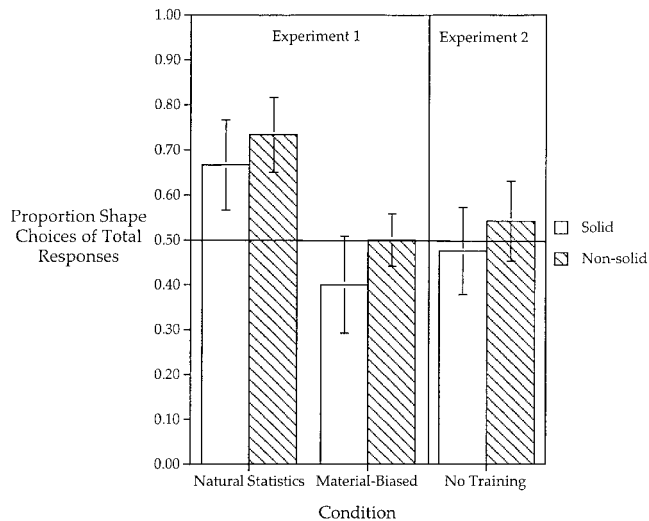
Figure 4. Mean proportions of shape choices out of total responses in the novel noun generalization test with solid and nonsolid stimuli for children in the natural statistics and material-biased conditions of Experiment 1 (left panel) and children in the no-training condition of Experiment 2 (right panel). Chance performance equals .50. Vertical bars depict standard errors of the means.

by either shape or material similarity. Because of the lack of significant effects involving test object, follow-up t tests collapsed across this factor compared performance in each condition with chance performance (.5). This analysis confirmed that children in the natural statistics condition generalized names on the basis of shape at levels significantly different from chance, $t(9) = 3.31$, $p < .01$, whereas children in the material-biased condition performed at chance levels, $t(9) = -.91$, *ns*.

Vocabulary Development

An important finding was that the noun training had an effect on children's vocabulary development outside the laboratory. Figure 5 shows the mean total vocabulary, noun vocabulary, count noun vocabulary, and other words vocabulary (all words except nouns) of children in the natural statistics and material-biased conditions at the beginning (Session 1) and end (Session 9) of noun training and at the follow-up appointment 1 month after the last experimental session. As can be seen in the figure, over the course of noun training, children in both conditions acquired words outside the laboratory at rates that did not differ. This was true for total vocabulary as well as for the smaller segments of vocabulary. By the 1-month follow-up appointment, however, children in the natural statistics condition had acquired more words than children in the material-biased condition. This accelerated vocabulary development was not restricted to the noun or count noun segments of the vocabulary. Rather, children in the natural statistics condition acquired other words (nonnouns) at accelerated rates as well. These conclusions were confirmed by dependent samples t tests. These analyses revealed no significant differences in the numbers of words acquired by children in the natural statistics and material-biased conditions between Sessions 1 and 9, all $ts(9) < 1.0$, *ns*, and significant differences between the groups in the numbers of words

acquired between Session 9 and the 1-month follow-up: for total vocabulary acquisition, $t(9) = 2.59$, $p < .05$; for noun vocabulary acquisition, $t(9) = 2.82$, $p < .05$; for count noun vocabulary acquisition, $t(9) = 3.12$, $p < .01$; and for other words acquisition, $t(9) = 2.05$, $p < .05$.

Discussion

The results of Experiment 1 provide conditional support for the statistical regularities hypothesis. Giving very young children additional naming experience with a set of nouns designed to match the kinds of nouns they would normally learn—a set dominated by names for solid objects in shape-based categories—created a precocious shape bias that, as predicted, was overgeneralized to nonsolid stimuli. Furthermore, this intensive naming experience appears to have led to an acceleration in productive vocabulary development in much the same way that teaching children names for four categories of artificial objects accelerated the count noun acquisitions of children in Smith et al.'s (2002) study. However, giving very young children additional naming experience with a set of nouns that fit a material bias—names for nonsolid substances in material-based categories—did not create a precocious material bias even though this set of nouns presented a perfect correlation between mass noun syntax, nonsolidity, and material-based categories. In addition, these children appear to have learned fewer words outside the laboratory compared with children taught a more typical set of nouns. It is possible, however, that the difference in the vocabularies of children in the two conditions at the follow-up appointment was actually due to a suppression of the vocabulary development of children in the material-biased condition. This possibility was investigated in Experiment 2.

Experiment 2

Experiment 2 provides a control for the possibility that the vocabulary development of children in the material-biased condition of Experiment 1 was harmed by the unusual kind of noun categories they were taught. Ten children participated in a longitudinal study exactly like Experiment 1 with the exception that these children did not receive intensive naming experiences. Thus, this experiment provides a measure of the typical vocabulary development of a matched set of children who repeatedly visited the laboratory and practiced the novel noun generalization task but who did not receive intensive naming experiences.

Method

Participants

Ten children, 4 boys and 6 girls between the ages of 15 and 21 months, participated (mean age = 18 months 2 days; range = 15 months 9 days to 21 months 7 days). Children were recruited from the child participants file at Indiana University. All children were learning English as their first and only language. Parents of potential participants were contacted by mail and recruited in a follow-up phone call. Potential participants were screened by asking parents approximately how many words their children could produce. At the first experimental session, children were matched on productive vocabulary scores from a subsection of the MCDI to children from Experiment 1 so that the mean ages and vocabulary scores across experiments and conditions did not differ: $F(2, 29) = 0.14$, *ns* for age, and $F(2, 29) = 0.01$, *ns* for total vocabulary. All children received a small prize
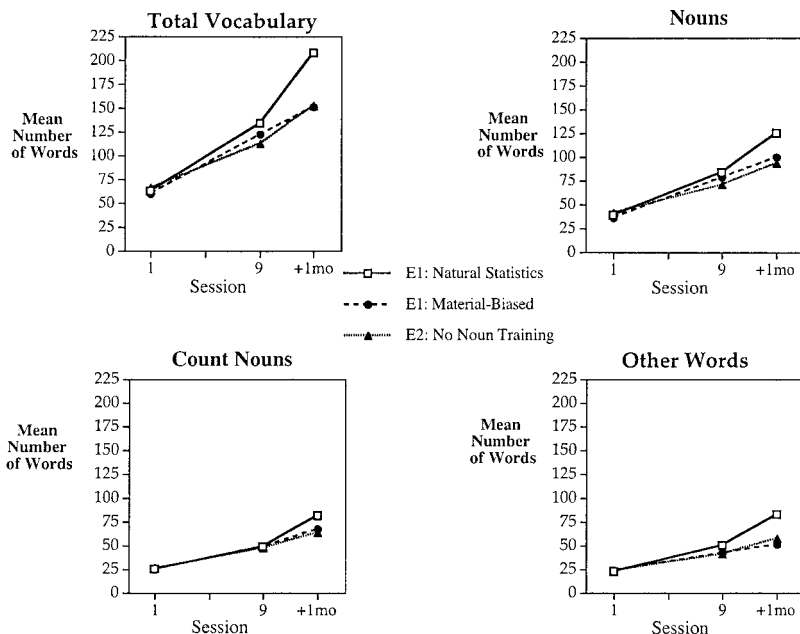
Figure 5. Mean productive total vocabulary, noun vocabulary, count noun vocabulary, and other words vocabulary for children in the natural statistics and material-biased conditions of Experiment 1 (E1) and the no-training condition of Experiment 2 (E2). Data are from three measurements of productive vocabulary: before (Session 1) and after (Session 9) noun training and at a 1-month (1 mo) follow-up appointment.

after each experimental visit, and the families received copies of experimental videotapes and T-shirts at the completion of the study.

Materials

The eight sets of novel noun generalization stimuli and 20 sets of familiar noun generalization stimuli were the same as those used in Experiment 1.

Design and Procedure

The design and procedure were the same as those in Experiment 1 with the exception that children did not participate in any naturalistic naming sessions. Thus, children in this experiment performed the familiar noun generalization task at every weekly session, practiced the novel noun generalization task with novel solid and nonsolid stimuli at Sessions 3 and 6, and participated in the novel noun generalization test at Session 9. And, like children in Experiment 1, the children's productive vocabulary development was measured via parental report at Sessions 1 and 9 and at a follow-up appointment 1 month after the novel noun generalization test.

Results

Novel Noun Generalization

The right-hand side of Figure 4 shows the mean proportion of shape choices out of the total responses on the novel noun generalization test (Session 9) for both the solid and nonsolid stimulus sets. As can be seen in the figure, children in Experiment 2 performed at chance levels (.50) on the novel noun generalization tasks for both stimulus sets: $t(9) = 0.26$, *ns* for the solid set, and $t(9) = 0.46$, *ns* for the nonsolid set. In addition, there was no difference in responding across the two stimulus sets, $t(9) = 0.04$, *ns*.

Vocabulary Development

Figure 5 shows the mean total vocabulary, noun vocabulary, count noun vocabulary, and other words vocabulary (all words except nouns) of children in Experiment 2 at the beginning (Session 1) and end of the experiment (Session 9) and at the follow-up appointment 1 month after the novel noun generalization test. As can be seen in the figure, there were no differences in the number of words acquired by children in Experiment 2 and children in either the natural statistics or material-biased conditions of Experiment 1 between Sessions 1 and 9, all dependent samples ts$(9) < 1.13$, *ns*. It is important to note that there were also no significant differences in the number of words acquired by children who did not receive noun training and children in the material-biased condition of Experiment 1 between the end of the experiment and the 1-month follow-up appointment, all dependent samples ts$(9) < 1.07$, *ns*. However, there were significant differences in the number of words acquired by children in Experiment 2 and children in the natural statistics condition of Experiment 1 between Session 9 and the follow-up appointment: for total vocabulary, dependent samples $t(9) = 2.06$, $p < .05$; for noun vocabulary, $t(9) = 2.19$, $p < .05$; for other words vocabulary, $t(9) = 1.84$, $p < .05$; and for count noun vocabulary, $t(9) = 2.28$, $p < .05$.

Discussion

The results of this experiment bolster the conclusions of Experiment 1. The children in this experiment, who visited the laboratory once a week for 9 consecutive weeks but did not receive intensive naming experiences, did not develop a shape or material bias and did not show accelerated vocabulary development. In fact,

the vocabulary development of children in this experiment was similar to that of children in the material-biased condition of Experiment 1. Therefore, giving the young children in the material-biased condition of Experiment 1 additional experiences naming categories of nonsolid substances well organized by similarity in material did not harm their vocabulary development. By contrast, the children in Experiment 2 did acquire significantly fewer new words between the end of the experiment and the 1-month follow-up appointment relative to children in the natural statistics condition of Experiment 1. Thus, giving the young children in the natural statistics condition of Experiment 1 additional naming experience with a set of nouns typical of the kind learned early in vocabulary development *did* accelerate their vocabulary growth relative to a matched sample of children who did not receive intensive noun training.

General Discussion

The goal of the research presented here was to test the hypothesis that word learning biases such as the shape bias reflect generalizations across the statistical structure of the nouns children learn early. Most of the nouns that children learn early are count nouns that name solid objects in shape-based categories, and in novel noun generalization tasks, children interpret novel nouns for novel solid objects as referring to shape-based categories. Fewer of the nouns that children learn early are mass nouns that name nonsolid substances in material-based categories, and in novel noun generalization tasks, children are inconsistent in their interpretations of names for novel nonsolid substances.

Three specific questions concerning the developmental relation between these phenomena were addressed. First, are the statistical regularities found in the early noun corpus of young children sufficient to support the development of a shape bias and insufficient to support the development of a material bias? Second, what aspects of these statistics might lead a statistical learner to a shape bias but not a corresponding material bias? And third, is it possible to create precocious word learning biases in children? The first two questions were addressed by using connectionist simulations to study the relation between the statistical regularities in the early noun vocabulary and the abstraction of a shape bias. The third question was addressed in two longitudinal experiments in which very young children were given intensive naming experiences designed either to replicate the regularities presented in natural vocabularies or to perturb those regularities in a direction that might also produce a material bias. The results of these simulations and experiments are considered in turn.

Better Evidence for a Shape Bias

Why might children know to attend to shape when naming a novel solid object before they know to attend to material when naming a novel nonsolid substance? Simulation 1 showed that the statistical regularities found in the early noun vocabulary of young children are sufficient to support a shape bias in a simple, unbiased learner of statistical regularities. The regularities are not sufficient to support a material bias. In fact, networks taught a vocabulary that contained the same statistical regularities found in the early noun vocabulary of children overgeneralized the shape bias from solid stimuli to nonsolid stimuli. The basis for the overgeneraliza-

tion of the shape bias was investigated in Simulation 2. Networks taught a vocabulary that contained equal numbers of names for solid objects and nonsolid substances, for shape- and material-based categories, and for count and mass nouns but that contained tight correlations between solidity, category organization, and syntax learned to differentiate solid objects and nonsolid substances in the novel noun generalization task. These networks generalized novel names for novel solid objects by shape and novel names for novel nonsolid substances by material. In contrast, networks taught a vocabulary that contained a high frequency of names for solid objects, names for shape-based categories, and count nouns but that had only weak correlations between solidity, category organization, and syntax developed a shape bias that was overgeneralized to nonsolid stimuli. These results suggest that the early dominance of the shape bias over the material bias is the result of the greater number of nouns that fit a shape bias compared to nouns that fit a material bias in the young child's productive vocabulary. Together, the simulation results suggest that the statistical regularities in the early noun corpus could support the development of a shape bias and predict that this bias should be overgeneralized to nonsolid stimuli.

Creating a Shape Bias

In the natural statistics condition of Experiment 1, very young children were given intensive naming experience with a set of nouns that contained the same statistical regularities found in the early noun vocabulary. The goal of this condition was to accelerate the development of a shape bias. After 9 weeks of noun training, children in the natural statistics condition generalized novel names for solid objects by shape at above chance levels. However, these children also generalized novel names for novel nonsolid stimuli by shape at above chance levels. Thus, these children developed a precocious shape bias and overgeneralized this shape bias to nonsolid stimuli.

The connectionist simulations suggest that this result was due to the high frequency of names that fit a shape bias (both in and outside of the laboratory) over the 9 weeks of noun training. That is, networks taught a vocabulary that did not include tight correlations between solidity, syntax, and category organization (the balanced and shape-dominated vocabularies) did not differentiate between solid and nonsolid stimuli in the novel noun generalization task. Likewise, children in the natural statistics condition may have been responding to the high frequency of nouns that fit a shape bias in their vocabulary but may not have learned that solid things, in general, are in categories organized by similarity in shape. Thus, when tested in the novel noun generalization task at the end of the experiment, children generalized novel names for both solid and nonsolid stimuli in accordance with the most dominant segment of their vocabulary.

A month after the completion of noun training, children in the natural statistics condition had significantly larger productive vocabularies than children in the other conditions. Intensive naming experience with a set of nouns that fit the shape bias led to a more general acceleration in word learning. The acceleration in object noun learning is a replication of the accelerated noun learning predicted and found by Smith et al. (2002). It is interesting, however, that children in the Smith et al. study were accelerated only in their learning of object names, whereas children in the

present study were accelerated in their acquisition of both nouns and other words. Several methodological differences between the two studies could be the source of this discrepancy.

In Smith et al.'s (2002) study, the four names taught to children were all count nouns, and the objects were all solid objects specially constructed to match exactly in shape. In contrast, the children in the present study were taught names for a more varied set of categories, and none of the exemplars matched exactly on any perceptual dimension. Perhaps the more richly structured categories used in the current experiment, along with the contrast between categories organized by shape and those organized by other dimensions, helped children learn something not just about how categories of solid objects are organized and named but about how other kinds of categories are organized and named. It is also possible that the use of real-world objects in the current study allowed parents to continue training outside the laboratory, thereby giving children more opportunities to learn both the nouns from the experiment and other related words. A final possibility is that the different results are due to the different ages of children in the current study (mean age = 18 months 28 days; range = 15 months 20 days to 20 months 3 days at the start of the study) and children in Smith et al.'s experiment (mean age = 17 months 1 day; range = 16 months 22 days to 17 months 15 days at the start). It is possible that children in the current study knew a higher proportion of "other" words at the beginning of the experiment and were thus farther along in the development of this section of their vocabularies.[5]

Creating a Material Bias

Children in the material-biased condition of Experiment 1 were given intensive naming experience with a set of nouns that presented a very different set of statistical regularities from those normally learned by young children. These children were taught 12 mass nouns that named nonsolid substances in categories well organized by material. The goal of this condition was to see whether a precocious material bias could be created by giving children intensive and frequent naming experiences with nouns that presented a perfect correlation between mass noun syntax, nonsolidity, and material-based categories.

After 9 weeks of training, children in the material-biased condition performed at chance levels with both the solid and nonsolid stimulus sets in the novel noun generalization task. Even though solidity, category organization, and syntax were perfectly correlated for these children, they did not demonstrate a material bias. Further, these children did not learn as many new words between the end of the experiment and the follow-up appointment as did children in the natural statistics condition.

Why did this group of children fail to develop a material bias? One possibility is that even though the noun training presented a perfect correlation between nonsolidity, material-based categories, and mass noun syntax, the frequency of these kinds of nouns was low relative to the number of names for solid objects, shape-based categories, and count nouns in the children's naturally developing noun vocabulary. This is consistent with the demonstrated importance of frequency in Simulation 2. A second and related reason children in the material-biased condition failed to develop a material bias may be because the experiment ended too soon. Perhaps the development of a material bias would simply have taken more

time, especially given that the statistical regularities of the nouns children were learning outside the laboratory worked against the statistics being emphasized in the laboratory. A third possibility is that children's previous experiences using materials such as Play Doh and paint to form representations of objects and then naming the represented objects rather than the material made learning material-based categories more difficult[6] (see Gelman & Ebeling, 1998, for data suggesting that children understand the use of materials to represent objects).

A fourth and most intriguing possibility is that something in the children's previous developmental history prevented them from developing a material bias. Perhaps the few nouns that 15–20-month-olds had when they entered the laboratory, although not enough to generate a reliable shape bias in the novel noun generalization task, were enough to prevent the development of a material bias. Data from young infants in habituation tasks suggest that names influence categorization well before children perform systematically in the novel noun generalization task (e.g., Schafer & Plunkett, 1996; Waxman & Markow, 1995; Woodward & Hoyne, 1999). Likewise, recent data from Waxman and Booth (2001) suggest that in very supportive word extension tasks involving repeated presentations of objects from familiar categories, explicit comparison of category members with nonmembers, and practice extending a novel name to a previously named example of a category, 14-month-olds extend novel count nouns in different ways than they extend novel adjectives.

These results clearly suggest that children even younger than those who participated in the current experiments understand something about how names link to object categories and—presumably, given that Waxman and Booth's (2001) stimuli were all rigid objects—that nouns refer to categories organized by shape. Thus, the data appear to conflict with previous findings that children do not demonstrate a shape bias until they already have many nouns in their productive vocabularies (Gershkoff-Stowe & Smith, 2001; Samuelson & Smith, 1999). This conflict is less apparent, however, in the larger context of the four-step process that has been proposed to underlie the development of the shape bias (Smith et al., 2002). Waxman and Booth's (2001) data suggest that by 14 months of age, infants have begun to extract the within-category similarities necessary for generalizing a newly learned name to novel instances of familiar categories—the second step in Smith et al.'s (2002) proposal. It is not clear, however, that these young infants have developed a generalized word learning bias that could be applied to completely novel stimuli in much more abstract and less supportive tasks such as the novel noun generalization task. This raises the larger issue, however, of where the statistical regularities that support the next steps in the development of a shape bias—and possibly hinder the development of a material bias—come from.

The Origins of the Statistical Regularities

Why do children learn so many names for solid objects in shape-based categories and so few names for nonsolid substances in material-based categories early in language development? One

---

[5] I would like to thank an anonymous reviewer for this suggestion.

[6] I would like to thank a second anonymous reviewer for this suggestion.

possibility is that the statistical structure of the early productive noun vocabulary reflects a prelinguistic perceptual bias for solid objects in shape-based categories. That is, names for solid objects may dominate because children's perceptual systems are tuned to pick out rigid, bounded shapes and solid, rigid objects that retain their shape over movement and other transformations. This idea fits Gentner's natural partitions hypothesis that nouns are easier to learn than relational terms because nouns refer to objects that are easily individuated whereas relational terms refer to nondiscrete, nonindividuated entities (Gentner, 1982; Gentner & Boroditsky, 2001). This idea also fits with data suggesting that even very young infants know something about solidity: They expect solid objects to follow connected paths through space, and they know that two solid objects cannot occupy the same space at the same time (Baillargeon & DeVos, 1991; Needham, Baillargeon, & Kaufman, 1997; Streri & Spelke, 1989; Xu & Carey, 1996). Moreover, data from Tremoulet, Leslie, and Hall (2000) and Wilcox (1999) suggest that even very young infants use shape information to individuate objects. A perceptual bias for solid, easily individuated objects cannot be the whole story behind the statistical regularities in the early noun vocabulary, however. Samuelson and Smith (1999) found no evidence that names for solid objects in shape-based categories are privileged in young children's acquisitions. Children do not learn names for solid objects in shape-based categories before they learn names for solid objects in other kinds of categories.

A second possible reason why names for solid objects in shape-based categories dominate the early noun vocabulary is that these kinds of nouns are more frequent in the input to children. According to this view, the statistical regularities in the early noun corpus that support the development of a shape bias are inherent not just in the nouns children learn but also in the nouns children hear. Data from a recent study by Sandhofer, Smith, and Luo (1999) support this possibility. Using a procedure similar to that used by Samuelson and Smith (1999), Sandhofer et al. (1999) examined the statistical regularities presented by 43 of the most common nouns produced by parents in naturalistic play sessions with their children. They found that most of these nouns—73%—were count nouns that named solid objects in shape-based categories. Thus, there may be a shape bias in the nouns children hear as well as in the nouns children learn.

It is possible, however, that a bias in parents' speech to children could be derived from a perceptual bias in children. Perhaps parents name more solid things in shape-based categories because children are most interested in these things. If this is true, then the most likely cause of the biased statistics in the early noun vocabulary is a combination of these factors. A small prelinguistic perceptual bias in the child together with a small bias in parents' speech to children could work together to produce the dominance of names for solid objects and shape-based categories found in the early noun vocabulary. Applied to the results of the present experiments, this idea suggests that one reason the natural statistics condition of Experiment 1 was so effective is that the nouns presented to children in this condition matched what they were already primed to learn. The statistical regularities created a precocious shape bias and accelerated vocabulary acquisition in these very young children because their developmental history of perceiving and naming set them up to learn just the kinds of statistical regularities presented by the nouns in this condition. Further, the

reason children in the material-biased condition did not develop a material bias—even though they were presented with a perfect correlation between nonsolidity, category organization by material, and mass noun syntax—was that the statistical regularities presented in this condition went against what their developmental history had primed them to learn.

Conclusions

The results of these simulation studies and experiments strongly suggest that the statistical regularities in the early noun corpus support the development of the shape bias but not the development of a material bias. In this way, the simulations and experiments presented here suggest a mechanistic account of how children abstract a specific word learning bias from regularities found in their productive noun vocabularies. Further, the fact that children who developed a precocious shape bias also demonstrated accelerated vocabulary acquisition suggests the possibility that the statistical regularities in the early noun corpus also support children's amazing word learning skills more generally. These conclusions are tempered, however, by the failure to create a precocious material bias.

This later finding suggests the important role that the longer developmental history of word learning plays in the development of word learning biases. This research has focused on only two of the many steps taken in the development of the shape bias. The more general goal in this program of research is to develop a mechanistic understanding of the longer trajectory of word learning—from the initial links between words and categories that shape the correlations underlying the statistical regularities, to the brute-force acquisition of individual object names based on specific inputs provided by parents and the environment, to the development of the many context-sensitive word learning biases that make children such smart word learners. The present results inform this understanding by suggesting how the development of the shape bias fits in this longer chain of events. That is, perhaps a small perceptual bias together with a bias in the input children receive creates a statistical bias in the early noun vocabulary. This statistical regularity in the words children know, in turn, leads to the development of a shape bias, which, in the next developmental step, leads to an acceleration in vocabulary acquisition. Each development guides the next, and each subsequent development depends on the previous advances such that, in the end, young children *become* smart noun learners.

References

Baillargeon, R., & DeVos, J. (1991). Object permanence in young infants: Further evidence. *Child Development, 62,* 1227–1246.

Brown, R. (1957). Linguistic determinism and the part of speech. *Journal of Abnormal and Social Psychology, 55,* 1–5.

Clark, E. V. (1973). What's in a word? On the child's acquisition of semantics in his first language. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 65–110). New York: Academic Press.

Dale, P. S., & Fenson, L. (1993). *LEX: A lexical development norms database.* Seattle, WA: University of Washington, Department of Psychology.

Dickinson, D. K. (1988). Learning names for material: Factors constraining

and limiting hypotheses about word meaning. *Cognitive Development, 3,* 15–35.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D., & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5, Serial No. 242).

Gelman, S. A., & Ebeling, K. S. (1998). Shape and representational status in children's early naming. *Cognition, 66,* B35–B47.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj II (Ed.), *Language development: Vol. 2. Language, thought and culture* (pp. 301–334). Hillsdale, NJ: Erlbaum.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). New York: Cambridge University Press.

Gershkoff-Stowe, L., & Smith, L. B. (2001). *Shape and the first hundred nouns.* Manuscript submitted for publication.

Gordon, P. (1985). Evaluating the semantic categories hypothesis: The case of the count/mass distinction. *Cognition, 20,* 209–242.

Gordon, P. (1988). Count/mass category acquisition: Distributional distinctions in children's speech. *Journal of Child Language, 15,* 109–128.

Hall, D. G., Lee, S. C., & Belanger, J. (2001). Young children's use of syntactic cues to learn proper names and count nouns. *Developmental Psychology, 37,* 298–307.

Harris, M., & Coltheart, M. (1986). *Language processing in children and adults: An introduction.* London: Routledge & Kegan Paul.

Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist, 39,* 1372–1388.

Imai, M. (2000). Universal ontological knowledge and a bias toward language-specific categories in the construal of individuation. In S. Niemeier & R. Dirven (Eds.), *Evidence for linguistic relativity* (pp. 139–160). Amsterdam: John Benjamins.

Imai, M., & Gentner, D. (1997). A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition, 62,* 169–200.

Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development, 9,* 45–75.

Jones, S. S., & Smith, L. B. (1993). The place of perception in children's concepts. *Cognitive Development, 8,* 113–139.

Jones, S. S., & Smith, L. B. (1998). How children name objects with shoes. *Cognitive Development, 3,* 323–334.

Jones, S. S., Smith, L. B., & Landau, B. (1991). Object properties and knowledge in early lexical learning. *Child Development, 62,* 499–516.

Katz, N., Baker, E., & Macnamara, J. (1974). What's in a name? A study of how children learn common and proper names. *Child Development, 45,* 469–473.

Kelly, M. H., & Martin, S. (1994). Domain-general abilities applied to domain-specific tasks: Sensitivity to probabilities in perception, cognition, and language. *Lingua, 92,* 105–140.

Kemler Nelson, D. G., Russell, R., Duke, N., & Jones, K. (2000). Two-year-olds will name artifacts by their functions. *Child Development, 71,* 1271–1288.

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

Macnamara, J., & Reyes, G. E. (1994). Foundational issues in the learning of proper names, count nouns and mass nouns. In J. Macnamara & G. E. Reyes (Eds.), *The logical foundations of cognition. Vancouver Studies in Cognitive Science* (pp. 144–176). New York: Oxford University Press.

Markman, E. M. (1989). *Categorization and naming in children.* Cambridge, MA: MIT Press.

Markman, E. M. (1992). Constraints on word learning: Speculations about their nature, origins, and domain specificity. In M. R. Gunnar & M.

Maratsos (Eds.), *Minnesota Symposia on Child Psychology: Vol. 25. Modularity and constraints in language and cognition* (pp. 59–101). Hillsdale, NJ: Erlbaum.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88,* 375–407.

Mervis, C. B. (1987). Child-basic object categories and lexical development. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 201–233). Cambridge, England: Cambridge University Press.

Mervis, C. B., Mervis, C. A., Johnson, K. E., & Bertrand, J. (1992). Studying early lexical development: The value of the systematic diary method. *Advances in Infancy Research, 7,* 291–378.

Movellan, J. R. (1991). Contrastive Hebbian learning in the continuous Hopfield model. In D. S. Touetzky, J. E. Elman, T. J. Sejnowski, & G. E. Hinton (Eds.), *Connectionist models, proceedings of the 1990 summer school* (pp. 10–17). San Mateo, CA: Morgan Kaufmann.

Needham, A., Baillargeon, R., & Kaufman, L. (1997). Object segregation in infancy. *Advances in Infancy Research, 11,* 1–44.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development, 38*(1–2, Serial No. 149).

Quine, W. V. O. (1960). *Word and object.* Cambridge, MA: MIT Press.

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

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

Samuelson, L. K., & Smith, L. B. (2000). Children's attention to rigid and deformable shape in naming and non-naming tasks. *Child Development, 71,* 1555–1570.

Sandhofer, C., Smith, L. B., & Luo, J. (1999). Counting nouns and verbs in the input: Differential frequencies, different kinds of learning? *Journal of Child Language, 27,* 561–585.

Schafer, G., & Plunkett, K. (1996). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development, 69,* 309–320.

Smith, L. B. (1995). Self-organizing processes in learning to learn words: Development is not induction. In C. A. Nelson (Ed.), *Minnesota Symposia on Child Psychology: Vol. 28. Basic and applied perspectives on learning, cognition, and development* (pp. 1–32). Mahwah, NJ: Erlbaum.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. K. (2002). Object name learning provides on-the-job training for attention. *Psychological Science, 13,* 13–19.

Soja, N. N. (1992). Inferences about the meanings of nouns: The relationship between perception and syntax. *Cognitive Development, 7,* 29–45.

Soja, N. N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition, 38,* 179–211.

Streri, A., & Spelke, E. S. (1989). Effects of motion and figural goodness on haptic object perception in infancy. *Child Development, 60,* 1111–1125.

Subrahmanyam, K., Landau, B., & Gelman, R. (1999). Shape, material, and syntax: Interacting forces in children's learning of novel words for objects and substances. *Language and Cognitive Processes, 14,* 249–281.

Tomasello, M., & Akhtar, N. (1995). Two-year-olds use pragmatic cues to differentiate reference to objects and actions. *Cognitive Development, 10,* 201–224.

Tomasello, M., & Mervis, C. B. (1994). The instrument is great, but measuring comprehension is still a problem. Commentary on Fenson et al. *Monographs of the Society for Research in Child Development, 59*(5, Serial No. 242).

Tremoulet, P. D., Leslie, A. M., & Hall, D. G. (2000). Infant individuation and identification of objects. *Cognitive Development, 15,* 499–522.

Waxman, S. R., & Booth, A. E. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology, 43,* 217–242.

Waxman, S. R., & Hall, D. G. (1993). The development of a linkage between count nouns and object categories: Evidence from fifteen- to twenty-one-month-old-infants. *Child Development, 64,* 1224–1241.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology, 29,* 257–302.

Wilcox, T. (1999). Object individuation: Infants' use of shape, size, pattern, and color. *Cognition, 72,* 125–166.

Woodward, A. L., & Hoyne, K. L. (1999). Infants' learning about words and sounds in relation to objects. *Child Development, 70,* 65–77.

Xu, F., & Carey, S. (1996). Infant's metaphysics: The case of numerical identity. *Cognitive Psychology, 30,* 111–153.

Appendix

The Network's Activation, Input, and Weight Change Functions

The activation function for all units in the network is the interactive activation rule of McClelland and Rumelhart (1981).

If $h_i^t > 0$,

$$\Delta a_i^t = h_i^t(a_i^{max} - (a_i^{t-1} - D_i a_i^{t-1})),$$

else,

$$\Delta a_i^t = h_i^t((a_i^{t-1} - D_i a_i^{t-1}) - a_i^{min}), \tag{A1}$$

where h_i^t is the input to unit i at time t; a_i^t is the activation of unit i at time t; and a_i^{max}, a_i^{min}, and D_i are the constant maximum activation, minimum activation, and decay rate, respectively, associated with i.

The input to a unit from other units is given by

$$h_i^t = \sum_{j=1}^{n} a_j^t \cdot w_{ij}, \tag{A2}$$

where n is the number of units in the network and w_{ij} is the weight connecting units i and j.

Finally, the weight update rule for a connection between two units is

$$\Delta w_{ij}^t = L \cdot a_i^t \cdot a_j^t, \tag{A3}$$

where L is a constant learning rate. In the present simulations, L was always .01.