PAPER

Developmental changes in visual object recognition between 18 and 24 months of age

Alfredo F. Pereira and Linda B. Smith

Department of Psychological and Brain Sciences, Indiana University Bloomington, USA

Abstract

Two experiments examined developmental changes in children's visual recognition of common objects during the period of 18 to 24 months. Experiment 1 examined children's ability to recognize common category instances that presented three different kinds of information: (1) richly detailed and prototypical instances that presented both local and global shape information, color, textural and featural information, (2) the same rich and prototypical shapes but no color, texture or surface featural information, or (3) that presented only abstract and global representations of object shape in terms of geometric volumes. Significant developmental differences were observed only for the abstract shape representations in terms of geometric volumes, the kind of shape representation that has been hypothesized to underlie mature object recognition. Further, these differences were strongly linked in individual children to the number of object names in their productive vocabulary. Experiment 2 replicated these results and showed further that the less advanced children's object recognition was based on the piecemeal use of individual features and parts, rather than overall shape. The results provide further evidence for significant and rapid developmental changes in object recognition during the same period children first learn object names. The implications of the results for theories of visual object recognition, the relation of object recognition to category learning, and underlying developmental processes are discussed.

Introduction

Human visual object recognition is impressive in several ways: it is fast, seemingly automatic, robust under degraded viewing conditions, and capable of recognizing novel instances of a very large number of common categories (Cooper, Biederman & Hummel, 1992; Fize, Fabre-Thorpe, Richard, Doyon & Thorpe, 2005; Pegna, Khateb, Michel & Landis, 2004). For example, in their everyday lives, people routinely recognize the dog whose nose is sticking out from the blanket, the highly unique modernistic chair, and the cup on the table as a particular and favorite cup. Competing theories of object recognition (Biederman, 1987; Edelman, 1999; Ullman, 1996) often pit different kinds of hypothesized processes and representations against each other. However, it seems likely that human object recognition is dependent on a multitude of partially distinct and partially overlapping processes (Hayward, 2003; Hummel, 2000; Marr, 1982; Peissig & Tarr, 2007; Peterson, 1999). That is, no single mechanism is likely to explain the full range of contexts in which people recognize objects as individuals and as instances of categories.

The experiments reported in this paper are concerned with developmental changes in children's recognition and categorization of common objects, changes that occur during the same period that children first learn object names. A connection between the representation of shape and the learning of objects names makes sense as many common object categories are (by adult judgment) well organized by shape (Rosch, 1973; Samuelson & Smith, 1999). But an open question in the object recognition and categorization literature is the proper psychological description of object shape. A critically important developmental question is whether that description changes as children learn object categories.

Shape and learning object names

As young children learn object names, they appear to increasingly attend to object shape in lexical categorization tasks. One widely used task is novel noun generalization (Landau, Smith & Jones, 1988). For example, children might be shown a novel object of a particular shape and told its name 'This is a dax!' They are then asked what other objects have the same name. Two- and 3-year-old children systematically generalize the name to new instances by shape (e.g. Colunga & Smith, 2005; Gathercole & Min, 1997; Imai, 1999; Imai, Gentner & Uchida, 1994; Keil, 1994; Soja, 1992; Yoshida & Smith, 2003). Further studies show that this shape bias develops. Very young children (12- to 18-month-olds) do not attend to object shape in naming tasks as systematically as do older children. Instead, attention to shape increases during the period between 18 and 30 months (e.g. Gershkoff-Stowe

Address for correspondence: Alfredo F. Pereira, Department of Psychological and Brain Sciences, 1101 East Tenth Street, Bloomington, IN 47405, USA; e-mail: afpereir@indiana.edu

& Smith, 2004; Rakison & Butterworth, 1998a). In addition, attention to shape is developmentally related to the number of object names in children's vocabularies, emerging when children have between 50 and 150 object names in their productive vocabulary (Gershkoff-Stowe & Smith, 2004; Smith, 2003). Longitudinal studies suggest further that the shape bias is temporally linked in individual children to a measurable spurt in the growth of object name vocabulary (Gershkoff-Stowe & Smith, 2004). Finally, training studies show that teaching children to attend to shape facilitates novel noun acquisitions and accelerates the rate of real world vocabulary development (Smith, Jones, Landau, Gershkoff-Stowe & Samuelson, 2002). All these results point to a link between learning common object category names and attention to shape in categorization tasks.

One unresolved issue central to understanding these phenomena is just how children perceive, represent and compare object shapes. In order for a shape bias to work in learning real object categories, children must be able to recognize sameness in shape. This is a trivial problem in laboratory versions of the shape bias task (in which all objects are simple and in which same-shaped objects are the exact same shape), but it is not trivial in the real world. In order for children to learn, for example, that chairs are 'chair-shaped' and to use that knowledge to recognize a new chair, they must be able to abstract the common shape from the whole array of experienced chairs, each with its own unique detailed shape. Members of the same real object category, even one seemingly well organized by shape, are not exactly the same shape, but only similar in shape at some appropriate level of abstraction. Thus, two critical questions are: What is the proper description of shape for common object categories? When and how do children discover that description?

Holistic and structural representations of shape

Theories of adult object recognition suggest several different ways of specifying shape. According to 'view-based' theories, people store representations of specific views of experienced instances. Identification, recognition and categorization are accomplished with reference to these stored exemplars (e.g. Edelman, 1999; Edelman & Bülthoff, 1992; Tarr & Pinker, 1989; Ullman, 1996). Edelman and his colleagues (Edelman, 1995; Edelman & Duvdevani-Bar, 1997; Edelman & Intrator, 1997) suggest further that the shape representations relevant for object categorization are a product of learning those categories. In this account, category learning creates prototypes of the holistic shape of category members. Novel instances are subsequently categorized by their overall similarity to these representations. Two critical ideas from this account are that shape representations are holistic blends of experienced instances and that they are learned as categories are learned.

'Object-based' theories such as Biederman's (1987) Recognition-by-Components (RBC) account present
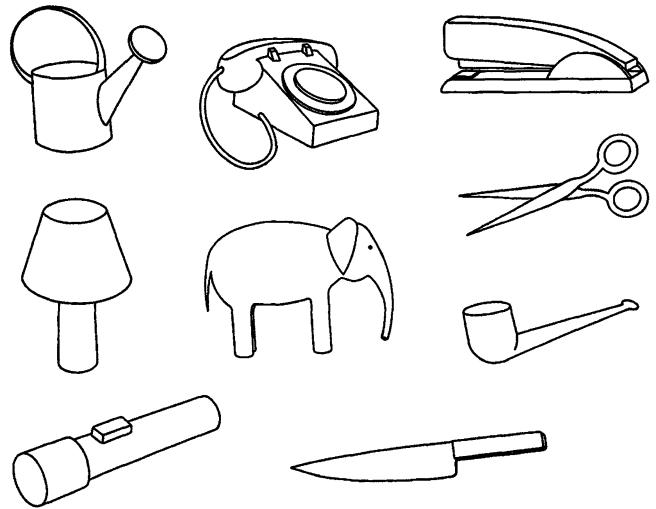
Figure 1 *Example of stimuli used in Biederman's 1987 study. Adults can recognize common object categories from only 2–4 major parts present. Stimuli taken from Biederman (1987, p. 130).*

another idea about what constitutes 'sameness in shape'. This theory proposes that objects are perceptually parsed, represented, and stored as configurations of geometric volumes ('geons'). Within this account, object shape is defined by two to four geometric volumes in the proper spatial arrangement, an idea supported by the fact that adults need only two to four major parts to recognize instances of common categories (Biederman, 1987; Biederman & Gerhardstein, 1993; Hummel & Biederman, 1992) as illustrated in Figure 1. This account thus posits sparse and impoverished representations that, through their high level of abstraction, can gather all variety of highly different things into a 'same shape' category. The critical idea from this account is that category relevant descriptions of object shape are abstract descriptions of the relational structure of a few major parts.

Both classes of theories suggest sparse representations of global shape and both fit aspects of the adult data, which include strong view dependencies in object recognition and also knowledge of part structure and relations. Accordingly, there is a growing consensus that both kinds of theories may capture important but different processes in mature object recognition (Hayward, 2003; Peissig & Tarr, 2007; Peterson, 1999; Stankiewicz, 2003; Tarr & Vuong, 2002).

Development of object recognition

There are few studies of the early development of either aspect of object recognition (Kellman, 2001). However, one recent study examined whether very young children (18 to 24 months) could recognize instances of common object categories from sparse representations of the structure of major geometric parts, as proposed in Biederman's RBC theory (Smith, 2003). The experiment

Figure 2 *Examples of Typical (bottom row) and Shape Caricature (top row) objects from the Smith 2003 study (p. 245).*

specifically contrasted richly detailed typical examples with Shape Caricatures as shown in Figure 2. The task was name comprehension ('get the camera'), and the 18- to 24-month participants were grouped into developmental level by the number of object names in their productive vocabulary.

The main results were that children with smaller and larger vocabularies (below 100 object names versus more than 100 object names) recognized the richly detailed instances equally well. However, children with smaller noun vocabularies performed at chance levels when presented with the Shape Caricatures, whereas the children with high noun vocabularies recognized the Shape Caricatures *as well as they did the richly detailed and typical instances*. These results have been replicated in a second study (Son, Smith & Goldstone, under review). Further, a study of older late talkers with limited object names in their productive vocabularies also found a deficit in the recognition of Shape Caricatures but not richly detailed typical instances (Jones & Smith, 2005). These results suggest a potentially significant change in how young children represent and compare object shape that is developmentally linked to the learning of objects names. In particular, sparse representations of object shape appears to emerge between 18 and 24 months.

One other line of research also suggests possible developmental changes in the stimulus information used to categorize and recognize objects. These studies suggest that children younger than 20 months attend to the individual parts or local details of objects rather than overall shape (Quinn, 2004a; Rakison & Butterworth, 1998a). In a series of programmatic studies, Rakison and colleagues (Rakison & Butterworth, 1998b; Rakison & Cohen, 1999) showed that 14- and 22-month-old children based category decisions on highly salient parts (such as legs and wheels) and not on overall shape. For example, when, children were presented with cows whose legs had been replaced by wheels, they classified the cows with vehicles rather than animals; likewise they categorized a vehicle as an animal when it had cow legs. Similarly,

Colunga (2003) showed that 18-month-olds tended to look at only a small part of any pictured object, using clusters of local features such as the face when recognizing animals, or the grill and headlights when recognizing vehicles. These results raise the possibility that very young children – perhaps before they develop more sparse representations of object structure – recognize objects via what Cerella (1986) called 'particulate perception', concentrating on local components unintegrated into the whole. Younger children's 'part'-based object recognition is also suggestive of an approach to object recognition that has emerged in the machine vision literature: in particular, Ullman has developed a procedure through which objects are successfully recognized via stored representations of category-specific fragments (Ullman & Bart, 2004; Ullman, Vidal-Naquet & Sali, 2002).

Motivation for the present study

The purpose of the two empirical studies that follow is to provide greater insight into developmental changes in the recognition of common object categories between the period of 18 to 24 months. It seems likely that mature perceivers use many different sources of information, including (but probably not limited to) local clusters of features or fragments (enabling, for example, the recognition of the dog from the dog nose sticking out from the blanket), holistic descriptions of overall shape prototypes, perhaps of the kinds hypothesized by Edelman (1999), and sparse descriptions of the structural relations among a few major parts as proposed by Hummel and Biederman (1992). However, the relative importance and availability of these different sources of information relevant to object categorization and recognition may also change with development and perhaps as a direct consequence of category learning.

In this study, we examine three potential sources of information as illustrated in Figure 3. The first is the information available in a small local region of the object. One need not necessarily take in or integrate across the whole object to know the object's category. Fine-grained information about texture, color, and shape of a local area or fragment might well be sufficient if the properties of that local region are typical of past experienced instances. A second kind of information that may be relevant to children's object recognition is the detailed shape of the whole object at multiple spatial frequencies. This is the kind of information, for example, that might be holistically compared to a prototype representation of the shape of frequently experienced instances. If the whole detailed shape is sufficiently similar to previous instances then it should be recognizable as a member of the category. A third kind of information concerns the geometric structure of the whole devoid of any surface details and limited to lower spatial frequencies. This is the kind of information relevant to recognition and categorization via a few geometric volumes in the proper arrangement as proposed by Biederman's RBC, a kind
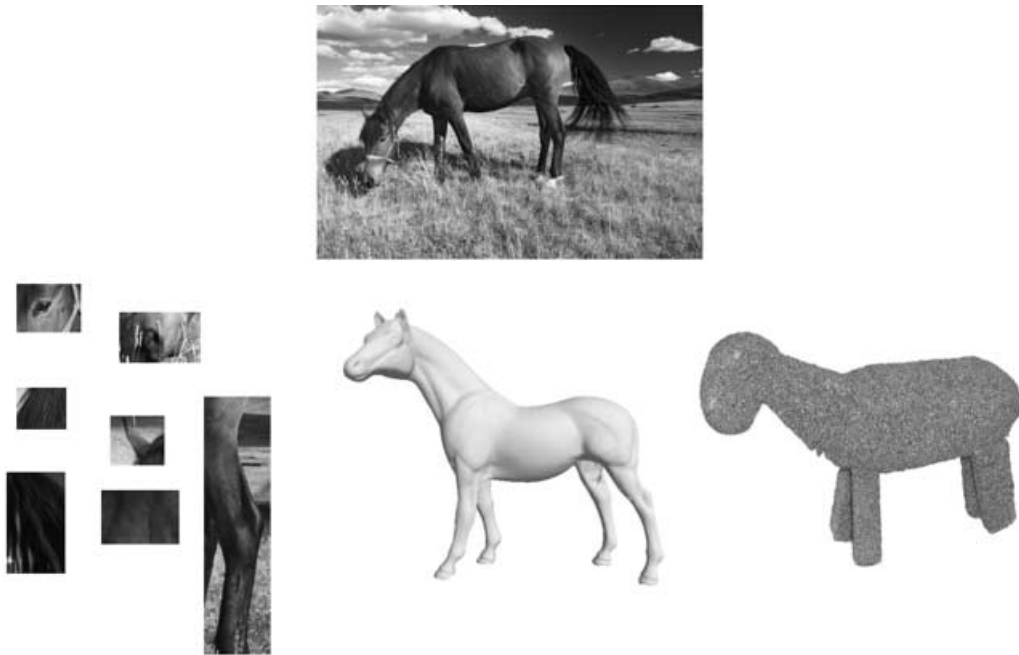
Figure 3 *Recognizing that the object in the scene (top row) contains a horse can potentially be achieved by using multiple sources of information. One possibility is to statistically combine the most predictive object fragments (potentially at multiple spatial scales) as depicted in the left image, bottom row. These are highly predictive of the object's category by similarity to previously encountered instances (e.g. Ullman et al., 2002). Another possibility is to segment figure from ground and holistically compare this specific object view to previously stored views of object. This can be accomplished in multiple ways, for example by interpolation with stored views (e.g. Edelman, 1999; Ullman, 1996) as illustrated in center image, bottom row. A final possibility is to extract an edge-detected version of the horse and from that, detect the most important geometric volumes and their spatial relations. These are placed in a structural description of the object in the scene. Such structural description can be matched with a previously learned representation of the category. This representation could contain, for example, the structural elements illustrated in the right image, bottom row (e.g. Hummel & Biederman, 1992).*

of representation that previous work suggests may emerge between 18 and 24 months.

The two experiments that follow examine $1\frac{1}{2}$- to 2-year-old children's use of these three different kinds of information. As in Smith's (2003) previous study, children are grouped by productive vocabulary size, as reported by the child's parent, rather than age. Parent report of productive vocabulary is an imperfect measure of children's individual word and category knowledge, and a conservative one in that receptive vocabulary, particularly at young age levels, is typically much larger than productive vocabulary (Tomasello, 1994) and, moreover, may not even be straightforwardly related to productive vocabulary in individuals (Bates, Dale & Thal, 1995). However, parent report of productive vocabulary has proven a reliable global measure of lexical development and highly predictive of performances in categorization tasks (e.g. Bates *et al.*, 1995; Smith *et al.*, 2002). Thus, parent report of productive vocabulary may be a more relevant index of object category knowledge in this period of rapid development than is age.

Experiment 1 compares children's recognition of common object categories given three different kinds of stimulus sets: (1) richly detailed and *Typical* instances that present rich shape information as well as typical texture, color, and surface features; (2) *Rich Shape* instances

that present the highly detailed and prototypical shapes but with no color, texture or surface featural information, and (3) *Shape Caricatures* that provide only a sparse description of shape via a few major parts in their proper spatial arrangement. Experiment 2 examines children's use of local part information to recognize objects and directly compares that to their use of global geometric structure.

Experiment 1

Following the procedure of Smith (2003), the task is name comprehension. Children are presented with three alternatives, all instances of everyday categories, and asked to indicate one named by the experimenter. The main manipulation is kind of stimulus information available as indicated in Figure 4.

Method

Participants

Sixty-four children (33 female, 31 male) were recruited from a working- and middle-class population in a Midwestern college town. All were native speakers of

Figure 4 *Photos of the stimuli used in Experiment 1. This example shows the three levels of shape detail in six of the 18 test categories (airplane, butterfly, car, hammer, horse and turtle). The black line close to each object is 1 inch in length.*

English and had no known neurological or language disorders. They ranged in age from 16.5 months to 29.0 months. Eleven additional subjects began the experiment but did not contribute data, because the parent coached the child (contrary to instructions) during the experiment, for fussiness, or for failure to understand the task.

Stimuli

Eighteen object categories and prototypical instances of those categories were selected such that by pilot testing of receptive word knowledge all were recognizable by 75% of a sample of 18- to 24-month-olds. The categories were: airplane, boat, butterfly, cake, car, cow, dog, fish, frog, girl, hamburger, hammer, horse, pig, sheep, shoe, tree and turtle. The Typical instances were store-bought richly detailed and prototypical toy instances of the target categories. Each Rich Shape instance was constructed from a duplicate of the corresponding Typical toy; the store-bought originals were covered with clay, coated with wax and then painted black so as to maintain most of the shape details of the original but to remove information about texture, color, and fine-grained surface features. The Shape Caricature instances were constructed from Styrofoam volumes and designed to represent the major part structure with the minimum number of parts and as such roughly fit the global and sparse structure

of both category level shape representations as proposed by Biederman (Hummel, 1992, p. 211) and Edelman (Edelman, 1999, p. 244). Across instances, the number of parts varied from one to seven ($M = 4.0$, $SD = 2.0$). In all conditions, all instances averaged 90 cm^3 in volume. Figure 4 shows a subset of the stimuli used.

An additional nine toy objects were used in a warm-up phase prior to the main experiment: bear, ball, banana, bottle, duck, flower, cup, block, carrot and spoon.

Language measure

Parents were asked to indicate the number of count nouns in their children's productive vocabulary using the Bates-MacArthur Communicative Development Inventory (MCDI) (Fenson, Dale, Reznick, Bates, Thal & Pethick, 1994). We measured only count nouns because these are the nouns that label common object categories and because past research suggests that it is specifically the size of count noun vocabulary that predicts children's attention to shape (see Gershkoff-Stowe & Smith, 2004; Smith, 2003; Smith *et al.*, 2002). Parents were specifically asked to indicate the nouns they had heard their child produce.

Procedure

The experiment proper began with a warm-up phase. The purpose of this phase was to make clear to the child that their task in the main experiment was to select the one object from three alternatives that was named by the experimenter. The warm-up began with the experimenter presenting the child with three objects from the warm-up set. These were placed in segregated sections on a 72 cm by 23 cm tray. With the tray held so the child could see all three objects but out of reaching distance, the experimenter directed the child's attention to each object without naming. Then with the tray still out of reach, the experimenter named the target object several times (e.g. 'I want the carrot! Get me the carrot! The carrot!'). The tray was then pushed forward. Pointing responses or picking up the object was taken as a response. On these warm-up trials, feedback was given and, if necessary, the child was helped to reach to the correct object. Placement of the target object (left, center or right) was counter-balanced. There were a minimum of four warm-up trials and a maximum of six. All children who contributed data moved to the main experiment when they had successfully reached for the named object at least three times without help.

The main experiment was structured identically to the warm-up trials except that no feedback was given. On each trial, children were presented with three objects from the same stimulus condition (Typical instances, Rich Shape instances, or Shape Caricatures) and asked to get one by name. All children received six trials with Typical instances, six with Rich Shape instances, and six with Shape Caricatures. No child ever saw a different version of the same object (e.g. if an individual child was

Table 1　*Number of subjects, mean value and range for age and number of object names in productive vocabulary as measured by the MCDI, for Group I, II and II in Experiment 1*

	Group I – Children knowing less than 51 object names (*n* = 20)		Group II – Children knowing between 51 and 150 object names (*n* = 22)		Group III – Children knowing more than 150 object names (*n* = 22)	
	M	Range	*M*	Range	*M*	Range
Developmental indicator						
Age (months)	19.3	16.5–24.4	22.0	16.7–27.2	24.1	19.6–29.2
Number of object names in productive vocabulary	18.9	0–50	92.8	51–149	221.4	151–309
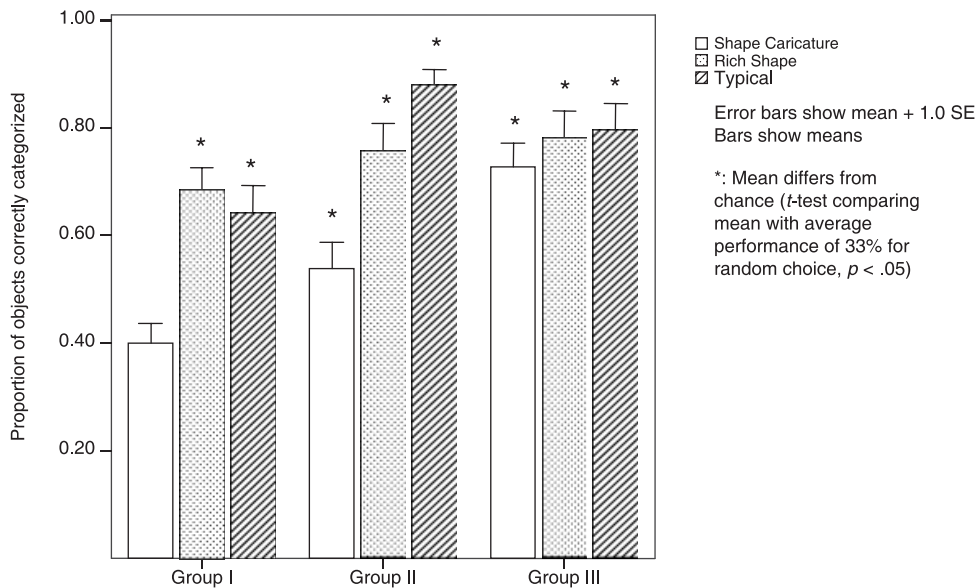
Figure 5　*Mean proportion of number of objects correctly categorized (out of six trials) in Experiment 1 across the three groups of vocabulary level for Typical, Rich Shape and Shape Caricature objects.*

assigned the Typical turtle, they did not see on any other trial the Rich Shape turtle, or the Shape Caricature turtle). Across children, each target object served equally often in each stimulus condition and all objects served equally often as distracters. The placement of the target object (left, center or right) was also counterbalanced across trials. The parent was instructed not to name objects or indicate correct answers and children's data were excluded if parents did not follow instructions. The experimental session lasted less than 15 minutes.

Results and discussion

The original Smith (2003) study compared two groups of children – those above and below 100 count nouns in productive vocabulary. Previous longitudinal research, however, suggests potentially relevant changes in children's categorization and attention to shape, particularly from the 50 to 150 count noun mark in productive vocabulary. Of further interest are children's strategies for object recognition in the earliest stages of vocabulary growth. Accordingly, for the main analyses, children were placed into three developmental groups: (1) Group I – those at the earliest stages of word learning, with fewer than 50

count nouns in productive vocabulary; (2) Group II – those whose productive vocabularies fall in this suggested transition period, 50 to 150 count nouns, and (3) Group III – those with more extensive productive vocabularies, greater than 150 count nouns. Table 1 shows the ages and numbers of count nouns in productive vocabulary for the three groups of children. As is common during this period of rapid development, there is a wide range of overlapping ages for each level of productive vocabulary development, although vocabulary size and age are also correlated, $r = .59$, $p < .001$.

Figure 5 shows the main results. The children with the most advanced count noun vocabularies recognize all three kinds of stimuli equally well. The middle group of children, with vocabularies between 50 and 150 count nouns, recognize the Typical instances that provide shape, texture, color, and fine-detail information as well as they recognize the Rich Shape instances that provide only shape information but do so at a high degree of local detail. These children, however, recognize the Shape Caricatures less well than the two more detailed kinds of instances. The children with the smallest count noun vocabularies overall comprehend fewer of the nouns – in all conditions – than do children in the two more

advanced groups, a not surprising result given the size of their productive vocabularies. But, critically, these children show the same pattern of recognition as do the middle group of children, recognizing Typical and Rich Shape instances equally well and better than the Shape Caricatures.

These conclusions were confirmed by an ANOVA for a 3 (Vocabulary level) × 3 (Stimulus condition) mixed design which yielded significant main effects of Vocabulary level, $F(2, 61) = 10.09$, $p < .001$, and Stimulus condition, $F(2, 122) = 24.97$, $p < .001$, and a reliable interaction between these two factors, $F(4, 122) = 4.10$, $p < .01$. Post-hoc comparisons (Tukey's HSD, $\alpha = .05$) also confirm the following pairwise comparisons: Children in Groups I and II perform less well given the Shape Caricatures than do the children in Group III given these same stimuli. The children in Groups I and II also perform less well on the Shape Caricatures than they do on the Rich Shape and Typical instances, which do not differ from each other. Finally, children in Group III perform equally well on all stimulus types. Performance of all children in all stimulus conditions exceeded chance except the least advanced children given Shape Caricatures, $t(19) = 1.98$, $p > .05$. These effects are not due to the particular definition of vocabulary groups. In a second analysis, we fitted the data to an ANCOVA model that included the productive vocabulary measure as a covariate. We found significant main effects of Stimulus condition, $F(2, 124) = 21.78$, $p < .001$, and number of count nouns in productive vocabulary, $F(1, 62) = 21.21$, $p < .001$, and a significant two-way interaction between these two variables, $F(2, 124) = 4.53$, $p < .02$, indicating again that less and more advanced children in productive vocabulary differ in their recognition of shape caricatures, and not in recognition of detailed instances.

The observed relation between the recognition of shape caricatures and vocabulary size in these analyses could indicate a causal relation between processes of visual object recognition and lexical development or perhaps both are related to some other factor. We will consider this issue in the general discussion. Here we note only that, in this sample, recognition of shape caricatures is more strongly correlated with count noun vocabulary, $r = .58$ and $R^2 = .34$, than with age, $r = .38$, $R^2 = .15$. Subsequent analyses also examined the possibility of individual stimulus effects. None were observed. For example, children were equally likely to recognize shape caricatures of animals as of nonanimals (58% and 52% correct, respectively) and to recognize shape caricatures with more than four parts and less than four parts (59% and 53% correct).

The main finding then is that between 18 and 24 months, children are well able to recognize objects from highly detailed and prototypical information about object shape alone, but there is a marked increase in the ability to recognize objects from abstract representations of global geometric shape. The developmental trend is not strictly about being able to use only shape information to recognize an object (rather than, for example, also requiring color
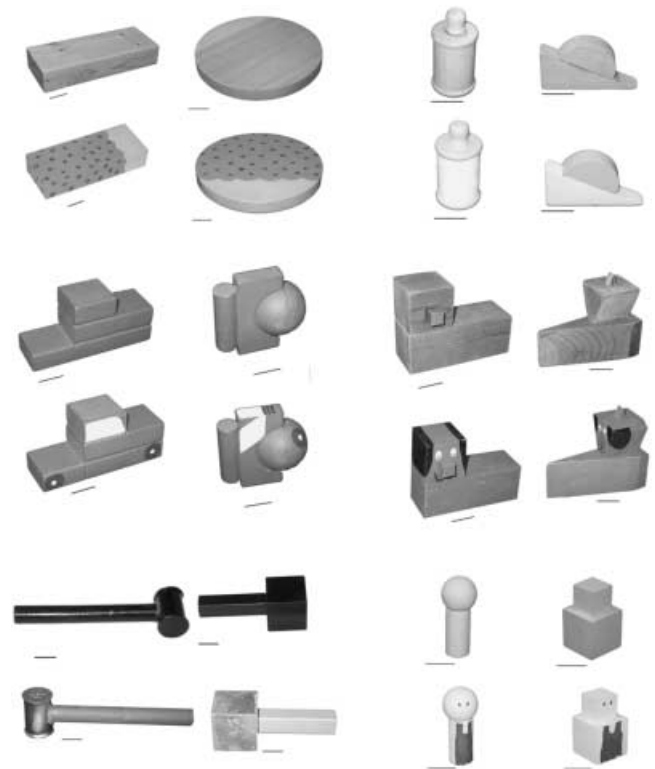
Figure 6 *Photos of the stimuli in Experiment 2. Each set of four pictures contains clockwise from the upper left: (–Local Details, +Shape Caricature), (–Local Details, –Shape Caricature), (+Local Details, –Shape Caricature), (+Local Details, +Shape Caricature). The black line close to each object is 1 inch in length.*

or texture information) but rather is about abstract representation of global shape. These developments occur during a period of rapid growth in children's knowledge of object names and thus may play a role in supporting that growth, be a consequence of that learning, or both.

Experiment 2

Children's recognition of the Rich Shape objects could be based on an overall prototype of the whole or it could be based on fragments and localized clusters of features. To examine this issue, and children's possibly joint use of local and global information, Experiment 2 consisted of a 2 × 2 design examining the presence and absence of global information about geometric structure (which we will label by +*Shape Caricature* and –*Shape Caricature*) and localized and fine detailed information predictive of the category (which we will label by +*Local Details* and –*Local Details*). Examples of the four stimulus conditions are in Figure 6. More specifically, the +Shape Caricatures, structured as in Experiment 1, were made from one to four geometric components in the proper spatial relations. These representations are sparser than those in Experiment 1 in order to increase sensitivity to the possible

Table 2 *Number of subjects, mean value and range for age and MCDI's number of object names for Group I, II and III in Experiment 2*

	Group I – Children knowing less than 51 object names (*n* = 21)		Group II – Children knowing between 51 and 150 object names (*n* = 27)		Group III – Children knowing more than 150 object names (*n* = 44)	
	M	Range	*M*	Range	*M*	Range
Developmental indicator						
Age (months)	20.0	16–29	21.7	17–31	25.9	17–31
Number of object names in productive vocabulary	20.0	5–48	103.9	54–150	247.4	153–373

contributions of local details. The −Shape Caricatures were alterations of the +Shape Caricatures: the shape of at least one component volume was altered and if possible the spatial arrangement of two volumes relative to each other was rearranged. The presence of detailed local information was achieved by painting surface details on these volumes that were predictive of the target category, for example, the face of a dog, wheels, and so forth.

Method

Participants

Ninety-two children (44 female, 48 male) were recruited from a working- and middle-class population in a Midwestern college town. All were native speakers of English and had no known neurological or language disorders. They ranged in age from 16.0 months to 31.0 months. Fifteen additional subjects began the experiment but did not contribute data because the parent coached the child (contrary to instructions) during the experiment, for fussiness, for failure to understand the task, or experimenter error.

Stimuli

Instances of six common categories were selected: dog, truck, person, hammer, bed and bottle. For every category, a +Shape Caricature instance was made from one to four geometric volumes to represent the overall shape. The −Shape Caricature instance was made by changing the shape of at least one component and (if possible) rearranging the spatial structure as shown in Figure 6. The +Local Details instances were made by painting localized surface features predictive of the target category. These were positioned on the most appropriate location. Each stimulus object was approximately 75 cm^3.

Language measure

As in Experiment 1, parents were asked to indicate the number of count nouns in their children's productive vocabulary using the Bates-MacArthur Communicative Development Inventory (MCDI) (Fenson *et al.*, 1994).

Procedure and design

The + and −Local Details conditions were tested between subjects and the + and −Shape Caricature conditions were within subjects. Within each between-subject condition, three categories were assigned to be targets in the +Shape Caricature condition and three were assigned to be targets in the −Shape Caricature condition. Category assignments to the +/− Shape Caricatures were counterbalanced across children such that each object served equally often across children in the + or − version. Each object both in its + and its −Shape Caricature version also served as distracters. Targets and distracters on every trial were from the same stimulus condition. Children were questioned about each unique target twice with different randomly selected distracters serving on each trial. Thus there were six +Shape Caricature trials (three target categories repeated twice) and six −Shape Caricature trials (three target categories repeated twice) or a total of 12 trials in each between-subject (+/− Local Details) condition. These 12 trials were presented in a randomly determined order with the constraint that two successive trials did not have identical targets. All other aspects of the procedure were the same as in Experiment 1.

Results

Table 2 shows the ages and numbers of count nouns in productive vocabulary for the three groups of children. Again and as is common during this period of rapid development, there is a wide range of overlapping ages for each level of productive vocabulary development, although vocabulary size and age are also correlated (r = .61, p < .001).

Figure 7 shows the main result; the darker bars indicate performance when local details were present (+Local Details) and the solid bars indicate performance given the +Shape Caricatures, that is, when the appropriate though sparse global shape structure was present. The children in the lowest vocabulary group show their highest level of performance (the darker bars) when the stimuli present local details, and for these stimuli the presence or absence of appropriate shape structure does not matter. The children in the most advanced vocabulary group show their best performance given the appropriate sparse representations of global shape. These results suggest
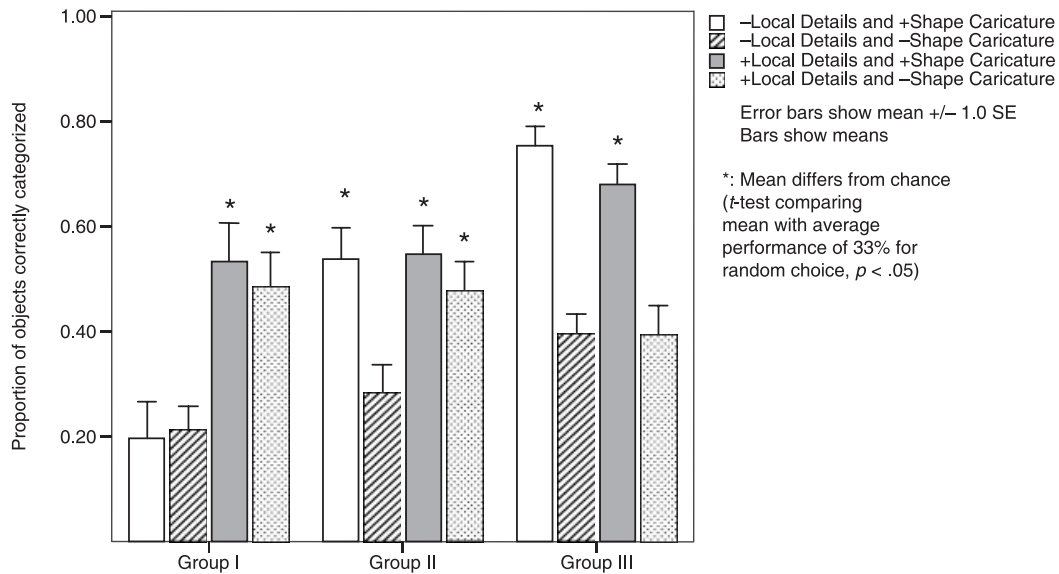
Figure 7 *Mean proportion of number of objects correctly categorized (out of six trials) in Experiment 2 across the three groups of vocabulary level for the Local Details × Shape Caricature interaction. Between-subjects conditions −Local Details and +Local Details are shown in white and black bars, respectively. Within-subjects conditions +Shape Caricature and −Shape Caricature are shown in solid and patterned bars, respectively.*

increasing recognition of the shape caricatures with increasing vocabulary size and a greater dependence on local features earlier in their vocabulary development.

The conclusions were confirmed by a 2 (+/− Local Details) by 3 (Vocabulary) by 2 (+/− Shape Caricatures) ANOVA mixed design. The analysis yielded a reliable main effect of Shape Caricatures, $F(1, 86) = 30.51, p < .001$. Children chose the target instance more often when it was composed of appropriate volumes in the proper arrangement than when it was not. The analysis also yielded a reliable main effect of Local Details, $F(1, 86) = 12.82, p < .001$; the presence of category-appropriate features led to improved recognition. There was also a main effect of Vocabulary group, $F(2, 86) = 11.80, p < .001$; children with less advanced vocabularies recognized fewer objects than the children with more advanced vocabularies. More critically, the analysis also yielded a reliable interaction between Shape Caricature and Vocabulary Group, $F(2, 86) = 8.96, p < .001$, and between Local Details and Vocabulary Group, $F(2, 86) = 8.51, p < .001$. The three-way interaction between Shape Caricature, Local Details and Vocabulary Group was not significant, $F(2, 86) = 1.19, p > 0.3$.

Post-hoc analyses (Tukey's HSD, $\alpha < .05$) indicate that children with smaller and larger noun vocabularies recognized the targets equally well when they had localized features predictive of the category (for both +/− Shape Caricatures) but children with the most advanced vocabularies recognized targets better than the two other vocabulary groups (which did not differ) when the targets were shape caricatures (without local details) and thus could be recognized only via the global geometric structure. There were no reliable differences among the groups when the targets presented neither the predictive local features

nor global structure. Figure 7 also provides the results of individual means compared to chance. It is noteworthy that the children with the least advanced vocabularies performed above chance only when Local Details were present whereas the most advanced children performed above chance when the stimuli had the correct global shape.

The performance of individual children as a function of vocabulary level is shown in Figure 8. In order to investigate the effect of using categorical levels in the productive vocabulary measure, we submitted these data to a mixed design ANCOVA model that included this measure as a covariate. This analysis revealed the same results as in the initial ANOVA model: significant main effects of Shape Caricature, $F(1, 88) = 52.26, p < .001$, Local Details, $F(1, 88) = 8.13, p < .01$, and number of count nouns in productive vocabulary, $F(1, 88) = 22.04, p < .001$; significant two-way interaction between Shape Caricature and number of count nouns in productive vocabulary, $F(1, 88) = 25.44, p < .001$, and between Local Details and number of count nouns in productive vocabulary, $F(1, 88) = 8.13, p < .01$. Figure 8 also illustrates the robustness of these effects by showing the sample in terms of productive vocabulary. The children with the smallest vocabularies tend to perform better when the local details are added; this can be seen by the higher density of the cross symbol on the top part of the scatterplot as compared to the higher density of the circle symbol in the bottom part. In contrast, children with the highest vocabularies perform best when the global shape information is added and this can be seen by the higher density of dark symbols on the top part of the scatterplot. There is a strong correlation between vocabulary and recognition of the +Shape Caricature when there are no Local Details ($r = .75, p < .001$), and
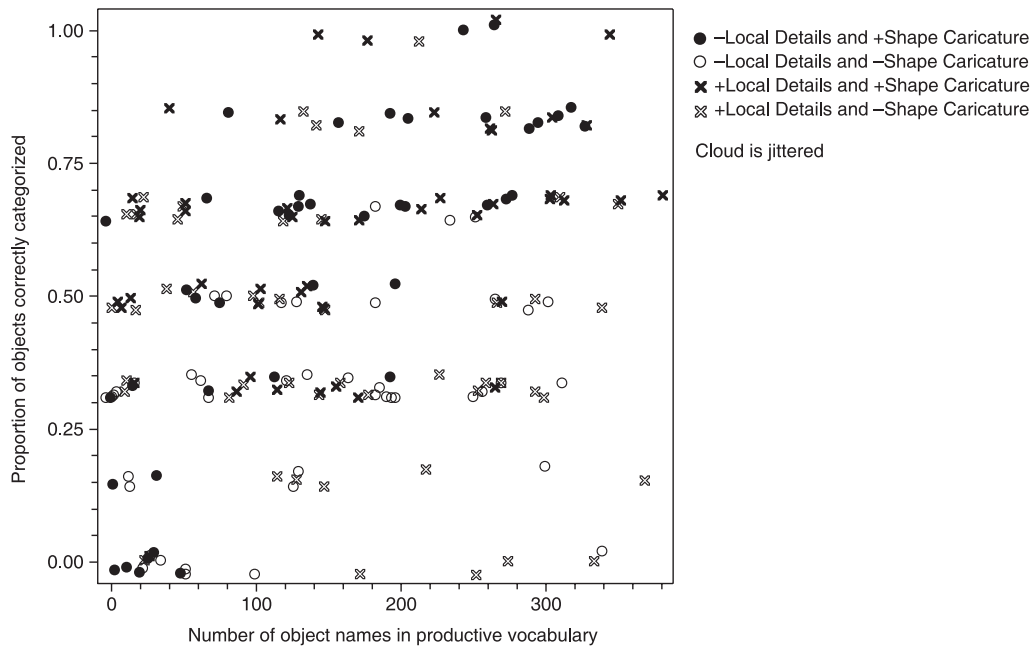
Figure 8 *Scatterplot of the individual subject's proportion of number of objects correctly categorized (out of six trials) in Experiment 2 by number of object names in the child's productive vocabulary by parent report on the MCDI. To ease visualization, a random jitter was added to each point.*

also when there are Local Details ($r = .41$, $p < .01$). The recognition of targets presenting *only* local features (−Shape Caricatures +Local features) is negatively, though not reliably, related to vocabulary level ($r = −.17$, $p > .20$). This pattern of individual results, like the group results, strongly suggests that it is not just object recognition in general that is increasing, but object recognition based on global geometric structure.

Because vocabulary and age are themselves strongly correlated, one cannot be certain that the developmental effects in object recognition reported above reflect vocabulary-specific effects. However, recognition of the shape caricatures with no local features is more strongly related to vocabulary level ($r = .75$, $p < .001$) than to age ($r = .58$, $p < .001$). However, recognition of shape caricatures with local features is equally predicted by vocabulary ($r = .41$, $p < .005$) and age ($r = .40$, $p < .005$). Thus, category knowledge as measured by productive vocabulary appears to be specifically related to the formation of sparse representations of global shape.

Finally, these conclusions appear to be appropriate for the individual six categories included in the study. In each condition, children were categorized as recognizing an item if they chose it at least once (out of two trials) when it was the target. By this measure, there were no reliable item effects in any of the four conditions, $\chi^2(5, N = 92) = 5.56$, $p > .35$.

General discussion

Research in machine vision, a field that tries to build devices to recognize objects, makes clear that object

recognition is not trivial (Peissig & Tarr, 2007). The specific computational goal in most approaches to machine vision is to find the proper internal representation that is sensitive to relevant differences for object discrimination but tolerant of irrelevant variation within a class. A variety of different kinds of representations have been proposed, for example, holistic prototypes, fragments, low-level features, and relations among geometric primitives (Biederman, 1987; Edelman & Intrator, 2003; Hummel & Biederman, 1992; Tarr & Bülthoff, 1998; Ullman, 1996; Ullman *et al.*, 2002). Adult humans are experts in recognizing common objects – from partial and occluded views, from various perspectives – and more dauntingly, they can readily recognize novel instances for many different categories (Standing, 1973; Thorpe, Fize & Marlot, 1996). Importantly, this expertise may arise not from the use of a single representation, but instead may include variations of *all* the approaches in machine vision (Riesenhuber & Poggio, 2000; Ullman, 1996).

Despite the importance of object recognition for all areas of learning, categorization, and cognition, remarkably little is known about the development of object recognition. This is so even though many computational models of object recognition explicitly involve two stages: the *acquisition* stage during which labeled representations are constructed from explicit training experiences and a subsequent *recognition* stage which uses these representations to recognize novel instances and to form novel categories. A full theory of human object recognition requires a description of this acquisition phase. Moreover, a developmental description may be needed for several different types of representations.

The present two experiments contribute first by adding to the developmental studies suggesting potentially

significant changes in visual object recognition in the second year of life (Smith, 2003; Son *et al.*, under review). The very idea that object recognition may change substantially in this period is not one commonly considered in studies of categorization and concepts in infancy and early childhood. This is so even though we know that there is at last one domain in which recognition undergoes significant changes as a function of development and experience. Specifically, face recognition is characterized by strong early sensitivities in infancy yet also shows a slow and protracted course of development with adult-like expertise not achieved until adolescence (e.g. Johnson, Dziurawiec, Ellis & Morton, 1991; Mondloch, Lewis, Budreau, Maurer, Dannemiller, Stephens & Kleiner-Gathercoal, 1999; Carey, Diamond & Woods, 1980; Mondloch, Le Grand & Maurer, 2002). In this context, the idea of significant changes in object recognition and a possibly protracted course of development seem less surprising (as also suggested by Abecassis, Sera, Yonas & Schwade, 2001). The present results specifically indicate early significant changes in the same period during which children learn the names for many common categories. Further, they suggest that early object recognition may be piecemeal and based on fragments and parts (see also Rakison & Lupyan, in press). During this developmental period, children appear to *add* to this earlier recognition process one based on sparse representation of global geometric structure. These findings raise several issues relevant to a complete understanding of the development of object recognition and the role of category learning in that development. We discuss these next.

Why is early recognition by local features or fragments? Research in a number of domains suggests that perceptual learning and development progress generally from being based on parts to being based on wholes (Mareschal, 2000). This is certainly evident in the case of face perception, which proceeds from piecemeal to configural (e.g. Carey & Diamond, 1977; Maurer, Le Grand & Mondloch, 2002). This has also been suggested by studies of adult experts (Bukach, Gauthier & Tarr, 2006; Gauthier & Tarr, 1997, 2002; Tanaka & Taylor, 1991). The form of 'expert' representations relevant to identifying faces, or bird species, or car makes is likely to be different from that relevant to recognizing common objects since those forms of experiences require fine-grained discriminations whereas common object recognition requires the treatment of a broad array of instances as equivalent (Biederman & Kalocsai, 1997; Nelson, 2001). Yet still, the starting point for learning in all these domains may be local parts or fragments.

Interestingly, as Mareschal (2000) notes, this developmental trend from more piecemeal to more integrated representations has been observed at varying ages in different kinds of tasks. For example, Younger and Furrer (2003) showed a progression from categorization based on features to integrated holistic form in 6- to 12-month-old infants in a habituation task using line drawings. Rakison and Cohen (1999) report a similar trend in 12- to 24-month-olds in their categorization of three-dimensional instances of common categories (e.g. animals and vehicles; see also Mareschal, Quinn & French, 2002; Quinn & Eimas, 1996; Rakison & Butterworth, 1998b; Smith, Jones & Landau, 1996; Younger & Furrer, 2003). Since this trend has also been suggested in comparisons of adult novices and experts, it may reflect a general pattern of experience with object categories. One testable prediction is that for categories with which children have more experience, there should be greater sensitivity to holistic shape, a result found in Quinn (2004b).

Presumably, the visual system develops the kinds of representations that support the task that needs to be done (Biederman & Kalocsai, 1997; Nelson, 2001). The task in object recognition and categorization requires the treatment of many different things with unique detailed shapes (all variety of chairs, for example) as equivalent. It has been argued (Biederman, 2000; Biederman & Gerhardstein, 1993) that for this task the most reliable type of object information that is available from an image is *abstract geometric* information. Consistent with this idea, Son *et al.* (under review) recently showed that young children learned and generalized novel categories better when they were presented with training instances that highlighted the abstract geometric structure of the whole. In that study, children who were given the abstractions first, benefited later in category learning. In real world development, children presumably have to build these abstractions on their own from the experience of individual and richly unique instances.

Two developmental studies suggest that the developmental course in object recognition, like that in face recognition, is slow and long. One study by Abecassis *et al.* (2001) investigated 2- to 4-year-olds' categorization of volumes that varied on categorical and metric properties. This is a relevant question because Biederman has suggested that primitive volumes are categorical, and that metric variability within a category does not matter for visual recognition of object part structure. Abecassis *et al.* (2001) found that young children were highly sensitive to metric variations within categories of volumes, and that their responses were at best weakly organized by Biederman's categorical distinctions. This suggests that if Biederman's account is an accurate description of adult visual object recognition, it is not a good description of young children's object representations. Instead, the relevant parts in children's perceptions may well be richer fragments, which include metric properties and local details.

In a related study, Mash (2006) presented two-dimensional images of novel objects to children and adults that differed in both part shapes and part positions in relation to each other. Mash varied part shape and part position metrically and used a triad task to ask children to match a category exemplar. Both kinds of differences were discriminated by 5- and 8-year-olds, but the younger age group attended only to matching parts and not to relational structure. This result suggests that the relational structure may be particularly late, an idea that fits with

other general principles of development (e.g. Gentner & Rattermann, 1991). Importantly, both the Abecassis *et al.* (2001) and the Mash (2006) studies concerned children's categorizations of *novel* things not real common object categories as examined here. It may well be that children first learn the relevant kinds of object representations to support recognition for specific well-learned categories and only later generalize those principles to novel things. It is also possible that children's emerging recognition of shape caricatures documented here is not based on abstractions of the kind described by Biederman's theory of geons but may instead be based on components lying somewhere between highly detailed fragments and categorical volumes. This is a developmental question requiring programmatic empirical investigation.

In this context, an important question is the role of category learning and, specifically, object names in these developments. There is circumstantial evidence for a causal role for word learning in that, in the present study, the recognition of shape caricatures is strongly predicted by the number of count nouns – the nouns that refer to the shape-based categories in children's vocabularies (Samuelson & Smith, 1999). Further, Jones and Smith (2005) found that late talkers who had small productive vocabularies for their age also had difficulty in recognizing shape caricatures and performed in that task more on a par with language-matched children than age-matched children. The case, however, is circumstantial because the evidence is merely correlational. It could well be that the recognition of shape caricatures emerges through some independent set of processes but then is related to word learning because attention to abstract geometric shape makes the learning and generalization of noun categories easier. Still, an important possibility is that learning the range of instances that fall into natural object categories educates processes of visual object recognition and may well promote highly abstract representations of global structure that, while not necessary to object recognition (after all, all the children recognize the richly detailed instances), make it more robust in certain task circumstances, including generalization to new instances. Although category learning itself seems of likely relevance, it is an open question as to whether word learning *per se* is critical or whether other forms of category learning through functional uses or action make critical contributions (Smith, 2005).

A final question is the relation between children's recognition of shape caricatures and the so-called shape bias in children's novel noun generalizations. Past research using objects of highly simplified shape suggests that children's attention to shape in the laboratory task of generalizing names for novel things is strongly related to the number of count nouns in children's productive vocabulary. Further, teaching very young children (17-month-olds) to attend to shape when naming these simple things induces a generalizable shape bias and increases the rate of count noun learning outside of the laboratory (Smith *et al.*, 2002). These results suggest a developmental feedback loop between learning object names and attention to shape: Learning names provides a context in which children learn the relevance of shape for object categories. Each name learned enhances attention to shape, progressively creating a generalized bias to extend names to new instances by shape, and as a consequence yields more rapid learning of common noun categories. The additional question, and one not answered by the present results, is whether such experiences also create a more abstract and minimalist description of object shape, one that may in fact be necessary for a shape bias to be useable at all in categorizing the richly detailed things that populate the real world. Thus, an important next question in this program of research is whether the recognition of shape caricatures developmentally precedes a systematic bias to extend object names to new instances by shape or, perhaps, is somehow dependent on increased attention to shape over such other object properties as color, texture, material, and size.

In conclusion, the results from these two experiments provide converging evidence for significant changes in visual object recognition during the developmental period in which children's object name learning is rapidly expanding. They demonstrate the multiple sources of information that may be used to recognize objects and show that young children, at the start of a period of rapid category learning, can use detailed local information to recognize instances of common categories but not more abstract information about geometric structure. Children only slightly more advanced, however, do recognize common objects from such shape caricatures. This period of rapid developmental change seems crucial to understanding the nature of human object recognition and may also provide a crucial missing link in our understanding of the developmental trajectory in early object name learning, a trajectory of vocabulary growth that begins slowly but progresses to quite rapid learning characterized by the fast-mapping of names to whole categories of similarly shaped things.

Acknowledgements

References

Abecassis, M., Sera, M.D., Yonas, A., & Schwade, J. (2001). What's in a shape? Children represent shape variability differently than adults when naming objects. *Journal of Experimental Child Psychology*, **78** (3), 213–239.

Arguin, M., & Saumier, D. (2004). Independent processing of parts and of their spatial organization in complex visual objects. *Psychological Science*, **15** (9), 629–633.

Bates, E., Dale, P.S., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 96–151). Oxford: Blackwell.

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review*, **94**, 115–117.

Biederman, I. (2000). Recognizing depth-rotated objects: a review of recent research and theory. *Spatial Vision, Special Issue: Object recognition and image understanding by brain and machines*, **13** (2–3), 241–253.

Biederman, I., & Gerhardstein, P.C. (1993). Recognizing depth-rotated objects: evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, **19** (6), 1162–1182.

Biederman, I., & Kalocsai, P. (1997). Neurocomputational bases of object and face recognition. *Philosophical Transactions of the Royal Society London, B Biological Sciences*, **352** (1358), 1203–1219.

Bukach, C.M., Gauthier, I., & Tarr, M.J. (2006). Beyond faces and modularity: the power of an expertise framework. *Trends in Cognitive Sciences*, **10** (4), 159–166.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science*, **195** (4275), 312–314.

Carey, S., Diamond, R., & Woods, B. (1980). Development of face recognition: a maturational component? *Developmental Psychology*, **16** (4), 257–269.

Cerella, J. (1986). Pigeons and perceptrons. *Pattern Recognition*, **19** (6), 431–438.

Colunga, E. (2003, April). Where is the cow hiding? A new method for studying the development of features. Paper presented at the Biennial meeting of the Society for Research on Child Development, Tampa, FL.

Colunga, E., & Smith, L.B. (2005). From the lexicon to expectations about kinds: a role for associative learning. *Psychological Review*, **112** (2), 347–382.

Cooper, E.E., Biederman, I., & Hummel, J.E. (1992). Metric invariance in object recognition: a review and further evidence. *Canadian Journal of Psychology*, **46** (2), 191–214.

Edelman, S. (1995). Representation, similarity, and the chorus of prototypes. *Minds and Machines*, **5** (1), 45–68.

Edelman, S. (1999). *Representation and recognition in vision*. Cambridge, MA: MIT Press.

Edelman, S., & Bülthoff, H.H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Research*, **32** (12), 2385–2400.

Edelman, S., & Duvdevani-Bar, S. (1997). Similarity, connectionism, and the problem of representation in vision. *Neural Computation*, **9** (4), 701–720.

Edelman, S., & Intrator, N. (1997). Learning as extraction of low-dimensional representations. In R.L. Goldstone, D.L. Medin, & P.G. Schyns (Eds.), *Perceptual learning: The psychology of learning and motivation, Vol. 36* (pp. 353–380). San Diego, CA: Academic Press.

Edelman, S., & Intrator, N. (2003). Towards structural systematicity in distributed, statically bound representations. *Cognitive Science*, **27** (1), 73–109.

Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., Thal, D.J., & Pethick, S.J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, **59** (5, serial 242).

Fize, D., Fabre-Thorpe, M., Richard, G., Doyon, B., & Thorpe, S.J. (2005). Rapid categorization of foveal and extrafoveal natural images: associated ERPs and effects of lateralization. *Brain and Cognition*, **59** (2), 145–158.

Gathercole, V.C.M., & Min, H. (1997). Word meaning biases or language-specific effects? Evidence from English, Spanish and Korean. *First Language*, **17** (49, Pt 1), 31–56.

Gauthier, I., & Tarr, M.J. (1997). Becoming a 'greeble' expert: exploring mechanisms for face recognition. *Vision Research*, **37** (12), 1673–1682.

Gauthier, I., & Tarr, M.J. (2002). Unraveling mechanisms for expert object recognition: bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance*, **28** (2), 431–446.

Gentner, D., & Rattermann, M.J. (1991). Language and the career of similarity. In Susan A. Gelman & James P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 1225–1277). New York: Cambridge University Press.

Gershkoff-Stowe, L., & Smith, L.B. (2004). Shape and the first hundred nouns. *Child Development*, **7** (10), 1098–1114.

Hayward, W.G. (2003). After the viewpoint debate: where next in object recognition? *Trends in Cognitive Sciences*, **7** (10), 425–427.

Hummel, J.E. (2000). Where view-based theories break down: the role of structure in shape perception and object recognition. In E. Dietrich & A.B. Markman (Eds.), *Cognitive dynamics: Conceptual and representational change in humans and machines* (pp. 157–185). Hillsdale, NJ: Erlbaum.

Hummel, J.E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, **99** (3), 480–517.

Imai, M. (1999). Constraint on word-learning constraints. *Japanese Psychological Research, Special Issue: Lexical development: Evaluation of word-learning constraints*, **41** (1), 5–20.

Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning: the role of shape similarity in early acquisition. *Cognitive Development*, **9** (1), 45–75.

Johnson, M.H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, **40** (1–2), 1–19.

Jones, S.S., & Smith, L.B. (2005). Object name learning and object perception: a deficit in late talkers. *Journal of Child Language*, **32** (1), 223–240.

Keil, F.C. (1994). The birth and nurturance of concepts by domains: the origins of concepts of living things. In Lawrence A. Hirschfeld & Susan A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 1234–1254). New York: Cambridge University Press.

Kellman, P.J. (2001). Separating processes in object perception. *Journal of Experimental Child Psychology, Special Issue: Reflections*, **78** (1), 84–97.

Kimchi, R., Hadad, B., Behrmann, M., & Palmer, S.E. (2005). Microgenesis and ontogenesis of perceptual organization: evidence from global and local processing of hierarchical patterns. *Psychological Science*, **16** (4), 282–290.

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

Mareschal, D. (2000). Object knowledge in infancy: current controversies and approaches. *Trends in Cognitive Sciences*, **4** (11), 408–416.

Mareschal, D., Quinn, P.C., & French, R.M. (2002). Asymmetric interference in 3- to 4-month-olds' sequential category learning. *Cognitive Science: A Multidisciplinary Journal*, **26** (3), 377–389.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* New York: W.H. Freeman.

Mash, C. (2006). Multidimensional shape similarity in the development of visual object classification. *Journal of Experimental Child Psychology*, **95** (2), 128–152.

Maurer, D., Le Grand, R., & Mondloch, C.J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, **6** (6), 255–260.

Mondloch, C.J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, **31** (5), 553–566.

Mondloch, C.J., Lewis, T.L., Budreau, D.R., Maurer, D., Dannemiller, J.L., Stephens, B.R., & Kleiner-Gathercoal, K.A. (1999). Face perception during early infancy. *Psychological Science*, **10** (5), 419–422.

Nelson, C.A. (2001). The development and neural bases of face recognition. *Infant and Child Development, Special Issue: Face processing in infancy and early childhood*, **10** (1–2), 3–18.

Pegna, A.J., Khateb, A., Michel, C.M., & Landis, T. (2004). Visual recognition of faces, objects, and words using degraded stimuli: where and when it occurs. *Human Brain Mapping*, **22** (4), 300–311.

Peissig, J.J., & Tarr, M.J. (2007). Visual object recognition: do we know more now than we did 20 years ago? *Annual Review of Psychology*, **58** (1), 75–96.

Peterson, M. (Ed.) (1999). *The MIT encyclopedia of the cognitive sciences.* Cambridge, MA: MIT Press.

Quinn, P.C. (2004a). Development of subordinate-level categorization in 3- to 7-month-old infants. *Child Development*, **75** (3), 886–899.

Quinn, P.C. (2004b). Is the asymmetry in young infants' categorization of humans versus nonhuman animals based on head, body, or global gestalt information? *Psychonomic Bulletin & Review*, **11** (1), 92–97.

Quinn, P.C., & Eimas, P.D. (1996). Perceptual cues that permit categorical differentiation of animal species by infants. *Journal of Experimental Child Psychology*, **63** (1), 189–211.

Rakison, D.H., & Butterworth, G.E. (1998a). Infants' attention to object structure in early categorization. *Developmental Psychology*, **34** (6), 1310–1325.

Rakison, D.H., & Butterworth, G.E. (1998b). Infants' use of object parts in early categorization. *Developmental Psychology*, **34** (1), 49–62.

Rakison, D.H., & Cohen, L.B. (1999). Infants' use of functional parts in basic-like categorization. *Developmental Science*, **2** (4), 423–431.

Rakison, D.H., & Lupyan, G. (in press). Developing object concepts in infancy: an associative learning perspective. *Monographs of the Society for Research in Child Development*.

Riesenhuber, M., & Poggio, T. (2000). The individual is nothing, the class everything: psychophysics and modeling of recognition in object classes. AI Memo 1682, CBCL Paper 185, Artificial Intelligence Lab and Center for Biological & Computational Learning, Massachusetts Institute of Technology.

Rosch, E.H. (1973). Natural categories. *Cognitive Psychology*, **4** (3), 328–350.

Saiki, J., & Hummel, J.E. (1998). Connectedness and the integration of parts with relations in shape perception. *Journal of Experimental Psychology: Human Perception and Performance*, **24** (1), 227–251.

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

Smith, L.B. (2003). Learning to recognize objects. *Psychological Science*, **14** (3), 244–250.

Smith, L.B. (2005). Action alters shape categories. *Cognitive Science*, **29** (4), 665–679.

Smith, L.B., Jones, S.S., & Landau, B. (1996). Naming in young children: a dumb attentional mechanism? *Cognition*, **60** (2), 143–171.

Smith, L.B., Jones, S.S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, **13** (1), 13–19.

Soja, N.N. (1992). Inferences about the meanings of nouns: the relationship between perception and syntax. *Cognitive Development*, **7** (1), 29–45.

Son, J.Y., Smith, L.B., & Goldstone, R.L. (under review). Simplicity and generalization: the role of abstraction in children's object categorizations. *Cognition*.

Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology*, **25** (2), 207–222.

Stankiewicz, B.J. (2003). Just another view. *Trends in Cognitive Sciences*, **7** (12), 526.

Tanaka, J.W., & Taylor, M. (1991). Object categories and expertise: is the basic level in the eye of the beholder? *Cognitive Psychology*, **23** (3), 457–482.

Tarr, M., & Vuong, Q.C. (2002). Visual object recognition. In S. Yantis (Ed.), *Steven's handbook of experimental psychology: Vol. 1. Sensation and perception* (3rd edn., Vol. 1, pp. 287–314). New York: John Wiley & Sons.

Tarr, M.J., & Bülthoff, H.H. (1998). Image-based object recognition in man, monkey, and machine. *Cognition, Special Issue: Image-based recognition in man, monkey and machine*, **67**, 1–20.

Tarr, M.J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology*, **21** (2), 233–282.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, **381** (6582), 520–522.

Tomasello, M.M. (1994). The instrument is great, but measuring comprehension is still a problem. *Monographs of the Society for Research in Child Development*, **59** (5), 174–179.

Ullman, S. (1996). *High-level vision: object recognition and visual cognition.* Cambridge, MA: MIT Press.

Ullman, S., & Bart, E. (2004). Recognition invariance obtained by extended and invariant features. *Neural Networks*, **17** (5/6), 833–848.

Ullman, S., Vidal-Naquet, M., & Sali, E. (2002). Visual features of intermediate complexity and their use on classification. *Nature Neuroscience*, **5** (7), 682–687.

Yoshida, H., & Smith, L.B. (2003). Known and novel noun extensions: attention at two levels of abstraction. *Child Development*, **74** (2), 564–577.

Younger, B.A., & Furrer, S.D. (2003). A comparison of visual familiarization and object-examining measures of categorization in 9-month-old infants. *Infancy*, **4** (3), 327–348.