ELSEVIER

Cognition 85 (2002) 223–250

COGNITION

www.elsevier.com/locate/cognit

The role of language in acquiring object kind concepts in infancy

Fei Xu

Department of Psychology, 125 Nightingale Hall, Northeastern University, Boston, MA 02115, USA

Received 25 April 2001; received in revised form 18 January 2002; accepted 29 May 2002

Abstract

Four experiments investigated whether 9-month-old infants could use the presence of labels to help them establish a representation of two distinct objects in a complex object individuation task. We found that the presence of two distinct labels facilitated object individuation, but the presence of one label for both objects, two distinct tones, two distinct sounds, or two distinct emotional expressions did not. These findings suggest that language may play an important role in the acquisition of sortal/object kind concepts in infancy: words may serve as "essence placeholders". Implications for the relationship between language and conceptual development are discussed. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Language; Object kind concepts; Infancy

1. Introduction

What is the relationship between language and thought? This question has intrigued psychologists and philosophers for decades (Brown, 1976; Carey, 1994; Gopnik & Meltzoff, 1997; Nelson, 1996; Quine, 1960; Whorf, 1956; among others; see Bloom & Keil, 2001, for a review). Contemporary discussions of the issue often begin with Whorf and Sapir, who advocated a radical version of linguistic determinism: the linguistic distinctions made in each language determine the conceptual distinctions made by speakers of that language. This strong Whorfian view has long been abandoned on both theoretical and empirical grounds. Many have argued that the position is conceptually incoherent (Fodor, 1975; Pinker, 1994) and empirical studies have shown that speakers of languages that mark different linguistic distinctions can nevertheless make the same conceptual distinctions. For example, color perception in Dani is fundamentally similar to ours even though Dani language only has two color terms (Heider, 1972; Rosch, 1975, 1978; see Brown,

E-mail address: fxu@neu.edu (F. Xu).

PII: S0010-0277(02)00109-9

1976, for a review). A recent surge of interest in these issues has focused on developing weaker versions of the linguistic relativity view. Most investigators agree that studies may reveal differences in *habitual* thought across languages and that some of these linguistic differences may influence memory, categorization, and other aspects of non-linguistic cognition (Boroditsky, 2001; Gentner & Boroditsky, 2001; Gumperz & Levinson, 1996; Hunt & Agnoli, 1991; Levinson, 1996; Lucy, 1992; Pederson et al., 1998; Slobin, 1996; though see Li & Gleitman, 2002, for a critique of some of the findings).

Parallel to the discussion on how cross-linguistic differences may shape cognition in adults, both Vygotsky (1962) and Quine (1960) have proposed theories of conceptual development in which language plays a crucial role. Vygotsky emphasized how universal aspects of language may exert influence on cognitive development, thus separating us from alinguistic non-human primates and pre-linguistic human infants. Quine, on the other hand, considered how language may be used to build our ontology. Most widely discussed is the case of the conceptual distinction between objects and substances and the linguistic distinction of count/mass syntax. Quine proposed that the infant's world is profoundly different from ours for lack of representations of enduring objects, and that it is by learning the count/mass syntax of a natural language, e.g. English, that the infant is able to "bootstrap" herself into a more adult-like conceptual scheme. On this view, cross-linguistic differences result in profound conceptual differences in adults; e.g. speakers of Japanese, which lacks the count/mass distinction, would not represent the ontological distinction between individuated entities, such as objects, and non-individuated entities, such as substances. Furthermore, children learning different languages would follow rather different developmental trajectories. Empirical investigations of this issue have found that the strong version of this view is wrong: even English-speaking children who have not mastered the count/mass syntax already differentiate objects from substances in extending word meanings (Soja, Carey, & Spelke, 1991). However, Imai and Gentner (1997) and Yoshida and Smith (in press) have argued that cross-linguistic differences may result in more subtle, quantitative differences.

Every child goes through the first few years of life from being prelinguistic to having a language, be it English, Chinese, or Hebrew. This universal fact of development provides us with an opportunity to ask questions regarding how language may shape cognition: what role does language play in shaping the child's conceptual representations? This question encompasses three sub-questions: first, how do universal aspects of language shape conceptual development? Are these conceptual representations built up in the aid of language (if any) uniquely human? Second, when do we begin to find the influence of cross-linguistic differences on cognition and thought (if ever)? Do these effects (if any) become stronger over time as the child is more and more immersed in her native language? This line of investigation could potentially contribute to the question of whether language contributes to habitual thought so the amount of exposure and entrenchment may result in quantitative, or even qualitative, differences among adults. Third, do children acquiring different languages follow a somewhat different developmental course in conceptual development? Perhaps certain linguistically-marked distinctions could facilitate the process of acquiring concepts but these differences are transient and they do not result in differences in adults (e.g. Sera, Bales, & Del Castillo Pintado, 1997). Thus, taking a developmental perspective may bring new ways of addressing the question of language and thought and enrich our understanding of the issues discussed in the adult literature.

Advocates of strong developmental continuity argue that learning language is simply a matter of mapping elements of the linguistic code onto pre-existing conceptual representations, and therefore that no fundamental change takes place in the process of language acquisition (e.g. Fodor, 1975; Pinker, 1984). In contrast, some have argued in recent years that language may play a causal role in changing the child's conception of the world in the domains of theory of mind, number, and spatial reorientation (de Villiers & de Villiers, 2000; Hale & Tager-Flusberg, 2000; Hermer & Spelke, 1997; Spelke & Tsivkin, 2001; Xu & Spelke, 2001). Some students of cognitive development have suggested a fairly radical view that language provides certain computational powers for the child (Hermer-Vazquez, Spelke, & Katsnelson, 1999). Others have suggested weaker versions in which language influences thought during "thinking for speaking" (Slobin, 1996) or in which cross-linguistic differences produce transient developmental differences (e.g. Yoshida & Smith, in press).

This paper investigates the first question of whether certain universal aspects of language may drive conceptual development. We explore the following hypothesis in the domain of sortal/kind concepts: does the process of learning count nouns towards the end of the first year change infants' conceptual representations of kinds?

Sortal concepts are concepts that provide criteria for individuation (where one object ends and another one begins) and identity (whether an object is the same one seen on a different occasion) (Gupta, 1980; Hirsch, 1982; Macnamara, 1986; Macnamara & Reyes, 1994; Wiggins, 1980). These concepts are lexicalized as count nouns in languages that make the count/mass distinction, e.g. *dog*, *cup*, *person*. The sortal concept *dog* provides criteria for deciding whether we are in the presence of one dog or two dogs, and it provides criteria for deciding whether we have seen the same dog on two different occasions, or two numerically distinct dogs. Sortal concepts such as *dog*, *cup*, or *ball* are a subset of what cognitive scientists refer to as "kind concepts" (which also include other concepts, such as *gold*, *water*, etc.).

Several recent studies have investigated when children begin to represent sortal/kind concepts. First, Spelke, Kestenbaum, Simons, and Wein (1995) (see also Spelke & Kestenbaum, 1986) and Xu and Carey (1996) found that at both 4 and 10 months, infants are able to use spatiotemporal criteria for object individuation, thus representing the sortal concept *physical object*. In contrast, not until 10–12 months are infants able to use basic-level sortals/kinds for object individuation. That is, by 12 months, infants are able to use the differences between a duck and a ball, or a cup and a bottle to establish a representation of two objects. In these studies, infants were shown a toy duck emerging from behind a screen and returning behind it, followed by a ball emerging from behind the same screen and returning. On the test trials, the screen was removed to reveal two objects (a duck and a ball, the expected outcome) or one object (the duck or the ball, the unexpected outcome). Looking times were recorded. The rationale behind the study was that if the infants had established a representation of two distinct objects, they should look longer at the unexpected outcome of one object. Xu and Carey (1996) found that at 10 months infants' looking time pattern on the test trials was not different from their baseline preference for one or two objects, whereas at 12 months infants overcame their baseline preference for two objects and looked longer at the unexpected outcome of one object. Two important control experiments showed that the methods were sensitive. First, when spatiotemporal evidence was given (by showing the two objects simultaneously for a few seconds at the beginning of the experiment), infants looked longer at the unexpected outcome of a single object. In other words, infants were able to overcome their baseline preference for two objects under some circumstances. Second, infants were able to perceive the property/featural differences between the two objects. Infants who were shown the alternating sequence duck–ball–duck–ball during familiarization trials habituated slower than the ones who were shown the same object over and over again. Xu and Carey (1996) hypothesized that 10-month-old infants do not represent basic-level sortal/kind concepts such as *duck*, *ball*, *cup*, or *bottle*. Subsequently, using both looking time and manual search measures, Xu, Carey, and Welch (1999) and Van de Walle, Carey, and Prevor (2001) provided convergent evidence for the shift between 10 and 12 months (see Xu, 1997, 1999, in press, for discussions and reviews).

However, recent studies by Wilcox, Needham and Baillargeon have challenged this claim (Needham, 1998; Needham & Baillargeon, 1997, 1998; Wilcox, 1999; Wilcox & Baillargeon, 1998a,b; Wilcox & Schweinle, 2002). Using several experimental paradigms, these researchers have shown that infants are able to use object property/featural information for object individuation under some circumstances as early as 4.5 months of age. In their studies, the experimental tasks were drastically simplified. In particular, Wilcox and Baillargeon's studies contrasted "event-monitoring" (in which the infants only have to detect inconsistencies in an event without having to compare their mental representations of the event to an outcome) and "event-mapping" (in which the infants have to compare their mental representations of the event with an outcome, as in Xu & Carey's studies), arguing that the latter imposed more information-processing demands on the infant.

Xu and Carey (2000), in a rejoinder to Needham and Baillargeon (2000), acknowledged that the strong claim made by Xu and Carey (1996) – that infants before 12 months do not use object property/featural information for object individuation – is untenable. At the same time, they argued that the use of object property/featural information is not the same as using object kind/sortal information in the service of object individuation. One hypothesis is that perhaps the simpler tasks used by Wilcox and her colleagues can be solved using property/featural information and infants succeed at a fairly young age, whereas the more complex tasks used by Xu and her colleagues require representations of sortals/kinds and these representations only begin to emerge at around 12 months. What is the evidence that the conceptual distinction between object kind/sortal and property/feature applies in the case of object individuation in infancy? Two studies are relevant for this purpose.

Xu, Carey, and Quint (2002) reported a series of experiments with 12-month-old infants, asking if the success at using the differences between, say a duck and ball, for object individuation is based on property differences (i.e. yellow, irregularly shaped, rubbery vs. red, round, and shiny) or kind differences (i.e. a member of the kind duck vs. a member of the kind ball). To this end, they asked if infants would individuate objects based on property differences alone, e.g. color differences (e.g. a red ball vs. a green ball), size differences (e.g. a small red ball vs. a big red ball), or combinations of these properties (e.g. a small red ball vs. a big green ball). The results showed that infants failed to use these property differences for object individuation. In the last experiment of this series, they contrasted within-kind shape changes (e.g. a regular cup vs. a sippy cup of the same size and surface pattern) and cross-kind shape changes (e.g. a regular cup vs. a bottle of the same size and surface pattern). The infants failed to use within-kind shape changes for object individuation whereas they succeeded in using cross-kind shape changes, even when the two types of shape changes are roughly equally salient to the infants. These data suggest that at 12 months, infants' success in a complex object individuation task may be based on object kind representations as opposed to property representations. Because this series of experiments also used simple objects, e.g. a red ball vs. a green ball, it is likely that the discrepancy between Xu and Carey (1996) and Wilcox and Baillargeon (1998a) is due to the different experimental paradigms. Xu and Carey (2000) suggested that when a more complex procedure is employed, representations of sortals/kinds may be required, as in the Xu and Carey experiments.

In addition, Waxman and Markow (1995) and Waxman (1999) found that by 13 months, the distinction between property and kind plays a role in infants' categorization. They found that infants were sensitive to whether they heard a count noun or an adjective while examining the objects. If they heard an adjective, they were more likely to generalize based on a property (e.g. color or texture) whereas if they heard a count noun, they were more likely to generalize to objects of the same kind, based on shape. These results are consistent with the Xu et al. (2002) findings that the distinction between kinds and properties is present by about 13 months, and it maps onto the linguistic distinction between count nouns and adjectives, just as it does in adults.

These studies suggest that although infants' ability to process more complex objects and events improves dramatically over the course of the first year, sortal/kind representations also emerge towards the end of the first year. Thus, we are left with the following question: how do infants acquire sortal/kind concepts? That is, what is the mechanism that underlies the change between 10 and 12 months?

Two findings in the literature directly motivated the current studies. Balaban and Waxman (1996) found that words, but not tones, facilitate categorization in 9-month-old infants. The infants were familiarized with a set of pictures of a given category, say, rabbits. Some infants heard a word when shown the picture, e.g. "a rabbit". For other infants, a tone accompanied the presentation of the picture on some trials. The results showed that although both words and tones heightened the infants' attention to the objects, it was only in the label condition that the infants categorized the objects. That is, they preferentially looked at an exemplar from a new category (e.g. a pig) compared to an exemplar from the old category (e.g. a new rabbit). These results suggest that in the presence of a label, infants group together the exemplars into a single category more readily than in the absence of a label. In addition, a post hoc analysis from Xu and Carey (1996) provides a hint that language may play a role in acquiring sortal/kind concepts. Xu and Carey analyzed the looking time data from 10-month-old infants as a function of whether they understood any of the words that named the objects in the experiment by parental report. They found a correlation between the two variables: the infants who were reported to understand some words looked longer at the single-object, unexpected outcome whereas those who did not understand any words didn't. Perhaps knowing the words for these objects is a means of establishing that they belong to different kinds which in turn allowed the infants to succeed in the object individuation task.

The four experiments reported in this paper put this idea to a direct test by investigating the role of language in a complex object individuation task. We ask the question whether verbal labels may help 9-month-old infants to establish a representation of two numerically distinct objects behind an occluder.

2. Experiment 1

Experiment 1 investigates the role of labeling in a complex object individuation task. Infants were shown two objects, a toy duck and a ball, emerging from behind a screen, one at a time. In the two-word condition, as each object appeared, the infants heard the appropriate label, "Look, a duck" or "Look, a ball". In the one-word condition, the infants heard the same label for both objects, "Look, a toy". On the test trials, the screen was removed to reveal one or two objects and looking times were recorded. A baseline condition was included to assess the infants' intrinsic preference for the outcomes of one or two objects. The question of interest was whether two labels would help infants succeed on this task at an earlier age than 12 months. Nine-month-old infants were chosen for this experiment because this was the age at which Waxman and colleagues have found word facilitation effects on categorization.

2.1. Methods

2.1.1. Participants

Thirty-six full-term 9-month-old infants (half girls and half boys) participated in the experiment (mean age 9 months, 3 days; range 8 months, 18 days to 9 months, 16 days). Equal numbers of boys and girls were randomly assigned to one of three conditions: the Two-word Condition (mean age 9 months, 2 days; range 8 months, 20 days to 9 months, 2 days), the One-word Condition (mean age 9 months, 3 days; range 8 months, 18 days to 9 months, 16 days), and the Baseline Condition (mean age 9 months, 3 days; range 8 months, 18 days to 9 months, 15 days). All infants were recruited from the Greater Boston area by mail and subsequent phone calls. Most infants came from a middle-class non-Hispanic white background with about 10% of Asian, African-American, and Hispanic infants. The infants received a token gift (a T-shirt, a bib, or a sippy cup with a university logo) after the study. Three additional infants were excluded because of fussiness or parental interference. English was the primary language spoken at home for all infants.

2.1.2. Materials

Two objects were used in the study: a toy duck and a ball. The toy duck was bright yellow with an orange bill, black eyes, and orange feet (approximately $6 \times 6 \times 9$ cm in size) and the ball was felt with pink and green stripes (approximately 6 cm in diameter). Each object had a stick attached to the bottom so the experimenter could move it along a track on the stage. Four foam core screens (red, lavender, orange, and pink) were used; each measured 34 cm in width and 26 cm in height.

2.1.3. Apparatus

The events were presented on a three-sided stage (80 cm in length, 31 cm in width, and

30 cm in height) with a light blue surface. A black curtain hung behind the stage to put the objects in high contrast and to conceal the movement of the experimenter. Another black curtain hung over the front part of the stage to conceal the video camera under the stage as well as to prevent the experimenter from seeing the infant's face. The actual display area measured 80 cm in width and 24 cm in height. A slit ran across the stage so the objects could be moved when they were on sticks.

The video camera was connected to a 19 inch color TV that was placed in one corner of the room. An observer watched the infant on the TV monitor and recorded the looking times. The observer could not see what was presented on the stage and was blind to which condition the infant was in. A push button was connected to a computer that recorded looking times. A computer program written specifically for looking time studies (Xhab; Pinto, 1995) was used to record the looking times in all the experiments. White noise masked any sounds produced by the movements of the experimenter.

The stage was lit from above and from the two sides; otherwise the room was dark. The infant sat in a high chair, about 70 cm from the stage, with eye level slightly above the floor of the stage (about 5 cm). The parent sat next to the infant with his/her back toward the stage, and was instructed not to look at the displays and not to draw the infant's attention in any way. A video camera was set up under the stage, focusing on the infant's face and recording the entire session. The videotape record provided no information about what was presented on the stage so an observer scoring from the videotapes was completely blind to the condition and the order of the trials.

2.1.4. Design and procedure

All parents were asked to fill out a questionnaire where they were asked if their infants were familiar with the objects and whether they comprehended any words. The word list included the top 30 nouns for 9-month-old infants from the MacArthur Communicative Development Inventory (Fenson et al., 1993).

Equal numbers of infants were randomly assigned to three conditions: the Two-word Condition, the One-word Condition, and the Baseline Condition. The experimenter began by waving a set of keys at all six corners of the stage (the top and bottom of left, right, and center of the stage) in order to draw the infant's attention to the stage as well as to define the window of looking for the observer. During the experiment, the experimenter stood behind the back curtain and only her hands were sometimes visible to the infant.

2.1.4.1. Two-word Condition (Fig. 1)

2.1.4.1.1. Familiarization trials Each infant received 14 familiarizations: 12 were accompanied by a label and two were not. A screen was put on the stage. Each familiarization trial began when an object, e.g. a toy duck, was brought out from behind the screen. In infant-directed speech, the experimenter said, "Look, [baby's name]". To draw the infant's attention, she tapped the object a few times on the stage by lifting and lowering the stick from beneath, then said, "a duck". The word "duck" was always in the final position and stressed. The toy duck then returned behind the screen, marking the end of the familiarization trial. After a 2 s pause, the second familiarization trial began. A second object, a ball, was brought out from behind the screen to the other side of the stage. After the experimenter tapped it and labeled it, "a ball", it was returned

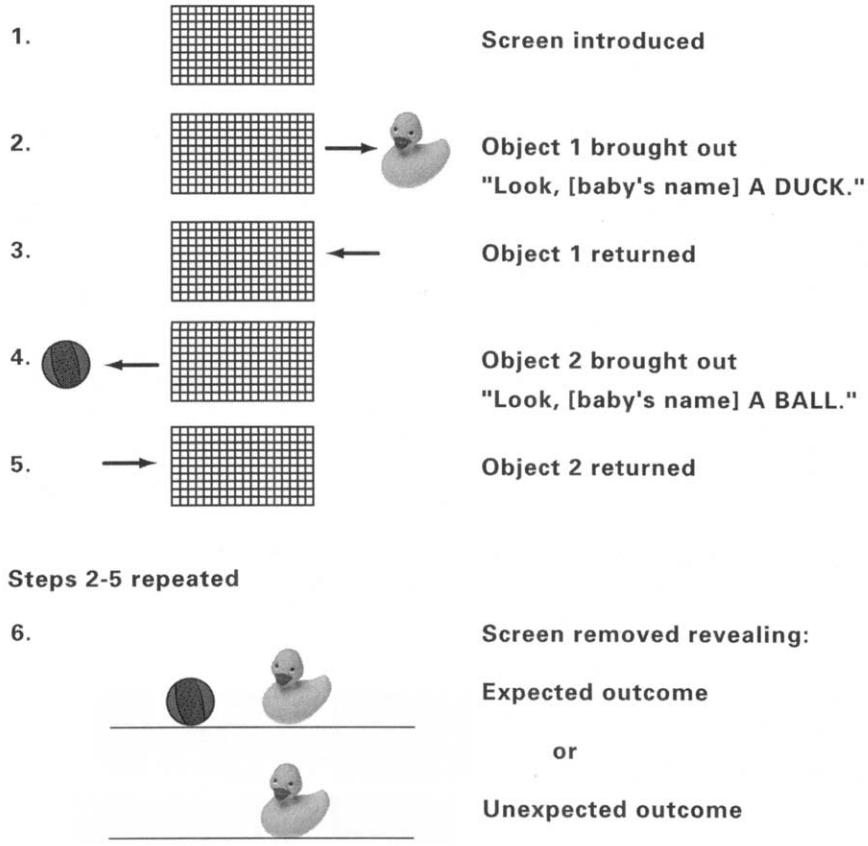

Fig. 1. Schematic representation of the events in Experiment 1: Two-word Condition.

behind the screen. Each infant saw each object seven times for a total of 14 familiarizations. On four of these 14 familiarizations (the 3rd and 4th, and the 7th and 8th), the object was left stationary for the infant to look at until she turned away; looking times were recorded. On two of these four familiarizations, the objects were labeled (labeled trials); on the other two trials, the objects were not labeled (silent trials). On the silent trials, the experimenter simply called the infant's name and drew her attention to the object. The silent trials were included to replicate an earlier finding that when an object was labeled, infants looked longer at it (Baldwin & Markman, 1989). Which object appeared first, which side of the screen it appeared on, and the order of the silent and the labeled trials were counterbalanced across infants.

2.1.4.1.2. Test trials At the end of the familiarization trials, each object was shown briefly, then the screen was turned to the side, revealing one or two objects. As the screen was turned, the experimenter drew the infant's attention, "Look, [baby's name]". Looking time was then recorded. At the end of each trial, the stage was cleared. A screen of a different color was then lowered with the same two objects hidden behind it. Each object was shown briefly and labeled again before the screen was turned to the side to reveal the objects. Each infant was shown four test trials, alternating between two outcomes. In the expected outcome, the infant was shown two objects sitting side by side when the screen was turned; in the unexpected outcome, the infant was shown only one of the two objects (the other object had been removed swiftly by the experimenter at the end of the familiarization trials). The experimenter went through the motion of removing the object even on the expected trials in order to equate the timing of all the test trials. The order of outcome (1, 2, 1, 2 or 2, 1, 2, 1) and which object was the single object were counterbalanced across infants.

For both the familiarization trials (in which the object was left stationery) and the test trials, a trial ended when the infant looked away for 2 consecutive seconds. The minimum looking time was 0.5 s and the maximum looking time was 120 s.

2.1.4.2. One-word Condition The procedure was identical to that of the Two-word Condition with one important difference: one label (i.e. "a toy") was used to label both objects during familiarization trials.

2.1.4.3. Baseline Condition The infants in this condition saw a screen being put on the stage. The screen was then turned to the side to reveal one or two objects. Eight trials were run; looking times were recorded. This was to assess if the infants had any intrinsic preference for looking at the expected outcome of two objects or the unexpected outcome of one object. The order of outcome (1, 2 or 2, 1) was counterbalanced across infants.

2.2. Results

The main results of Experiment 1 are shown in Fig. 2. An alpha level of 0.05 was used in all statistical analyses. Preliminary analyses found no effects of sex, order of outcome, and which object was the single object. Subsequent analyses were collapsed over these variables. All infants were off-line observed by a second observer who was completely blind to the condition and order of outcome of the experiment. Interscorer reliability averaged 92%.

2.2.1. Familiarization trials

An analysis of variance was performed on the four familiarization trials in which the objects were left stationery with condition (one-word vs. two-word) and trial type (labeled vs. silent) as variables. There was a main effect of type of trial ($F(1,22) = 4.562$, $P < 0.05$). The infants looked longer when the objects were labeled than when they were not ($M_{\text{labeled}} = 11.0$ s, SD = 6.2; $M_{\text{silent}} = 8.6$ s, SD = 4.6). There was no interaction between the two variables ($F(1,22) = 0.048$, $P = 0.828$).

Planned comparisons were performed on the silent and labeled trials within each condition. There was a marginally statistically significant main effect of trial type in each condition (One-word Condition: $F(1,11) = 3.562$, $P = 0.07$; $M_{\text{labeled}} = 11.5$ s, SD = 7.3; $M_{\text{silent}} = 8.8$ s, SD = 5.4; Two-word Condition: $F(1,11) = 3.320$, $P = 0.09$; $M_{\text{labeled}} = 10.5$ s, SD = 8.4; $M_{\text{silent}} = 8.4$ s, SD = 3.8).

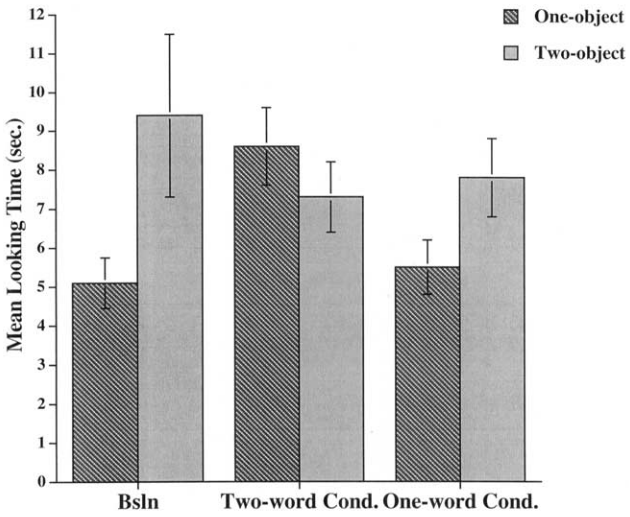

Fig. 2. Mean looking times of Experiment 1 with standard errors.

2.2.2. Test trials

Analyses of variance were performed comparing each of the experimental conditions with the baseline condition and the two experimental conditions with each other, with condition and outcome number (1 vs. 2) as variables.

2.2.3. Two-word Condition and Baseline Condition

There was an interaction between condition (baseline vs. two-word) and outcome number (1 vs. 2) ($F(1, 22) = 6.938$, $P < 0.02$). Planned comparisons were performed for each condition. The infants looked longer at the two-object outcome in the Baseline Condition ($M_{\text{one-object}} = 5.1$ s, SD = 2.3; $M_{\text{two-object}} = 9.4$ s, SD = 7.3) ($F(1, 11) = 6.786$, $P < 0.05$). They looked marginally longer at the one-object outcome in the Two-word Condition upon hearing two labels for the two objects ($M_{\text{one-object}} = 8.6$ s, SD = 5.0; $M_{\text{two-object}} = 7.3$ s, SD = 3.1) ($F(1, 11) = 3.524$, $P = 0.08$). Ten of the 12 infants in the Two-word Condition looked longer at the one-object, unexpected outcome (Wilcoxon $z = -2.080$, $P < 0.05$).

2.2.4. One-word Condition and Baseline Condition

There was a main effect of outcome number ($F(1, 22) = 8.988$, $P < 0.01$). The infants looked longer at the two-object outcome ($M_{\text{one-object}} = 5.2$ s, SD = 2.4; $M_{\text{two-object}} = 8.6$ s, SD = 5.6). There was no interaction between condition and outcome number ($F(1, 22) = 0.754$, $P = 0.395$). Planned comparisons were performed for each condition. The infants looked longer at the two-object outcome in the Baseline Condition ($M_{\text{one-object}} = 5.1$ s, SD = 2.3; $M_{\text{two-object}} = 9.4$ s, SD = 7.3) ($F(1, 11) = 6.786$, $P < 0.05$) as well as in the One-word Condition ($M_{\text{one-object}} = 5.4$ s, SD = 2.6; $M_{\text{two-object}} = 7.8$ s, SD = 3.5) ($F(1, 11) = 5.524$, $P < 0.05$). Ten of the 12 infants in the One-word Condition looked longer at the two-object, expected outcome.

2.2.5. One-word Condition and Two-word Condition

There was a marginally significant interaction between condition and outcome number ($F(1, 22) = 3.303$, $P = 0.08$). The infants showed different patterns of looking on the test trials of the two conditions.

2.2.6. Parental report of word comprehension

For the Two-word Condition, ten out of 12 parents reported that their infants were familiar with ducks and balls. Only two reported that their infants understand the words "duck" or "ball". The average estimated comprehension vocabulary for this group of infants was three words. For the One-word Condition, 11 out of 12 parents reported that their infants were familiar with ducks and balls. Three reported that their infants understand the word "duck" and five the word "ball". The average estimated comprehension vocabulary for this group of infants was four words. Thus, most of the infants were familiar with the objects but very few of them comprehended the words for these objects.

2.3. Discussion

In Experiment 1, 9-month-old infants succeeded in an object individuation task when provided with two distinct labels for two objects; they failed the same task when a single label was provided for both objects. That is, only when two distinct labels were given were the infants able to use them to establish two distinct objects behind the screen. Furthermore, the familiarization trials showed that in both the One-word and the Two-word Conditions, the infants looked longer when the objects were labeled relative to when they were unlabeled, and the presence of one word increased the infants' attention to the objects as much as the presence of two words. This result suggests that the success in the Two-word Condition was not simply due to an increase of attention when two words were present.

Before concluding that this facilitation effect is due to the presence of language per se, it is important to ask if other kinds of auditory signal could also help infants in an object individuation task. Experiment 2 addressed this question by providing the infants with two distinct tones for the objects. We also used a different pair of objects to test the generality of the findings in Experiment 1.

3. Experiment 2

This experiment used a similar procedure as that of Experiment 1. In the Two-word Condition, again two distinct labels were provided as the objects were shown to the infants; in the Two-tone Condition, two distinct tones were paired with the objects.

3.1. Method

3.1.1. Participants

Thirty-six full-term 9-month-old infants (18 girls and 18 boys) participated in the study (mean age 9 months, 0 days; range 8 months, 18 days to 9 months, 19 days). Equal numbers of boys and girls were randomly assigned to one of three conditions: the Two-word Condition (mean age 9 months, 0 days; range 8 months, 21 days to 9 months, 19 days), the Two-tone Condition (mean age 9 months, 0 days; range 8 months, 18 days to 9 months, 18 days), and the Baseline Condition (mean age 9 months, 1 day; range 8 months, 18 days to 9 months, 17 days). All infants were recruited from the same population as in Experiment 1. Four additional infants were excluded because of fussiness or parental interference. English was the primary language spoken at home for all infants.

3.1.2. Materials and apparatus

Two objects were used in the experiment: a sippy cup with yellow and orange vertical stripes (about $11 \times 8 \times 7.5$ cm in size) and a baby shoe with red, green, blue and white stripes (about $12.5 \times 6 \times 5$ cm in size). The same four foam core screens were used as in Experiment 1. The apparatus was the same as in Experiment 1.

3.1.3. Design and procedure

Equal numbers of infants were randomly assigned to three conditions: the Two-word Condition, the Two-tone Condition, and the Baseline Condition. The experiment began with a calibration as in Experiment 1.

3.1.3.1. Two-word Condition The procedure was identical to the Two-word Condition of Experiment 1.

3.1.3.2. Two-tone Condition The procedure was identical to the Two-word Condition except for one important difference: instead of hearing two labels for the objects, the infants heard two distinct tones played from a keyboard behind the stage (a D-sharp and a C one octave above the middle C). The tones were approximately the same volume as the words in the Two-word Condition.

3.1.3.3. Baseline Condition The procedure was identical to the Baseline Condition of Experiment 1.

3.2. Results

The main results of Experiment 2 are shown in Fig. 3. An alpha level of 0.05 was used in all statistical analyses. Preliminary analyses found no effects of sex, order of outcome, and which object was the single object. Subsequent analyses were collapsed over these variables. All infants were off-line observed by a second observer who was completely blind to the condition and order of outcome of the experiment. Interscorer reliability averaged 93%.

3.2.1. Familiarization trials

An analysis of variance was performed comparing the familiarization trials with condition (two-word vs. two-tone) and trial type (labeled/tone vs. silent) as variables. There was a marginal main effect of type of trial ($F(1,23) = 2.980$, $P = 0.09$) ($M_{\text{labeled}} = 6.5$ s, SD = 4.8; $M_{\text{silent}} = 5.2$ s, SD = 3.8). When planned comparisons were performed on each condition, the effect of trial type was not statistically significant (Two-word Condition: $M_{\text{labeled}} = 6.0$ s, SD = 5.2; $M_{\text{silent}} = 3.8$ s, SD = 2.3; $F(1,11) = 2.532$, $P = 0.14$; Two-tone Condition: $M_{\text{tone}} = 7.0$ s, SD = 4.5; $M_{\text{silent}} = 6.6$ s, SD = 4.5; $F(1,11) = 0.425$, $P = 0.528$).

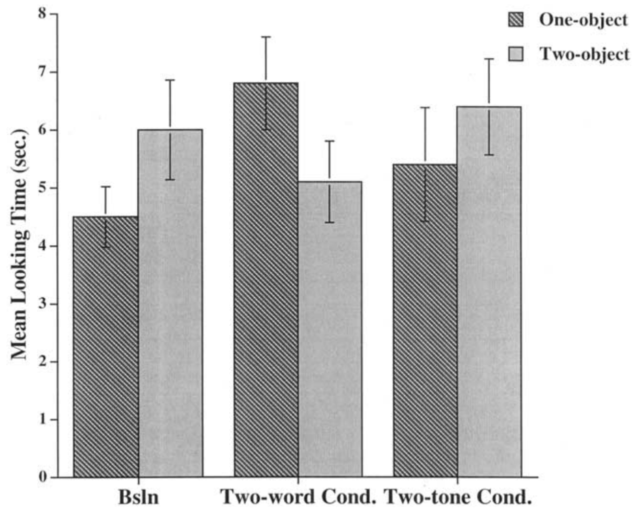

Fig. 3. Mean looking times of Experiment 2 with standard errors.

3.2.2. Test trials

Analyses of variance were performed comparing each of the experimental conditions with the baseline condition and the two experimental conditions with each other, with condition and outcome number (1 vs. 2) as variables.

3.2.3. Two-word Condition and Baseline Condition

There was an interaction between condition (baseline vs. two-word) and outcome number (1 vs. 2) ($F(1,22) = 4,386$, $P < 0.05$). Planned comparisons were performed within each condition. The infants looked slightly longer at the two-object outcome in the Baseline Condition ($F(1,11) = 2.609$, $P = 0.14$) ($M_{\text{one-object}} = 4.5$ s, SD = 1.8; $M_{\text{two-object}} = 6.0$ s, SD = 3.0). They looked reliably longer at the one-object outcome in the Two-word Condition upon hearing two labels for the two objects ($F(1,11) = 4.386$, $P < 0.05$) ($M_{\text{one-object}} = 6.8$ s, SD = 4.0; $M_{\text{two-object}} = 5.1$ s, SD = 3.3). Ten of the 12 infants in the Two-word Condition looked longer at the one-object, unexpected outcome (Wilcoxon $z = -2.510$, $P < 0.05$).

3.2.4. Two-tone Condition and Baseline Condition

There was a marginal main effect of outcome number ($F(1,22) = 3.442$, $P = 0.07$) ($M_{\text{one-object}} = 5.4$ s, SD = 3.3; $M_{\text{two-object}} = 6.4$ s, SD = 2.8). There was no interaction between condition and outcome number ($F(1,22) = 0.199$, $P = 0.660$). Planned comparisons were performed within each condition. The infants looked slightly longer at the two-object outcome in the Baseline Condition ($F(1,11) = 2.609$, $P = 0.14$) ($M_{\text{one-object}} = 4.5$ s, SD = 1.8; $M_{\text{two-object}} = 6.0$ s, SD = 3.0); they did not look reliably longer at either outcome in the Two-tone Condition ($F(1,11) = 1.008$, $P = 0.338$) ($M_{\text{one-object}} = 5.4$ s, SD = 3.3; $M_{\text{two-object}} = 6.3$ s, SD = 2.8). Ten of the 12 infants in the Two-tone Condition looked longer at the two-object, expected outcome.

3.2.5. Two-word Condition and Two-tone Condition

There was an interaction between condition and outcome number ($F(1, 22) = 4.386$, $P < 0.05$). The infants showed different patterns of looking on the test trials of the two conditions.

3.2.6. Parental report of word comprehension

For the Two-word Condition, nine out of 12 parents reported that their infants were familiar with shoes and cups. Only two parents reported that their infants understand the word "cup" and one the word "shoe". The average estimated comprehension vocabulary for this group of infants was three words. For the Two-tone Condition, eight out of 12 parents reported that their infants were familiar with shoes and cups. None reported that their infants understand the words "cup" or "shoe". The average estimated comprehension vocabulary for this group of infants was two words. Thus, most of the infants were familiar with the objects but very few of them comprehended the words for them.

3.3. Discussion

With a different pair of objects, Experiment 2 replicated the basic result of Experiment 1 that two contrastive labels helped the infants succeed in this task. In contrast, with two non-speech sounds (in this case two distinct tones), the infants failed at this task even though the procedure was identical to that of the Two-word Condition. These results suggest that words play a special role in facilitating object individuation.

However, there is an alternative account for the failure in the Two-tone Condition, namely that the tones were not distinct enough from each other so the infants could not differentiate them. In Experiment 3, we replaced the two tones with two very different sounds: one was a two-tone sequence, and the other was a sound produced with a gadget that allowed us to play a car-alarm sound followed by a spaceship sound. The two sound sequences were different in timbre, pitch, and chromaticity. The complexity of the sounds made them more distinct from each other.

In Experiment 3, we also explored the role of familiarity, that is, whether either the objects or the words need to be familiar to the infants for labeling to facilitate object individuation. Although the parental reports we collected in Experiments 1 and 2 showed very little comprehension of words such as duck, ball, shoe, and cup, we were not certain that parents' reports were accurate. Recent work by Tincoff and Jusczyk (1999) suggested that even 5-month-old infants may comprehend highly frequent words like "Mommy" and "Daddy". To rule out the possibility that we simply chose objects that at least some infants had words for, we used novel objects and nonsense words in Experiment 3.

4. Experiment 3

This experiment investigated whether the same facilitation effects found in Experiments 1 and 2 would replicate with novel objects and nonsense words, and whether two distinct sounds would help infants to succeed earlier.

4.1. Methods

4.1.1. Participants

Forty-eight full-term 9-month-old infants (half girls and half boys; mean age 9 months, 1 day; range 8 months, 15 days to 9 months, 15 days) participated in the experiment. Equal numbers of infants were randomly assigned to three conditions: the Two-word Condition (mean age 8 months, 29 days; range 8 months, 15 days to 9 months, 15 days), the Two-sound Condition (mean age 9 months, 2 days; range 8 months, 16 days to 9 months, 15 days), and the Baseline Condition (mean age 9 months, 2 days; range 8 months, 20 days to 9 months, 15 days). The infants were recruited from the same population as in Experiments 1 and 2. Six additional infants were excluded because of fussiness or parental interference. English was the primary language spoken at home for all infants.

4.1.2. Materials and apparatus

Two novel objects were used in the experiment: a purple lizard-like animal and an artifact-like made-up object that consisted of a dark blue rectangular box and a big orange arrow on top. The lizard-like animal was attached to a stick and presented vertically; it had two small black eyes on its head, silver dots on its body, and four legs and two gosamer wings (approximately $10 \times 13 \times 5$ cm in size). The rectangle box of the artifact-like object was decorated with three small orange triangles and glitter (approximately $10 \times 13 \times 4$ cm in size). The same four foam core screens were used as in Experiments 1 and 2. The apparatus was identical to those used in Experiments 1 and 2.

4.1.3. Design and procedure

Equal numbers of infants were randomly assigned to three conditions: the Two-word Condition, the Two-sound Condition, and the Baseline Condition. The experiment began with a calibration as in Experiment 1.

4.1.3.1. Two-word Condition The procedure was identical to the Two-word Condition of Experiment 1 except that two nonsense words were used, “fendle” and “toma”. The pairing of the objects and the words was counterbalanced across subjects.

4.1.3.2. Two-sound Condition The procedure was identical to the Two-word Condition except for one important difference: instead of hearing two labels for the objects, the infants heard two distinct sound sequences. One was a tone sequence played from a keyboard behind the stage (a D-sharp followed by a C one octave above the middle C); the other was a car-alarm sound followed by a spaceship sound played from a store-bought gadget. The two sound sequences were drastically different from each other. The sounds were approximately the same volume as the words in the Two-word Condition.

4.1.3.3. Baseline Condition The procedure was identical to the Baseline Condition of Experiment 1.

4.2. Results

The main results of Experiment 2 are shown in Fig. 4. An alpha level of 0.05 was used in all statistical analyses. Preliminary analyses found no effects of sex, order of outcome, and which object was the single object. Subsequent analyses were collapsed over these variables. Half of the infants were off-line observed by a second observer who was completely blind to the condition and order of outcome of the experiment. Interscorer reliability averaged 91%.

4.2.1. Familiarization trials

An analysis of variance was performed comparing the familiarization trials with condition (two-word vs. two-sound) and trial type (labeled/sound vs. silent) as variables. There was a marginal main effect of trial type ($F(1,30) = 3.764$, $P = 0.07$). The infants looked longer when the objects were accompanied by a label or a sound than when they were not ($M_{\text{labeled/sound}} = 6.4$ s, SD = 3.8; $M_{\text{silent}} = 5.4$ s, SD = 3.5).

Planned comparisons were performed within each condition. In the Two-word Condition, there was a marginal main effect of trial type ($F(1,15) = 3.978$, $P = 0.08$) ($M_{\text{labeled}} = 6.8$ s, SD = 3.7; $M_{\text{silent}} = 4.9$ s, SD = 3.3). In the Two-sound Condition, there was a main effect of trial type ($F(1,15) = 7.059$, $P < 0.05$) ($M_{\text{sound}} = 5.9$ s, SD = 2.8; $M_{\text{silent}} = 4.3$ s, SD = 3.3).

4.2.2. Test trials

Analyses of variance were performed comparing each of the experimental conditions with the baseline condition and the two experimental conditions with each other, with condition and outcome number (1 vs. 2) as variables.

4.2.3. Two-word Condition and Baseline Condition

There was an interaction between condition and outcome number ($F(1,30) = 8.889$,

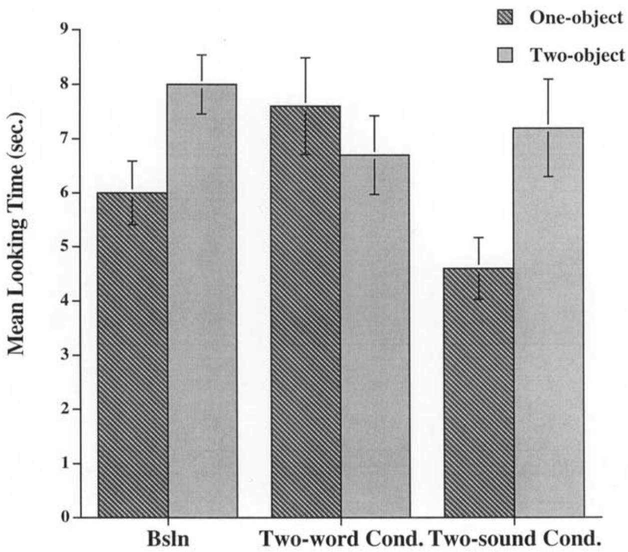

Fig. 4. Mean looking times of Experiment 3 with standard errors.

$P < 0.01$). The infants looked longer at the two-object outcome in the Baseline Condition ($M_{\text{one-object}} = 6.0$ s, $SD = 2.3$; $M_{\text{two-object}} = 8.0$ s, $SD = 2.1$) ($F(1, 15) = 21.754$, $P < 0.001$). They did not look reliably longer at the one-object outcome in the Two-word Condition upon hearing two nonsense words for the two novel objects ($M_{\text{one-object}} = 7.6$ s, $SD = 3.5$; $M_{\text{two-object}} = 6.6$ s, $SD = 2.9$) ($F(1, 15) = 0.741$, $P > 0.1$). Twelve of the 16 infants in the Two-word Condition looked longer at the one-object, unexpected outcome (Wilcoxon $P > 0.1$).

4.2.4. Two-sound Condition and Baseline Condition

There was a main effect of outcome number ($F(1, 30) = 36.936$, $P < 0.0001$). The infants looked longer at the two-object outcome than the one-object outcome. There was no interaction between condition and outcome number ($F(1, 30) = 0.798$, $P = 0.379$). Planned comparisons were performed within each condition. The infants looked longer at the two-object outcome in both the Baseline Condition and the Two-sound Condition ($M_{\text{one-object}} = 4.6$ s, $SD = 2.2$; $M_{\text{two-object}} = 7.2$ s, $SD = 3.6$) ($F(1, 15) = 17.576$, $P < 0.001$). Twelve of the 16 infants in the Two-sound Condition looked longer at the two-object, expected outcome.

4.2.5. Two-word Condition and Two-sound Condition

There was an interaction between condition and outcome number ($F(1, 30) = 10.961$, $P < 0.005$). The infants showed different patterns of looking on the test trials of the two conditions.

4.3. Discussion

In Experiment 3, we replicated the basic facilitation effects of Experiments 1 and 2 with a pair of novel objects and two nonsense words. Infants who heard the two novel words paired with the unfamiliar objects succeeded in establishing a representation of two distinct objects behind the screen whereas infants who heard two distinct sounds failed to do so. These results suggest that labeling plays an important role in establishing representations of distinct objects.

However, the comparison between words and sounds is perhaps unsatisfactory. A critical difference between words and sounds is that words are produced by the human vocal tract and they are intentional whereas the sounds used in Experiments 2 and 3 are neither uttered by humans nor intentional. Recent work by Bloom (2000) suggests that intention plays a critical role in word learning. Thus, a better comparison may be between words and other intentional sounds produced by human beings, e.g. emotional expressions.

5. Experiment 4

Experiment 4 sought to compare the effects of words and emotional expressions on object individuation.

5.1. Methods

5.1.1. Participants

Forty-eight full-term 9-month-old infants (half girls and half boys; mean age 9 months, 0 days; range 8 months, 15 days to 9 months, 15 days) participated in the experiment. Sixteen infants were randomly assigned to one of three conditions: the Word Condition (mean age 8 months, 25 days; range 8 months, 15 days to 9 months, 10 days), the Emotional-expression Condition (mean age 9 months, 1 day; range 8 months, 17 days to 9 months, 15 days), and the Baseline Condition (mean age 9 months, 2 days; range 8 months, 20 days to 9 months, 15 days). The infants were recruited from the same population as in previous experiments. Three additional infants were excluded because of fussiness (two) or experimenter error (one). English was the primary language spoken at home for all infants.

5.1.2. Materials and apparatus

The same two novel objects were used in this experiment: a purple lizard-like animal and an artifact-like “box and arrow”. The same four foam core screens were used as in previous experiments. The apparatus was identical to that used in previous experiments.

5.1.3. Design and procedure

Equal numbers of infants were randomly assigned to three conditions: the Word Condition, the Emotional-expression Condition, and the Baseline Condition. The experiment began with a calibration as in Experiment 3.

5.1.3.1. Word Condition The procedure was identical to the Two-word Condition of Experiment 3 except that the nonsense words were “blicket” and “tupa”. The pairing of the objects and the words was counterbalanced across infants.

5.1.3.2. Emotional-expression Condition The procedure was identical to the Word Condition except for one important difference: instead of hearing two labels for the objects, the infants heard two distinct emotional expressions produced by the experimenter. One emotional expression was an “Ah” sound carrying positive valence to express approval or satisfaction; the other emotional expression was an “Ewy” sound carrying negative valence to express dislike or disgust. The two expressions were produced for approximately the same duration and with approximately the same volume. The emotional expressions and the words were produced by the same female human voice at approximately the same volume.

5.1.3.3. Baseline Condition The procedure was identical to the Baseline Condition of Experiment 1.

5.2. Results

The main results of Experiment 4 are shown in Fig. 5. An alpha level of 0.05 was used in all statistical analyses. Preliminary analyses found no effects of sex, order of outcome, and which object was the single object. Subsequent analyses were collapsed over these vari-

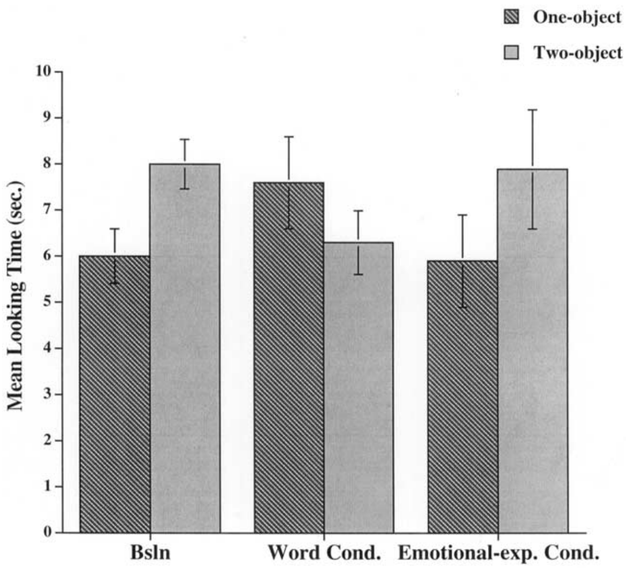

Fig. 5. Mean looking times of Experiment 4 with standard errors.

ables. Half of the infants were off-line observed by a second observer who was completely blind to the condition and order of outcome of the experiment. Interscorer reliability averaged 90%.

5.2.1. Familiarization trials

An analysis of variance compared the familiarization trials with condition (word vs. emotional-expression) and trial type (labeled/emotional expression vs. silent) as variables. There was a main effect of type of trial ($F(1, 30) = 4.502$, $P < 0.05$). The infants looked longer when the objects were accompanied by a label or an emotional expression ($M = 6.0$ s, SD $= 2.8$) than when they were not ($M = 5.0$ s, SD $= 3.0$).

Planned comparisons were performed within each condition. In the Word Condition, there was a marginal main effect of trial type ($F(1, 15) = 3.978$, $P = 0.08$) ($M_{\text{labeled}} = 5.8$ s, SD $= 3.7$; $M_{\text{silent}} = 4.9$ s, SD $= 3.3$). In the Emotional-expression Condition, there was also a marginal main effect of trial type ($F(1, 15) = 3.750$, $P = 0.08$) ($M_{\text{emo}} = 6.2$ s, SD $= 2.8$; $M_{\text{silent}} = 5.1$ s, SD $= 3.3$).

5.2.2. Test trials

Analyses of variance were performed comparing each of the experimental conditions with the baseline condition and the two experimental conditions with each other, with condition and outcome number (1 vs. 2) as variables.

5.2.3. Word Condition and Baseline Condition

There was an interaction between condition and outcome number ($F(1, 30) = 8.830$, $P < 0.01$). Planned comparisons were performed within each condition. The infants looked longer at the two-object outcome in the Baseline Condition ($M_{\text{one-object}} = 8.0$ s, SD = 2.1; $M_{\text{two-object}} = 6.0$ s, SD = 2.3) ($F(1,15) = 15.754$, $P < 0.01$). They did not look reliably longer at the one-object outcome in the Word Condition upon hearing two nonsense words for the two novel objects ($M_{\text{one-object}} = 7.6$ s, SD = 4.3; $M_{\text{two-object}} = 6.3$ s, SD = 2.7) ($F(1,15) = 1.572$, $P > 0.1$). Twelve of the 16 infants in the Word Condition looked longer at the one-object, unexpected outcome (Wilcoxon $P > 0.1$).

5.2.4. Emotional-expression Condition and Baseline Condition

There was a main effect of outcome number ($F(1,30) = 13.297$, $P < 0.001$). The infants looked longer at the two-object outcome ($M = 8.0$ s, SD = 5.2) than the one-object outcome ($M = 5.9$ s, SD = 4.3). There was no interaction between condition and outcome number ($F(1,30) = 0.001$, $P = 0.978$). Planned comparisons were performed within each condition. The infants looked longer at the two-object outcome in the Baseline Condition, and they looked marginally longer at the two-object outcome in the Emotional-expression Condition ($M_{\text{one-object}} = 5.9$ s, SD = 4.3; $M_{\text{two-object}} = 7.9$ s, SD = 5.2) ($F(1,15) = 3.874$, $P = 0.07$). Fourteen of the 16 infants in the Emotional-expression Condition looked longer at the two-object, expected outcome.

5.2.5. Two-word Condition and Emotional-expression Condition

There was an interaction between condition and outcome number ($F(1,30) = 5.175$, $P < 0.05$). The infants showed different patterns of looking on the test trials of the two conditions.

5.3. Discussion

Experiment 4 contrasted words with emotional expressions that are both intentional and are produced by the human vocal tract. The results showed that words facilitated object individuation whereas emotional expressions did not. Emotional expressions differ from words in that they are not referential, e.g. I may think that snakes are disgusting and sneer at them whereas other people may think that they are adorable and coo at them. But all of us would still use the same word and call these creatures “snakes”.

6. General discussion

Previous studies found that it is not until about 12 months of age that infants succeed in using sortal/object kind differences between, say, a duck and a ball, or a cup and a shoe, to establish two distinct objects behind an occluder in the object individuation task of Xu and Carey (1996). In the current studies, we found that when 9-month-old infants were given two distinct labels they succeeded in the same object individuation task. In all four experiments, the presence of two distinct labels led to success in object individuation, both with familiar objects and words (Experiments 1 and 2) and with unfamiliar objects and words (Experiments 3 and 4). Furthermore, the presence of a single label for both objects (Experiment 1), two tones (Experiment 2), two distinct sounds (Experiment 3), or two distinct emotional expressions (Experiment 4) did not help infants succeed on this task.

Several aspects of these data are important to note. First, because familiarity with the objects or the labels was not necessary, we suggest that it is the presence of distinct labels per se that allowed the infants to establish two distinct individual objects. This is supported by the finding that the presence of a single label did not lead to success. That is, it was not the mere presence of words that facilitated object individuation, perhaps via heightened attention to the objects. Second, these facilitation effects may be language-specific since in Experiments 2–4, other non-speech sounds (two tones, two very distinct sounds, or two emotional expressions) did not lead to success on this task. Although we have not exhausted all other non-speech sounds, the current findings are suggestive that language may play a special role in facilitating object individuation.

How do these results bear on the question of what role language may play in conceptual development? In particular, what is the role of language in these studies and how might labeling help with the construction of sortal/kind concepts? I will consider four possibilities, from the weakest to the strongest, and discuss each one in turn.

The first, and weakest, possibility is that language does not facilitate the acquisition of sortal/object kind concepts in infants per se, but it does come to the rescue of infants' limited memory and information-processing capacities. This is a possibility that should be taken seriously if the Wilcox and Baillargeon (1998a,b) studies reflected the use of sortal/object kind information. On their view, if the experimental task is sufficiently simplified, infants as young as 4.5 months are able to use property/featural information for object individuation *under limited circumstances*. If such early competence reflects the use of sortal/object kind information and not merely property/featural information, how would we then interpret the current results? A plausible candidate is that when the tasks are complex and when infants experience information overload, the presence of labels may allow them to hold the representations of two objects in mind and retrieve them later on in order to succeed on the task. In other words, the labels function as "summary representations" or mnemonics for the infants (Needham & Baillargeon, 2000; Wilcox & Baillargeon, 1998a). However, there are alternative interpretations of these studies and it remains an open question whether the younger infants used property or kind information to succeed in these simplified tasks (see Xu, in press, for a detailed discussion). In addition, preliminary evidence suggests that when perceptual information is pitted against linguistic information in an object individuation task, e.g. a single object is labeled by two distinct words or two distinct objects are labeled by a single word, 9-month-old infants rely on the linguistic information to infer the number of objects behind the occluder (Xu, 2002). These findings suggest that perhaps labels are not simply mnemonics that allow infants to remember and retrieve their object representations.

A second possibility is that distinct labels highlight the property/featural differences between objects. Once infants have taken further notice of these featural differences, they are more likely to conclude that there are two distinct objects behind the occluder. This is a stronger possibility than the first one because it allows language to influence infants' object representations: it is the presence of labels that draws the infants' attention to the object features and prompts them to further analyze these differences before drawing any conclusions about the number of objects in an event. One aspect of the data argues against this possibility, namely that the presence of two contrastive labels and a single label for both objects increased looking times to the objects by the same amount (in Experiment 1) compared to the silent trials when no label was provided. If two distinct labels simply increased the infants' attention to the objects, we would have predicted that the presence of two labels would have had a stronger attention-getting effect than the presence of a single label. That is, we would have expected a larger increase in looking time in the two-word condition than in the one-word condition when we compared the silent and the labeled trials. This was not the case. However, it might be argued that even though the increase of attention was the same in the two-word and the one-word conditions, how the infants used the information differed in the two conditions: the infants in the two-word condition used the extra looking time to discern the differences between the objects whereas the infants in the one-word condition used the extra time to analyze the sameness in object features. This is an empirical question that needs further research.

A third, and stronger, possibility is that words are "essence placeholders". Infants may expect that words for objects map onto distinct kinds in their environment. Given this expectation, the very fact that one object is called "a duck" and one object seen on a different occasion is called "a ball" is sufficient evidence for infants to posit two distinct kinds or essences. According to psychological essentialism (Gelman, in press; Medin & Ortony, 1989), essences determine the surface features and properties of objects. If two kinds of objects are indeed behind an occluder, it follows that there must be two distinct tokens of objects. This may be a mechanism by which infants first establish what *kinds of things* are in their environment and how these object kind concepts can play a role in object individuation. The evidence so far, however, is consistent with the view that words pick out distinct object *tokens* for infants as opposed to object *kinds*. That is, in the presence of labels, do infants expect two kinds of objects behind the occluder or just two individual objects? Further studies may address the issue. If it turns out that infants expect two kinds of objects, I suggest that words serve as "essence placeholders". Simply hearing and remembering words for object kinds does not give an infant or anyone else fully-fledged concepts of dog or chair, but words such as "dog" and "chair" may direct the child to set up "placeholders" for the relevant concepts and through interacting with the world, these concepts are elaborated and beliefs about these concepts are cumulated.

The last, perhaps also the strongest, possibility is that language may exert some influence on cognitive architecture. Following the Hermer-Vazquez et al. (1999) proposal that language may serve as the vehicle for conjoining distinct parts of cognitive architecture, Xu (1999) hypothesized that learning names for things may allow infants to conjoin two aspects of object representations. Young infants, like adults, may have two distinct visual pathways for encoding object motion and location information (roughly speaking, the 'where' pathway) and for encoding object features (roughly speaking, the 'what' pathway) (see also Leslie, Xu, Tremoulet, & Scholl, 1998; Scholl & Leslie, 1999, for related proposals). Early on in infancy, these pathways are largely separated and there is little connection between the two. When infants first begin to learn words for objects such as "a cup" or "a ball", two independently motivated word learning constraints have to be satisfied: the whole object constraint and the taxonomic constraint (Markman, 1989). The whole object constraint says that words refer to whole objects as opposed to parts of objects; the taxonomic constraint says that words should be generalized to members of the same kind. These constraints map onto the two aspects of object representations encoding "where" and "what" information since the former picks out whole objects and the latter provides a similarity metric that correlates with kind membership. In other words, in order to satisfy these constraints on word meaning, the infants would have to pay attention to both the location information and featural information. This conjecture also provides a way of connecting the literature on infant categorization (e.g. Eimas and Quinn, 1994; Quinn, Eimas, & Rosenkrantz, 1993) and the literature on object individuation (e.g. Xu & Carey, 1996). What the infant categorization literature may be tapping into are the characteristics of the 'what' system but this system is largely independent of the system that establishes distinct objects using spatiotemporal information. Object individuation, however, requires both the 'what' and the 'where' systems. Clearly, this proposal needs further empirical investigation (see Bonatti, Frot, Zangl, & Mehler, in press, for empirical evidence that the sortal/kind *human* does not need language to conjoin 'what' and 'where' information).

If any of the last three possibilities is correct, we may have in our hands a case of how *universal aspects of language* may influence conceptual development. Object kinds are largely universal (with the exception of artifact kinds) and words for object kinds are present in all languages, thus we may be studying a universal process by which children learning different languages acquire kind concepts. That is, a universal mapping between words for objects (or count nouns) and object kinds (or kinds of individuals in general, see Bloom, 1996, 2000) may underpin the earlier success by the 9-month-old infants in the present studies. We could also go on to ask questions about whether the absence of words for kind concepts (e.g. deaf infants who might be delayed in learning these words) hinders conceptual development and whether any cross-linguistic differences exist in this domain.

Several questions remain. Is language the only mechanism for acquiring sortal/kind concepts for human infants? Is language necessary for representations of sortals/kinds? Are words special in some sense in facilitating the process of acquiring sortal/kind concepts? I will discuss each of these questions in turn.

Is language the only mechanism for acquiring sortal/object kind concepts for human infants? Perhaps not, although we lack direct empirical evidence. For example, human infants may use functional information as a source of evidence that two objects are of different kinds – if a person can drink out of one object and use a second object as a hammer, an infant may conclude that there are two distinct objects with distinct functions. Infants may also use simple correlated features in object individuation. For example, if an object is red, square, and makes a "clunk-clunk" sound (internally generated) as it bounces across the floor and another object is blue, spherical, and makes a beeping sound as it rolls across the floor, the clusters of features may lead the infant to conclude that it is highly unlikely that there is only a single object involved.

A related question is whether non-human primates are capable of representing object kinds as we do. Recent studies with non-human primates suggest that language is perhaps not necessary for acquiring object kind concepts. Uller, Xu, Carey, and Hauser (1997) found that rhesus monkeys passed the Xu and Carey task with a carrot and a piece of squash. However, these data are still open to alternative interpretations, e.g. perhaps the monkeys were keeping track of food substances, carrot stuff vs. squash stuff, as opposed to individual food items, a carrot vs. a piece of squash. On this alternative they would not find an outcome of a single piece of food with some carrot stuff and some squash stuff unexpected. More recently, Santos, Sulkowski, Spaepen, and Hauser (2001) replicated the Uller et al. finding with a manual search paradigm. In a series of experiments, rhesus monkeys were shown, say a grape, being hidden in a box. They were then allowed to approach the box to retrieve the food item. On a consistent trial, they retrieved the grape; on an inconsistent trial, they retrieved a different food item, say, a chunk of coconut meat. The dependent measure was whether the monkeys would search the box again. More persistent search on the inconsistent trials was interpreted as success in using property/kind information for individuating food items. Santos et al. found that rhesus monkeys were able to use property/kind information for object individuation. To address the question of whether it was property or kind information that underlies the successes, Santos et al. contrasted shape with color using food items (e.g. a triangle vs. a disc, or a white vs. a blue piece of fruit). They interpreted their results as showing that rhesus macaques use color but not shape differences in individuating food items. They further argued that this reflected representations of food kinds because color, but not shape, tends to indicate change of kind membership in the domain of food. However, a closer look at the results suggests that the findings might not be as clear-cut. In Experiment 4 of Santos et al., there was no interaction between the shape and the color conditions, i.e. both groups searched longer on the violation trials (e.g. see disc then retrieve triangle) than the consistent trials (e.g. see disc then retrieve disc). Their interpretation was based on separate *t*-tests showing a statistically significant difference for the color but not the shape condition. Perhaps more subjects need to be tested to obtain a clearer result in the shape condition before any strong conclusions are drawn regarding whether the rhesus monkeys used kind distinctions for object individuation. In addition, size contrast apparently leads to successful individuation as well (Santos, pers. commun.). Thus, it is possible that adult rhesus monkeys are able to use all types of property contrasts in the service of object individuation.

Lastly, are words special in facilitating object individuation? Recently, Bloom and his colleagues (Bloom, 2000; Diesendruck & Markson, in press; Markson & Bloom, 1997) have examined whether words are special in the context of discussing the nature of word learning constraints (e.g. Markman, 1989). These studies contrasted words with facts conveyed verbally to see whether both types of information can be fast-mapped, and whether mutual exclusivity, a proposed word learning constraint, can be found with non-word stimuli. Three- and 4-year-old children were taught a new word or a new fact about an object and then recall was assessed ("Can you find the toma?" or "Can you find the one my uncle gave me?") immediately after, a week later, or a month later (Markson & Bloom, 1997). With both words and facts, children were able to choose the correct object well above chance under all conditions. Markson and Bloom argued that word learning does not require special, dedicated mechanisms; it recruits other abilities that are already present for other purposes, including general learning and memory abilities (see Bloom, 2000; Tenenbaum & Xu, 2000; Xu & Tenenbaum, in press). Similarly, Diesendruck and Markson (in press) found that mutual exclusivity can be demonstrated with both words and facts. For example, in the fact condition, if there are two unfamiliar objects present and one of them is described as given to the experimenter by his uncle, the child would choose the other object when asked to find the object that was given to the experimenter by his sister. In the word condition, if one object is labeled "a blicket", the child would choose the other object when asked to find "the fendle". They suggest that mutual exclusivity is a pragmatic constraint as opposed to a constraint on word meaning per se. The general view advocated by these researchers is that word learning is not special and words are not special since parallel phenomena can be found with non-word stimuli. How do we put together these findings with the current results in which words seem to play a special role in facilitating object individuation (while other auditory stimuli do not)?

The conceptual issues addressed by these two lines of research are clearly related. Several differences between these studies, however, make the comparison less direct and I will raise some possibilities for further research. First, the ages of the children studied are very different. Bloom and colleagues' studies have examined 3- and 4-year-old children, who are proficient word learners, whereas the current studies tested 9-month-old infants who are not yet speaking. Second, the tasks used in the studies were drastically different. The experiments reviewed above by Bloom and colleagues focused on whether words or facts can be remembered long-term after brief exposures and whether word learning constraints are best characterized as pragmatic principles. The current studies do not address the issue of whether the infants learned the words at all, but rather whether the presence of words may influence non-linguistic concept acquisition. It is an open question whether the infants in the current studies have learned the specific words and whether they can remember them long-term. Future studies may also ask whether other verbal contrasts (e.g. "This is pretty" vs. "This is nice") would facilitate object individuation.

Although language might not be the only mechanism for acquiring sortal/kind concepts and non-human primates may have at least some ability to represent kinds, it is nonetheless of interest that different aspects of language learning may shape children's conceptual representations in important ways. The current findings suggest a role of language in the acquisition of sortal/object kind concepts in infancy: words in the form of labeling may serve as "essence placeholders". We raised several possibilities for how labeling may influence conceptual representations in infants who are acquiring their first words.

Acknowledgements

I thank Allison Baker, Paul Bloom, Susan Carey, Dedre Gentner, Marc Hauser, Lori Markson, Cristina Sorrentino, Elizabeth Spelke, Joshua Tenenbaum, and three anonymous reviewers for helpful discussion and/or comments on earlier drafts. Thanks to the members of the MIT and Northeastern University Infant Cognition Laboratories for their help in data collection. This research was supported by NIH B/START grant (R03MH59040-02) and NSF grant (BCS-9910729). Experiment 1 was partially conducted in Spelke's laboratory and was supported by NIH grant (R37-HD23103).

References

Balaban, M., & Waxman, S. (1996). Words may facilitate categorization in 9-month-old infants. *Journal of Experimental Child Psychology*, *64*, 3–26.

Baldwin, D. A., & Markman, E. M. (1989). Establishing word-object relations: a first step. *Child Development*, *60*, 381–398.

Bloom, P. (1996). Possible individuals in language and cognition. *Current Directions in Psychological Science*, *5*, 90–94.

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Bloom, P., & Keil, F. C. (2001). Thinking through language. *Mind and Language*, *16*, 351–367.

Bonatti, L., Frot, E., Zangl, R., & Mehler, J. (in press). The human first hypothesis: identification of conspecifics and individuation of objects in the young infant. *Cognitive Psychology*.

Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology*, *43*, 1–22.

Brown, R. (1976). Reference: in memorial tribute to Eric Lenneberg. *Cognition*, *4*, 125–153.

Carey, S. (1994). Does learning a language require the child to reconceptualize the world? *Lingua*, *92*, 143–167.

de Villiers, J. G., & de Villiers, P. A. (2000). Linguistic determinism and the understanding of false beliefs. In P. Mitchell & K. Riggs (Eds.), *Children's reasoning and the mind* (pp. 189–226). Hove: Psychology Press.

Diesendruck, G., & Markson, L. (in press). Mutual exclusivity: a pragmatic constraint. *Developmental Psychology*.

Eimas, P., & Quinn, P. (1994). Studies on the formation of perceptually-based basic-level categories in young infants. *Child Development*, *65*, 903–917.

Fenson, L., Dale, P., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., Pethick, S., & Reilly, J. S. (1993). *The MacArthur Communicative Development Inventories: users guide and technical manual*. San Diego, CA: Singular Press.

Fodor, J. A. (1975). *The language of thought*, Cambridge, MA: MIT Press.

Gelman, S. (in press). *The essential child: origins of essentialism in everyday thought*. Oxford: Oxford University Press.

Gentner, D., & Boroditsky, L. (2001). Individuation, relational relativity and early word learning. In M. Bowerman & S. Levinson, *Language acquisition and conceptual development*. Cambridge: Cambridge University Press.

Gopnik, A., & Meltzoff, A. (1997). *Words, thoughts and theories*. Cambridge, MA: MIT Press.

Gumperz, J. J., & Levinson, S. S. (1996). *Rethinking linguistic relativity*. Cambridge: Cambridge University Press.

Gupta, A. (1980). *The logic of common nouns*. New Haven, CT: Yale University Press.

Hale, C. M., & Tager-Flusberg, H. (2000). The influence of language on theory of mind: a training study, submitted for publication.

Heider, E. (1972). Universals in color naming and memory. *Journal of Experimental Psychology*, *93*, 10–20.

Hermer, L., & Spelke, E. S. (1997). Modularity and development: the case of spatial reorientation. *Cognition*, *61*, 195–232.

Hermer-Vazquez, L., Spelke, E. S., & Katsnelson, A. S. (1999). Sources of flexibility in human cognition: dual-task studies of space and language. *Cognitive Psychology*, *39*, 3–36.

Hirsch, E. (1982). *The concept of identity*. New York: Oxford University Press.

Hunt, E., & Agnoli, F. (1991). The Whorfian hypothesis: a cognitive psychology perspective. *Psychological Review*, *98*, 377–389.

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

Leslie, A., Xu, F., Tremoulet, P., & Scholl, B. (1998). Indexing and the object concept: developing 'what' and 'where' systems. *Trends in Cognitive Sciences*, *2*, 10–18.

Levinson, S. (1996). Frames of reference and Molyneux's question: crosslinguistic evidence. In P. Bloom & M. Peterson (Eds.), *Language and space* (pp. 109–169). Cambridge, MA: MIT Press.

Li, P., & Gleitman, L. (2002). Turning the tables: language and spatial reasoning. *Cognition*, *83*, 265–294.

Lucy, J. (1992). *Grammatical categories and cognition: a case study of the linguistic relativity hypothesis*, Cambridge: Cambridge University Press.

Macnamara, J. (1986). *A border dispute: the place of logic in psychology*. Cambridge, MA: MIT Press.

Macnamara, J., & Reyes, G. E. (1994). *The logical foundations of cognition*. Oxford: Oxford University Press.

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

Markson, L., & Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature*, *385*, 813–815.

Medin, D., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). New York: Cambridge University Press.

Needham, A. (1998). Infants' use of featural information in the segregation of stationary objects. *Infant Behavior and Development*, *21*, 47–76.

Needham, A., & Baillargeon, R. (1997). Object segregation in 8-month-old infants. *Cognition*, *62*, 121–149.

Needham, A., & Baillargeon, R. (1998). Effects of prior experience on 4.5-month-old infants' object segregation. *Infant Behavior and Development*, *21*, 1–24.

Needham, A., & Baillargeon, R. (2000). Infants' use of featural and experiential information in segregating and individuating objects: a reply to Xu, Carey, and Welch (1999). *Cognition*, *74*, 255–284.

Nelson, K. (1996). *Language in cognitive development: emergence of the mediated mind*. Cambridge: Cambridge University Press.

Pederson, E., Danziger, E., Wilkins, D., Levinson, S., Kita, S., & Senft, G. (1998). Semantic typology and spatial conceptualization. *Language*, *74*, 557–588.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Pinker, S. (1994). *The language instinct*, New York: William Morrow.

Pinto, J. (1995). *Xhab. Program written for looking time experiments*. Palo Alto, CA: Author.

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

Quinn, P. C., Eimas, P., & Rosenkrantz, S. L. (1993). Evidence for representations of perceptually similar natural categories by three- and four-month-old infants. *Perception*, *22* (4), 463–475.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*, 192–233.

Rosch, E. (1978). Principles of categorization. In R. Rosch & B. B. Lloyd (Eds.) *Cognition and categorization*-Hillsdale, NJ: Erlbaum.

Santos, L. R., Sulkowski, G. M., Spaepen, G. M., & Hauser, M. D. (2001). Object individuation using property/kind information in rhesus macaques, submitted for publication.

Scholl, B. J., & Leslie, A. M. (1999). Explaining the infant's object concept: beyond the perception/cognition dichotomy. In E. Lepore & Z. Pylyshyn, *What is cognitive science?* (pp. 26–73). Oxford: Blackwell.

Sera, M. D., Bales, D. W., & Del Castillo Pintado, J. (1997). *Ser* helps Spanish speakers identify "real" properties. *Child Development*, *68*, 820–831.

Slobin, D. (1996). From "thought and language" to "thinking for speaking". In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge: Cambridge University Press.

Soja, 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.

Spelke, E. S., & Kestenbaum, R. (1986). Les origines du concept d'objet. *Psychologie Francaise*, *31*, 67–72.

Spelke, E. S., Kestenbaum, R., Simons, D. J., & Wein, D. (1995). Spatio-temporal continuity, smoothness of motion and object identity in infancy. *British Journal of Developmental Psychology*, *13*, 113–142.

Spelke, E. S., & Tsivkin, S. (2001). Language and number: a bilingual training study. *Cognition*, *78*, 45–88.

Tenenbaum, J. B., & Xu, F. (2000). Word learning as Bayesian inference. In L. Gleitman & A. Joshi (Es.), *Proceedings of the 22nd Cognitive Science Society conference* (pp. 371–376). Hillsdale, NJ: Erlbaum.

Tincoff, R., & Jusczyk, P. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*, 172–175.

Uller, C. M., Xu, F., Carey, S., & Hauser, M. (1997). Is language needed for constructing sortal concepts? A study with non-human primates. In E. Hughes, M. Hughes & A. Greenhill (Eds.), *Proceedings of the 21st annual Boston University conference on language development* (pp. 665–677). Somerville, MA: Cascadilla Press.

Van de Walle, G., Carey, S., & Prevor, M. (2001). Bases for object individuation in infancy: evidence from manual search. *Journal of Cognition and Development*, *1*, 249–280.

Vygotsky, L. (1962). *Thought and language*. Cambridge, MA: MIT Press.

Waxman, S. R. (1999). Specifying the scope of 13-month-olds' expectations for novel words. *Cognition*, *70*, B35–B50.

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

Whorf, B. (1956). *Language, thought, and reality*. Cambridge, MA: MIT Press.

Wiggins, D. (1980). *Sameness and substance*. Oxford: Basil Blackwell.

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

Wilcox, T., & Baillargeon, R. (1998). Object individuation in infancy: the use of featural information in reasoning about occlusion events. *Cognitive Psychology*, *37*, 97–155.

Wilcox, T., & Baillargeon, R. (1998). Object individuation in young infants: further evidence with an event-monitoring paradigm. *Developmental Science*, *1*, 127–142.

Wilcox, T., & Schweinle, A. (2002). Object individuation and event mapping: developmental changes in infants' use of featural information. *Developmental Science*, *5*, 132–150.

Xu, F. (1997). From Lot's wife to a pillar of salt: evidence that physical object is a sortal concept. *Mind and Language*, *12*, 365–392.

Xu, F. (1999). Object individuation and object identity in infancy: the role of spatiotemporal information, object property information, and language. *Acta Psychologica*, *102*, 113–136.

Xu, F. (2002). The effects of labeling on object individuation in 9-month-old infants, submitted for publication.

Xu, F. (in press). The development of object individuation in infancy. In J. Fagen, & H. Hayne (Eds.), *Progress in infancy research* (Vol. 3). Mahwah, NJ: Lawrence Erlbaum.

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

Xu, F., & Carey, S. (2000). The emergence of kind concepts: a rejoinder to Needham & Baillargeon. *Cognition*, *74*, 285–301.

Xu, F., Carey, S., & Quint, N. (2002). Kind-based object individuation in infancy, submitted for publication.

Xu, F., Carey, S., & Welch, J. (1999). Infants' ability to use object kind information for object individuation. *Cognition*, *70*, 137–166.

Xu, F., & Spelke, E. S. (2001). Large number discrimination in 6-month-old infants. *Cognition*, *74*, B1–B11.

Xu, F., & Tenenbaum, J. B. (in press). Rational statistical inference: a critical component for word learning. *Brain and Behavioral Sciences*.

Yoshida, H., & Smith, L. B. (in press). Shifting ontological boundaries: how Japanese- and English-speaking children generalize names for animals and artifacts. *Developmental Science*.