Developmental Psychology
1994, Vol. 30, No. 5, 639-650

0012-1649/94/$3.00

Learning Words in Nonostensive Contexts

Michael Tomasello and Michelle Barton

Four word learning studies with 24-month-old children are reported. In Studies 1 and 2, an adult used a novel word to announce her intention to perform an action or to find an object. It was found that a knowledge of what action or object was impending—established through scripted events before the word's introduction—was not necessary for children to learn the words. Studies 3 and 4 focused on what word learning cues children might be using in these contexts. In Study 3, it was found that children learned a novel verb for an intentional and not an accidental action. In Study 4, it was found that children learned a novel noun for an object the adult was searching for, not ones she had rejected while searching. Because none of the best-known constraints on lexical acquisition could have helped them in these contexts, it was concluded that children were relying on social-pragmatic cues to learn the new words.

Western middle-class parents often point to and name objects for young children in what is known as the *ostensive naming context* or the *original word game* (Brown, 1958). There is a widespread assumption among lexical acquisition researchers that this learning context is representative of lexical acquisition in general. Virtually all of the theoretical analyses and empirical research in the constraints approach to lexical acquisition, for example, address noun learning in the ostensive context (Markman, 1989). Similarly, the otherwise very different analyses and research of social-pragmatic theorists concerned with issues of joint attention and cultural learning also focus almost exclusively on this same object-naming game (Bruner, 1983; Tomasello, 1988). This narrow focus is unfortunate, however, because young children learn other types of words from early in their linguistic careers, and they are frequently exposed to new words of all types in contexts other than the ostensive context.

There is little empirical research on young children's acquisition of words in nonostensive contexts. With regard to object labels, the only research to date suggests that young children do not readily learn new object labels in nonostensive contexts. Whitehurst, Kedesdy, and White (1982) found much better object label learning by 2-year-olds in a condition in which word and object were presented simultaneously than in conditions in which there was a temporal discrepancy between presentation of word and object (see also Kuczaj, Carter, Sherman, & Borys, 1989). The problem, however, is that the instantiations of nonostensive contexts in this study were extremely artificial and contained conflicting nonverbal and discourse cues as to the new word's referent; for example, in the *past* condition the adult put away an object, waited 10 s, and then said, "That was a wick." In Baldwin (1991), an adult looked at and named an object that was concealed from the child in a bucket while the child was looking at another object. The 16–19-month-olds of this study did not assume that the object they were looking at was being named, but neither did they show strong evidence of having learned the name of the object in the bucket. It might be once again, however, that children did not learn the object name in this nonostensive context because there was an artificiality to the social situation in that the adult named the hidden object on four different occasions before finally showing it to the child. (Note that in the study of Baldwin, 1993b, in which children clearly did learn the new word, the named object was perceptually available to the child throughout.)

There is also little research on nonostensive learning contexts for word types other than object labels. Tomasello (1992a) reported that some of his daughter's early verbs seemed to be learned exclusively in nonostensive contexts such as when the adult requested child action. The suggestion was that the child learned new words in such contexts by understanding adult nonverbal behaviors indicating semantic intentions, for example, a parent pushing the child toward the car while saying "Let's go!" (see also Akhtar, Dunham, & Dunham, 1991). This suggestion was based solely on parental reports of probable learning contexts. Tomasello and Kruger (1992, Study 1), however, observed a relatively large sample of Western middle-class mothers in interaction with their young children and documented systematically the contexts in which mothers used verbs. They found that nonostensive contexts were in fact the norm: Mothers seldom used verbs to name ongoing actions in an ostensive manner for their young children (overall less than 30% of the time). Much more often mothers used verbs to request behaviors of the child, to anticipate either the child's or their own impending action, or to comment on the child's or their own just-completed action. Of most importance in this study was the *impending action* context in which the mother either requested or anticipated the child's action. This was not only the most frequent model type (approximately 60% of all maternal verb models) but was also the model type that led to the best child learning as measured correlationally and in terms of the child's responsiveness in the discourse context.

Michael Tomasello and Michelle Barton, Department of Psychology, Emory University.

This research was supported in part by a grant to Michael Tomasello from the Spencer Foundation. We thank Andrea Horowitz and Nameera Akhtar for help with data collection, and we also thank the mothers and children who participated in the study.

Correspondence concerning this article should be addressed to Michael Tomasello, Department of Psychology, Emory University, Atlanta, Georgia 30322.

The finding of superior child learning in the impending action context was further confirmed by a lexical training study in which 24-month-old children were taught a novel verb for a novel action (Tomasello & Kruger, 1992, Study 2). On 10 occasions an adult placed a doll on wheels at the top of a ramp, pushed a button beside the doll, and watched the doll roll down the ramp and through a hole. This action was referred to with the transitive verb *plunk* (which no children knew before the study). In the *impending action* condition, the adult announced as she placed the doll on the top of the ramp, "Look, Jason, I'll plunk the man," with the action occurring a few seconds later. In the *ongoing action* condition, she pushed the button and then said, "Look, Jason, I'm plunking the man," as the doll was rolling down the ramp. In the *completed action* condition, she waited until the doll was in its final resting place and then said, "Look, Jason, I plunked the man." Despite the presence of morphological cues in the ongoing and completed conditions (*-ing* and *-ed*), there was strong evidence in both production and comprehension that children learned novel verbs best in the impending action context.

The results of these studies raise several important questions about children's learning of words in nonostensive contexts, especially the impending action context. First, if the impending action context facilitates verb learning, at least two mechanisms for the effect can be imagined. On the one hand, it is possible that children only learn new verbs in this context when they know from previous experience (or some other cue) what action is impending; that is, the process is one in which the child perceives some cue (e.g., the doll at the top of the ramp) that indicates to him or her what action is to come. At the time the word is actually said, therefore, the action, while not physically present, is at least mentally present in the form of the child's expectation. On the other hand, it is also possible that 2-year-olds do not have to have such an expectation to learn a verb in the impending action context; that is, it may be that the process is one in which children keep the word in mind until they perceive the unanticipated action. If this latter mechanism is at work, it would mean that in this context young children are able to connect the new word with the action referent by relying on something other than a current knowledge of the action that is to come. The question is thus: To acquire a novel verb in the impending action context, do young children need to have specific knowledge of the action that is impending?

A second question concerns whether young children will also learn object labels in something like the impending action context, for example, in a finding game (which we may call the *impending object* context); if they do, it is also of interest which of the two previously outlined mechanisms might be at work. The Whitehurst et al. (1982) study suggests that perhaps young children will not acquire object labels in nonostensive contexts at all. But the impending object context as we conceive it is sufficiently different from the nonostensive contexts used in that study that this is still an open question. In the impending object context, the temporal discrepancy between word and referent is much briefer and the continuity of discourse context much greater, leading to the kind of naturalness that often makes a difference in young children's learning. The question is thus: Can young children acquire a novel object label in the impending object context, and, if so, do they need to have specific knowledge of the object whose presence is impending?

Returning to the impending action context, a third question concerns specifically what kinds of cues in adult behavior, if any, children are using to connect verb and action referent in this context. One possibility is that the children do not use any such cues. In the Tomasello and Kruger (1992) experimental study, after the adult announced her intention to plunk the toy, she then turned to the appropriate place and proceeded to perform the action. It might be, then, that children were simply following the relatively simple procedure of assuming that the action immediately following the new word was its appropriate action referent. On the other hand, it might be that the children were actively monitoring the adult's intentions in the experimental situation to discover the intended referent of the new word; they might have been using, for example, indications of intentionality involving such things as the adult's facial expression and body posture. These two possibilities might compete in cases in which temporal contiguity and indications of adult intentionality conflict, for example, when the adult announces an intention to perform a specific action and then engages in either preparatory or accidental actions before the intended action. The question is thus: When an adult uses a novel verb in the impending action context, will young children automatically assume that she is referring to the next action performed, or will they use adult intention cues to identify the action referent even if that referent is preceded by other actions?

A fourth and final question concerns whether this same basic process applies to noun learning. One indication that it does is provided by Baldwin (1993a), who found that 20-month-old children do not assume that the first object they see after the adult utters a novel noun is its intended referent. Thus, if the adult looks in a bucket and says "That's a *toma*" and then hands the child an object from another bucket before later giving her the toma, children will correctly learn the new word for the toma. As in Baldwin's other studies, children are presumably relying on the adult's gaze direction as a cue to referential intentions. It is possible, however, that young children are able to use other types of cues to make this same determination, for example, in a finding game, the adult expression of satisfaction that the sought object has been found (in which gaze direction toward unsought objects may even be a misleading cue). If it turns out that children are able to use multiple cues to determine which object an adult intends to indicate, the implication is that cues such as gaze direction are not highly specific indexes learned by a passive association, but rather that children are learning in a more flexible manner about adults and what they are doing with new pieces of language. The question is thus: When an adult uses a novel noun in the impending object context, in which gaze direction is not a helpful cue, will young children automatically assume that she is referring to the next object perceptually available, or will they use adult intention cues to identify the object referent even if that referent is preceded by other objects?

The present studies represent an attempt to address these four questions experimentally. In a series of four studies, we choreographed situations in which 2-year-old children were exposed to novel verbs and nouns. In the first two studies, we manipulated whether children knew what action or object might

be coming immediately after a novel word was uttered in the impending action or impending object context. In the third and fourth studies, an adult used a novel verb or noun to announce her intention to perform an action or find an object, then interjected an irrelevant action or object (indicating nonverbally that it was not the one intended) before performing the intended action or finding the intended object (indicating nonverbally that it was the one intended). In all cases, our attempt was to use patterns of learning and nonlearning to infer something about what cues in the social-pragmatic context 2-year-old children use in their acquisition of novel lexical items.

Study 1

The major question of this study is whether it is necessary for the child to know what action is coming in an impending action context to learn and use a novel verb for that impending action.

Method

Subjects. Forty 2-year-old children (M age = 2 years 0 months, range = 1 year 10 months to 2 years 2 months) participated as subjects. Children were recruited through a psychology department subjects file composed of parents who had voluntarily responded to a letter soliciting cooperation for studies in child development. Parents of each child were contacted by phone and were invited to participate if the child was currently showing productive language use, with more than one verb in his or her vocabulary. A brief inventory of common first verbs, including the target verb, was also conducted over the phone to exclude any child who already knew the target verb (no potential subjects did).

Materials. A set of giant waffle blocks was used to construct a play center with four activity stations. A set of five Sesame Street characters was used to perform the four actions. The action stations were (a) a tape measure that pulled in a ring (in which a character sat), (b) a basket (in which a character sat) that hoisted like a pulley, (c) a curved platform that launched a character into the air when hit on one side, and (d) a curved chute to drop a character through. The target action was always plunking a character down the curved chute. No verbs were ever used to describe any of the other three actions but only generic descriptions to "Look at this!" or "Watch!"

Experimental design. The design of the study was as follows. The model of the target verb *plunk* was manipulated along two dimensions: (a) the timing of the language model in relation to performing the action (impending vs. ostensive) and (b) the predictability of the upcoming action (known vs. unknown).

With regard to timing, in the ostensive condition, an experimenter (henceforth called E) gave the model "I'm (You're) *plunking* Big Bird. I'm (You're) *plunking* him" at the same time that she or the child dropped the character into the chute. In the impending condition, E held a character away from the action station, faced the child, and gave the language model "I'm going to (Can you) *plunk* Big Bird. I'm going to (Can you) *plunk* him" and then (after a 1- to 2-s delay) did the action or let the child do it. Thus for each ostensive or impending model, E used the target verb twice and paired it with one demonstration of the target action. These two conditions replicate the method of Tomasello and Kruger (1992), but with a more expeditious action (i.e., the action down the chute was very quick, whereas the rolling down the ramp of the previous study took several seconds). With regard to predictability, what varied was whether, on the basis of past experience, the child could predict which action should come next (known vs. unknown). We manipulated this, as described later, by setting up specific expectations with specific objects through games before and during the experimental models with the target verb.

These two variables were then crossed to form four experimental conditions: impending–known, impending–unknown, ostensive–known, and ostensive–unknown. Children were randomly assigned to one of the four conditions (n = 10 per condition) such that there were 5 girls and 5 boys in each group (and 2 or 3 firstborns per gender grouping).

Training procedure. Subjects came to a psychology laboratory for two visits, both during the same week. For both sessions, each child played with a female E, who was unaware of the hypotheses of the study, for approximately 30 min. All procedures were identical for a given child on both days.

Both sessions were videotaped by an observer (henceforth called O), who also kept a written record of the child's and E's actions and utterances during the modeling and testing phases of the study. On the first day, after an initial warm-up period, E began with a demonstration of each action station without any language models. During these demonstrations E used neutral language (e.g., "Watch this!") or invited the child to do the action (e.g., "Can you do this?", showed the child how to do it, and gave the child a chance to do it too). Then began a series of four rounds of action demonstrations and language models. The order of the actions and characters was rotated as specified according to condition. In the *unknown* conditions, all characters were rotated through all action stations in a random fashion throughout the four rounds; there was thus no way to know when E picked up a character what might be done with it. In the *known* condition, on the other hand, Big Bird was always paired with plunking and never with any of the other actions, whereas the other characters were randomly paired with actions other than plunking. Over time it thus became apparent that whenever E picked up Big Bird the action would be plunking. Each round ended with two verb models—that is, two performances of the plunking action, each of which was paired with two utterances of the word *plunk*—appropriate for that condition (ostensive or impending model as described earlier). Thus, at the end of the four rounds, a subject had observed the plunking action paired with the word *plunk* eight times (16 tokens of the target verb). Models were repeated if E and O both agreed that the child was not paying attention during the model.

The entire procedure, including the testing (next section), was repeated on a 2nd day for a total across the 2 days of 16 language models (32 tokens of *plunk* said in 16 pairs, each pair associated once with the plunking action).

Testing procedure. The observer, O, kept track of the completion of models and signaled E when the last model had been finished for a given day. She also recorded any spontaneous productions of the word *plunk* and contextual notes sufficient to tell what action it was used to indicate. Only if the child was performing the action, about to perform the action, or watching E perform the action when she uttered the target word was it considered an example of correct production. Use of the target word was checked by a second observer using the videotapes with almost perfect agreement (96%).

Following the completion of Round 4 on each day, the children were allowed a few minutes of free play with the toys. Then elicited production and comprehension tests were conducted. For elicited production, for the children in the known condition, E plunked Big Bird and asked, "What am I doing to Big Bird?" or "What's Big Bird doing?" In the unknown condition, E plunked one of the other characters, each time asking, "What am I doing to _____?" or "What am I doing to him?" If the child gave a verbal response, it was recorded by O, and E went on to the comprehension test (after a few minutes of intervening free play). If the child was unresponsive, the probe was repeated (up to three times) to encourage the child to answer. For testing comprehension in both conditions, E introduced a new character (Cookie Monster) and allowed the child to play with it for a few minutes. To test comprehension of the target verb, E then handed Cookie Monster to the child and asked, "Can you plunk Cookie Monster? Go plunk him." In making this request E looked directly at the child, not the action stations (still in the same

Table 1
Number of Children Who Produced, Comprehended, or Either Produced or Comprehended (Any Learning) Target Verb as a Function of Learning Condition, Study 1

Condition	Production	Comprehension	Any learning
Impending–unknown	6	9	10
Impending–known	4	8	8
Ongoing–known	3	9	9
Ongoing–unknown	4	6	7

Note. The sample size is 10 in each condition.

location as in the training phase). If the child performed one of the four possible actions following the prompt, O recorded that response and testing ended. If the child made no response, the prompt was repeated up to three times (occasionally by the parent) until the child made a response or clearly lost interest. A second observer also coded the videotapes for which action the child performed in the comprehension test, with perfect agreement (100%).

Results

The results of this experiment were clearcut (see Table 1). Equal numbers of children produced and comprehended the target verb in the four conditions on 1 of the 2 days (both chi-square tests were nonsignificant). Comprehension performance in each of the four conditions was greater than chance (binomial probability < .02 in all four cases). When these two measures of learning were combined to yield a judgment of whether any given child showed some signs of learning (in either comprehension or production; designated in Table 1 as "Any learning"), the values ranged from 7 to 10 children for the four conditions and were again statistically identical (chi-square test was nonsignificant). The similarity of the four conditions was not due to a ceiling effect. Testing at the end of the first day showed less learning overall, but the four groups were still essentially the same: Five to 8 children in each condition showed some signs of learning (chi-square tests for comprehension, production, and any learning were all nonsignificant).

Because the target action was the same for all of the subjects, it was also important to rule out the possibility that subjects in all conditions were choosing to plunk the toy in the comprehension test merely because of some preference for that activity. To test this possibility, we asked ten 24-month-old children (M age = 24.5 months, range = 24–25 months) with the same exposure to the experimental activities but without exposure to the word *plunk*, to plunk a toy as they confronted the same alternative activities as in the comprehension test. Two children chose the plunking activity, a value that does not differ from the 2.5 children expected by chance alone.

Discussion

Children in Study 1 were able to learn the novel verb whether or not it was used with an ongoing or an impending action, and they were able to do this whether or not they knew what action was coming at the time E first touched the toy. With regard to the issue of timing, it is important to note that there was no advantage in this study, as there was in the Tomasello and Kruger (1992) study, for the impending action context. The most likely reason for this difference is that in the Tomasello and Kruger study the target action was novel and exciting and temporally extended, which created high attentional demands. This made it difficult for children to coordinate attention to the action with the simultaneously occurring word as was required in the ongoing condition; in the impending and completed conditions the coordination could occur sequentially. In the present study, although the plunking action was similar, it was much quicker: In the previous study it took the toy over 2 s to complete the action, whereas in this study it took the toy less than 1 s to come through the pipe. The currently most plausible interpretation, then, is that the impending action and ongoing action contexts are both facilitative of child verb learning, with impending better than ostensive if the action is highly engaging and extended over time.

The findings with regard to the known–unknown variable were a surprise. We assumed that knowing the action that was impending would make learning easier for the children, because the known condition has all of the features of the unknown condition plus the knowledge of the upcoming action. It is of course possible that our manipulation did not work and that the children did not develop an expectation in the known condition. However, many of the children demonstrated on more than one occasion an orienting to the plunking activity station when Big Bird was first picked up. Moreover, if the manipulation did not work, it only makes more surprising the children's good performance in the known condition—it means that they were learning with weak or no expectations in all conditions. In any case, the present findings suggest that not only can 2-year-old children learn a verb for an impending action in predictable circumstances (as they did in Tomasello & Kruger, 1992), but they also can learn a verb when they need to hold it in mind until the action comes. It is still our hypothesis that this variable should be an important one, but perhaps it is only important at an earlier age when children's information-processing skills are less fully developed. In any case, these results attest to the robustness of the 2-year-old's verb learning capabilities.

Study 2

A second study was designed to investigate the same questions as those of the first study, but for object label learning. The questions were thus: (a) Do children learn object labels in the impending action context as well as in the ostensive context? and (b) Does knowing the object whose presence is impending facilitate word learning?

Method

Subjects. After completion of the verb game of Study 1 on each of the 2 testing days, children also participated in a second study. To select three distractor items to be used in addition to the target object, upon recruitment, we conducted a brief inventory of possible toys to be used (see next section, *Materials*) during the phone interview. Only children who did not know at least three of the distractor objects were invited to participate in the study.

Materials. A set of seven toys was available to be used in the study. We chose toys that were unlikely to be known by 2-year-olds. Toys were

a liquid-filled paperweight, a toy guitar, a bike horn, a bean-bag frisbee, a spinning top, a Ghostbusters creature, and an irregularly shaped noisemaker with six buttons that produced a variety of sounds when pushed. The target object noun for all of the children was the noisemaker dubbed a *gazzer;* no child had a name for this object (as determined by parental report). For each child, three of the remaining six toys (whose name the child did not know) were used in the study. In addition, a small house with doors on two opposite sides was used. The doors were large enough for each toy to enter and be hidden inside the house.

Experimental design. As in the first study, 10 children were randomly assigned to one of four experimental conditions created by the crossing of (a) the timing of the language model in relation to finding the toy (impending vs. ostensive) and (b) the predictability of the toy that would be found (known vs. unknown).

With regard to timing, in the impending condition E, after hiding the target object inside the house, said, "Where's the gazzer? Let's find the gazzer," and then the child or E opened the child's door and found the gazzer. In the ostensive condition E, after hiding the gazzer in the house, opened the child's door or waited until the child did so and then gave the language model, "You found the gazzer. That's the gazzer." Thus for both model types, each finding of the target was paired with two tokens of the target noun. With regard to predictability, in the known condition, the gazzer was the only toy ever hidden and found in the house. Thus it was expected that after several hidings and findings of the same toy, the child would know that E was going to hide the same toy again. In contrast, in the unknown condition, all four toys (the gazzer plus three selected toys) were hidden one at a time in random order, with the restriction that the same toy was never hidden again until all of the other toys had been found. Thus, on any given trial there was no way for the child to know what toy was hidden in the house.

These two variables were then crossed to form four experimental conditions: impending–known, impending–unknown, ostensive–known, and ostensive–unknown. Children were randomly assigned to one of the four conditions (n = 10 per cell) such that there were 5 girls and 5 boys in each group (and 2 or 3 firstborns per gender grouping).

Training procedure. On each of the 2 days described under Study 1, each child continued to play with E (who was still unaware of the experimental hypotheses) for an additional 20–30 min. All procedures were identical for a given child on both days. All sessions were videotaped by O, who also kept a written record of the child's and E's actions and utterances during the modeling and testing phases of the study.

After an initial warm-up period, E began a series of four rounds of hiding–finding toys and language models. A round began with a set of hiding–finding sequences without any language models. For any given hiding–finding sequence, E sat facing the child with the house placed on a small table between them, a door facing each of them. E then chose a toy from a bag hidden under the table and placed it in the house through the door facing her (out of view of the child). E then invited the child to open his or her door. A round ended with the performance of the assigned language model (impending or ostensive).

In the known condition, the gazzer was hidden one to four times without accompanying language models. Then it was hidden again, and the assigned language model was given for the last finding of the round. This sequence was repeated for four rounds. In the unknown condition, the gazzer was hidden first and found without an accompanying language model on Round 1. Then the other three toys were hidden one at a time, in random order, and each was found without an accompanying language model. Then the gazzer was hidden again, this time found by being paired with the assigned language model (impending or ostensive). The remaining three rounds each consisted of random order hiding–finding of the other three toys followed by the finding of the gazzer with its appropriate language model. Thus at the end of four rounds a subject had received four target noun–target object pairs.

Table 2
Number of Children Who Produced, Comprehended, or Either Produced or Comprehended (Any Learning) Target Noun as a Function of Learning Condition, Study 2

Condition	Production	Comprehension	Any learning
Impending–unknown	8	7	9
Impending–known	5	5	6
Ongoing–known	5	8	8
Ongoing–unknown	6	10	10

Note. The sample size is 10 in each condition.

The entire procedure, including the testing (see next section), was repeated on a 2nd day for a total across the 2 days of eight language models (16 tokens of *gazzer* said in eight pairs, each pair associated once with finding the gazzer).

Testing procedure. The observer, O, kept track of the completion of models and signaled E when the last model had been finished for a given day. She also recorded any spontaneous productions of the target word and contextual notes sufficient to tell what object it was used to indicate. Only if the child was performing the action, about to perform the action, or watching E perform the action when she uttered the target word was it considered an example of correct production. Use of the target word was checked by a second observer using the videotapes with almost perfect agreement (96%).

After the completion of Round 4 on each day, the children were allowed a few minutes of free play with the toys. Then elicited production and comprehension tests were conducted. First, to elicit production at the end of each day, E held up the gazzer and asked the child, "What's this?" If the child gave a verbal response, it was recorded by O, and E went on to the comprehension test. If the child did not respond, the probe was repeated (up to three times) to encourage the child to answer. After a few minutes of free play with the toys, E also attempted to test each child's comprehension. All four toys were laid on the floor in front of the child (in random order) and the child was asked, "Can you give me the gazzer? Give me the gazzer." In making this request E looked directly at the child, not at the objects. If the child chose one of the four objects, O recorded the choice and the session ended. If the child made no response, the prompt was repeated up to three times (occasionally by the parent) until the child made a choice. A second observer also coded the videotapes for which action the child performed in the comprehension test, with almost perfect agreement (96%).

Results

Once again, equal numbers of children produced and comprehended the target word in the four conditions across the 2 days (both chi-square tests were nonsignificant; see Table 2). When these two measures of learning were combined to yield a judgment of whether any given child showed some signs of learning (in either comprehension or production; designated in the table as "Any learning"), the values ranged from 6 to 10 children for the four conditions and were again statistically identical (chi-square test was nonsignificant). Testing at the end of the first day showed less learning overall, but the four groups were still essentially the same: Five to 7 children in each condition showed some signs of learning (chi-square tests for comprehension, production, and any learning were all nonsignificant). Comprehension performance in three of the four conditions was greater than chance (binomial probability $< .01$), with per-

formance in the impending–known condition being marginally different from chance (binomial test < .06).

Because the target object was the same for all subjects, it was important to rule out the possibility that subjects in all conditions were choosing the gazzer in the comprehension test merely because of some preference for that object. To test this possibility, we asked the same 10 children as the controls in Study 1 with exposure to the experimental objects but not to the word *gazzer*, to give E the gazzer as they confronted the same alternative objects as in the comprehension test. Two children chose the gazzer, a value that does not differ from the 2.5 children expected by chance alone.

Discussion

Children in this study were able to learn the novel object label whether or not it was used with an ongoing or an impending language model, and they were able to do this whether or not they knew what object was going to be found. To our knowledge, this is the first experimental demonstration of children learning object labels in nonostensive contexts. The findings are different from those of Whitehurst et al. (1982) and Kuczaj et al. (1989), in our opinion, because our nonostensive contexts were relatively natural modeling contexts, with brief delays within the same action and discourse context. The finding with regard to the known–unknown variable was again a surprise. It was assumed that knowing what object was coming would make learning easier for the children. Once again, it may be that our manipulation did not work totally, but again this only heightens the surprise at the excellent learning in all conditions. Once again, it is also possible that this variable is only important at an earlier age when children's information-processing skills are less fully developed.

Study 3

The third and fourth studies both investigated further the impending action and object contexts, the third with verbs and the fourth with nouns. Given that children do not need to know what action or object is impending to learn a new word in this context, the question of what kinds of cues they might be using arises. In Study 3, the question was whether young children in the impending action context determine the adult's intended referent by means of temporal contiguity (the action immediately following adult use of the new word) or instead use other behavioral cues to determine adult intentions.

Method

Subjects. Sixteen 2-year-old children (M age = 2 years 2 months, range = 2 years 1 month to 2 years 3 months) participated as subjects. There were equal numbers of male and female, firstborn and later-born children. Children were recruited in the same manner as the first two studies. To ensure that the subjects were currently acquiring verbs in their vocabularies (and that they did not know any of the four target verbs), we conducted a brief inventory of common first verbs and the target verbs over the phone. Only children who produced at least five verbs and who did not know the target verbs were invited to participate in the study.

Materials. Two toys were constructed to be used as stimuli: a crane and a merry-go-round. Each toy was designed to support two distinct, novel actions. A set of Sesame Street and Disney characters were used to perform all actions with the toys. An additional play center with several activity stations was also present to be used as a distraction when the experimental trials were set up.

The crane structure consisted of a 12-in. stiff spring mounted vertically on a support base to form the shaft of the crane. The cross-arm of the crane was a wooden, ladderlike piece attached to the top end of the spring and fixed at a 45° angle from horizontal. A clear plastic cup was hung by a string off the elevated end of the cross-arm. The string was laced through the cross-arm and had a large plastic ring tied to the opposite end. The two actions assigned to this toy were to *hoist* and to *twang* the characters who were placed in the clear plastic cup. To hoist Mickey Mouse entailed putting the character in the cup and pulling on the plastic ring until the cup was raised to the cross-arm. To twang Mickey Mouse entailed putting the character in the cup and pushing or pulling on the middle of the cross-arm (where it was attached to the spring) to make the cross-arm and cup jiggle from the spring's action.

The merry-go-round structure was a modified two-tiered lazy Susan. Each tier of the merry-go-round was 14 in. in diameter and was separated by 10 in. dowel rods. The top tier had a 3-in. hole cut in the center. Adjacent to the hole was a hinged platform (2 × 2 in.), with a small handle projecting past the outer edge of the upper tier. The two actions assigned to this toy were to *plunk* and to *whirl* the characters who were placed on the platform. To plunk Mickey Mouse entailed placing him on the platform and tipping it by raising the handle and causing the character to slide off and through the hole. To whirl Mickey Mouse entailed placing him on the platform and pushing against a dowel or the edge of the merry-go-round, making the whole toy (including the character) spin around.

Experimental design. The design of the study was as follows. Each child was taught two words, one associated with each toy. To model the target verbs, E paired a linguistic model with a demonstration of the target action. For every linguistic model, E said, for example, "Watch. I'm going to (*hoist*) Big Bird. I'm going to (*hoist*) him." The target verb was always said twice in the language model.

After the language model, E then used one of two modeling conditions to demonstrate the target action. In the *intentional-first* condition, E first performed the target action (e.g., hoisted Ernie) and then, in the same general activity sequence, performed the nontarget action associated with that toy as an accidental action (e.g., twanged Ernie). To mark these two actions as intentional and accidental, E performed the two actions according to the following script: (a) E says language model, "Let's _____ Big Bird. Let's _____ him." (b) E puts Big Bird in place and says "OK." (c) E performs the target action and says "There!" and (d) E performs the other action "accidentally" (clumsily) and says "Woops!" or "Uh-oh!" In the *intentional-last* condition, E simply had the accident before performing the intentional action. The script was therefore: (a) E says language model, "Let's _____ Big Bird. Let's _____ him." (b) E puts Big Bird in place and performs the nontarget action accidentally (clumsily) and says "Woops!" or "Uh-oh!" (c) E readjusts Big Bird and says "OK" and (d) E performs the target action and says "There!"

Each child served in both conditions, learning one word in the intentional-first condition, the other word in the intentional-last condition (one with each apparatus). The assignment of target words, the order of conditions, and the order of toys were all counterbalanced across conditions.

Training procedure. Subjects came to a psychology laboratory for two 30-min visits, both during the same week. The procedures and experimental conditions for a given child were identical on both days (however, the order of toys was reversed on Day 2). All sessions were videotaped by an observer, who also kept a written record of E's models and any target actions or productions of the child. After an initial warm-up period with the characters and the nonexperimental play center, E

(unaware of the experimental hypotheses) introduced the first toy. After a brief warm-up with the toy, E nonverbally demonstrated both actions that could be performed and encouraged the child to perform both actions as well.

When the child had done both actions at least two times, the modeling phase began. The modeling phase consisted of six pairings of the language models and action models. Thus E would give the language model, saying the target verb twice, and then do the appropriate sequence of intentional–accidental actions. A model was considered complete if the child was attending through its entirety to see both actions performed. If a child was not watching during either action or became distracted before either action was performed, O noted the model as incomplete and another was done later to replace it. Thus each child received six complete models for each toy. If the child or E performed additional nonverbal instances of the two actions, it was noted by O. To ensure that the child did not become preoccupied by just one of the actions, E either invited the child to perform actions an equal number of times or she demonstrated the less preferred action to maintain a balance of each action (always with no accompanying language). After the modeling phase of the first toy was completed, it was removed (out of sight), and the second toy was introduced. The procedure for the second toy was the same as with the first: initial warm-up of its two actions and modeling of the six language–actions pairs.

Testing procedure. The observer, kept track of the completion of models and signaled E when the last model had been finished for a given day. She also recorded any spontaneous productions of the target words and contextual notes sufficient to tell what action it was used to indicate. Only if the child was performing the action, about to perform the action, or watching E perform the action when she uttered the target word was it considered an example of correct production. Use of the target word was checked by a second observer using the videotapes with high agreement (88%).

After completion of the models with the second apparatus, it was removed, and E engaged the child in free play with the nonexperimental apparatus. O then brought both toys back out, placing them side by side in front of E and the child (randomly placed). Before administering the comprehension test, E allowed the child to play with both toys again, making sure the child noticed that both toys were present and performed each of the actions (with no accompanying E language). O then introduced a new character and handed it to the child. Either O or E asked the child to perform the target action of the first toy used on that day (e.g., "Can you go plunk Ernie?"). In making this request E looked directly at the child, not at either of the two toys. The prompt was repeated, up to three times, until the child did one of the four possible actions or clearly lost interest. O recorded the child's first response. (On the few occasions when the child performed the entire sequence of two actions as modeled, that was noted as well.) The second comprehension test was then begun. Again O introduced a new character and gave it to the child. The child was then asked (E looking at child) to perform the action associated with the second toy of that day. Again the prompt was repeated (up to three times) until the child performed one of the four possible actions or lost interest. The child's first response was recorded and the session ended. A second observer also coded the videotapes for which action the child performed in the comprehension test, with perfect agreement (100%).

Results

Children produced the novel verbs in similar ways across the two sessions in the two experimental conditions. As can be seen in Table 3, in each experimental condition there were 5 subjects who produced the target verb for the intentional action at some point in the study. The mean number of correct productions for the two groups also were similar, with just under one production per child in the intentional-last condition and just over one production per subject in the intentional-first condition (t test was nonsignificant). Productions during just the first day of training also were statistically identical (the intentional-first condition had 3 children with 10 total productions, whereas the intentional-last condition had 3 children with 9 total productions).

Table 3
Number of Children (Out of 16) Who Produced, Comprehended, or Either Produced or Comprehended (Any Learning) the Intentionally Marked Verb as a Function of Learning Condition, Study 3

Variable	Intentional first	Intentional last
Children producing for intentional action	5	5
Children misproducing for accidental action	0	0
Average no. of productions for intentional action	1.2	0.9
Children comprehending intentional action	6	10
Children with any learning of intentional action	7	11
Children choosing accidental action	2	2

It is also important that no child produced the target verb for the accidental action in either condition. This provides a comparison to something like chance, as all children had equal opportunities to produce the new word for either of these types of action. Thus, under the assumption that subjects would be equally likely to produce a target verb for the intentional or the accidental action, the five-to-one comparisons in each condition are unlikely to happen by chance (binomial test in each case, $p < .05$).

Also in Table 3 are comprehension scores for subjects in the two experimental conditions. Across both days, subjects comprehended target verbs equally well in the two conditions: Six of the 16 children comprehended the target verb correctly in the intentional-first condition, whereas 10 of 16 children comprehended the target verb correctly in the intentional-last condition (McNemar test was nonsignificant). Again it is important that very few children in either condition acted out the accidental action for the target verb. In the intentional-first condition there were 3 children who acted out the accidental action on the comprehension test on 1 of the 2 days; 1 of these children, however, did this only on the 1st day and then correctly produced the target verb on the 2nd day. If this subject is counted as correct, we may compare the 7 children who acted out the intentional action with the 2 who acted out the accidental action; the probability of this result occurring by chance alone is less than .10 (binomial test). In the intentional-last condition there were 2 children who acted out the accidental action on the comprehension test on 1 of the 2 days (as compared with 10 acting it out correctly); the probability of this result occurring by chance alone is less than .05 (binomial test). It is important to note, however, that both of these 2 children also performed the intentional action on 1 of the 2 days (one on the 1st day and one on the 2nd day—the latter also correctly producing the target word on the 2nd day). If these subjects are counted as correct or ex-

cluded from the analysis the result is even stronger. (One child acted out the entire action sequence—accidental then intentional—in response to the request for the target action on both days. If he is excluded from the analysis, the binomial test is still significant.)

Also in Table 3 is the number of children either comprehending or producing the target verb correctly on 1 of the 2 days (any learning). One child produced correctly but did not comprehend correctly in each condition, and thus there were 6 and 11 children with some form of learning in the two conditions, respectively; these are statistically identical. It should also be mentioned that after the 1st day only, 6 children comprehended the target verb correctly in the intentional-first condition, and 5 children comprehended the target verb correctly in the intentional-last condition (McNemar test was nonsignificant).

As an additional control for the comprehension test, tapes were scored for the number of times each subject performed each of the actions during the entire two sessions. This included both spontaneous performances, as well as performances on training trials. There were no differences (analysis of variance was nonsignificant): Children performed the intentional-first action 16.00 times (SD = 9.34), the intentional-last action 13.00 times (SD = 9.26), and the other two actions an average of 15.13 times (SD = 11.53). This indicates that before comprehension testing on each day, subjects indicated no preference for either of the target activities.

Discussion

Young children do not learn verbs in the impending action context by means of any simple rule such as "map the new word onto the next action to be performed by the speaker." They will ignore the next action and learn the new verb for the second action to follow if the intermediate action is marked as accidental (and the second is marked as intentional). Children in this situation clearly monitor the actions of E, discriminate accidental from intentional actions, and learn the new word for intentional actions only. This means that the process of learning a novel verb in this situation is not something passive and associative, depending on temporal contiguity of word and referent, but rather it is a much more active process in which the child is continuously seeking to understand the adult's intentions within the immediate action and discourse context.

It is important to note that the cues available to children in this study were composite. The intentional action was indicated by both verbal ("There!") and nonverbal (facial expression, smoothness of executing behavior, etc.) cues, as was the accidental action ("Woops!" and clumsy performance). We chose to use these composites because our initial attempts to dissociate the different components made for very unnatural behaviors (e.g., it is very difficult to keep oneself from saying "Woops!" in the presence of a child when an accident happens). We thus do not know precisely which of the specific behaviors specifying intentional and accidental actions in the present study were the ones that children used.

Study 4

The fourth and final study was an analogue of Study 3, but with concrete nouns. An adult who announced her intention to find an object picked up and rejected two objects before showing glee at finding the sought object on the third try. In comparing children's learning in this condition to that in a condition in which the adult found the intended object on the first try, our goal was to answer the question of whether children use temporal contiguity or more active social-pragmatic cues to determine adult referential intentions. Unlike the first three studies, the fourth study involved one session only.

Method

Subjects. Thirty 2-year-olds (M age = 2 years 0 months, range = 1 year 11 months to 2 years 2 months), mostly White and middle class and from the same pool as the other studies, participated as subjects. Parents of each child were contacted by phone and were invited to participate if the child was currently showing productive language (minimum 50 word vocabulary as determined by maternal estimate).

Materials. A set of five toys was available to be used in the study. All were chosen to be unlikely to be known by 2-year-olds. These included a liquid timer, a bean-bag frisbee, a bike horn, a rolling toy that clicked, and a spinning top. Four additional toys were available if a given child knew any of the preselected five. These were a plastic slinky, a Ghostbusters creature, a noise maker with six buttons to push, and a toy guitar. The hiding apparatus consisted of a set of five plastic buckets mounted in a horizontal row on a wooden plank. The buckets were large enough for each toy to fit comfortably inside, and all had lids that could be locked with a twisting motion. Also in the room was a play center, with a set of activity stations and Sesame Street characters; this setup was used during an initial warm-up period and as a distractor between the rounds of hide-and-seek.

Experimental design. Each child was taught one new noun, *toma.* The toy assigned to be the toma was counterbalanced across subjects and conditions. Each child in each condition played a finding game on three occasions (each game being called a round of the game). Each round of the game consisted of (a) E pointing out to the child the row of closed buckets (O having stocked each bucket with a toy, one of which was the child's target toy, the toma); (b) E saying, "Hey Annie! Where's the toma? Let's find the toma. Where's the toma?"; (c) E searching and finding the target toy (toma) in a predetermined way depending on condition (see below); and (d) E saying "Let's see what's in *here*?" and finding each of the four nontarget toys in the other buckets. In both conditions and for both target and nontarget objects, the actual finding event consisted of E opening the bucket, extracting and holding up the found toy with wide eyes and an "Ah!" gasp, and then giving the toy to the child. All objects were replaced before another object was found, and the row of buckets was returned to O when the round was over (i.e., after all objects were found and replaced). The point of the nontarget extractions was that when the child was asked in the comprehension task (see below) to get the toma, there would not be a bias in the direction of the target object simply because E had interacted with it more often or had shown more affect when finding it.

The two experimental conditions were *with search* and *without search.* Children were randomly assigned to one of these two conditions (n = 15 per group) such that there were 8 boys (4 firstborns and 4 later borns) and 7 girls (4 firstborns and 3 later borns) in each condition. All nontarget-finding events were the same in both conditions. The only difference was in the finding of the target object. In the with-search condition, the target-finding event consisted of E first taking out two objects and rejecting them before finding the target toy. Thus each target-finding event consisted of E (a) opening the bucket two to the right of the goal bucket, frowning while she extracted and held up the toy from within (E held the toy so the child could see it and continued to look at it disapprovingly, not making eye contact with the child at any time), then replacing it and the lid; (b) doing the same thing with the toy in the

bucket immediately to the right of the goal bucket; and then (c) opening the goal bucket and finding the target object (with the gleeful behaviors described in the immediately preceding paragraph). In the without-search condition, E simply opened the bucket containing the target object and found it immediately (with the gleeful behaviors described earlier). The target object was hidden in a different bucket for each of the three rounds of the game.

Training procedure. Subjects came to a psychology laboratory for one visit. On arrival, the parents were asked to look at a set of photos of the five preselected toys. They were asked if the child knew any of the toys by name. If the parent reported that the child knew one or more of the toys, the parent was asked to select a replacement toy (or toys) that the child did not know from the reserve set of four toys (also presented by photograph).

Each child then played with a female E (who was unaware of the experimental hypotheses) for approximately 30 min. An observer, O, was also present. O kept a written record of the child's and E's utterances during the modeling and testing phases of the study and rehid the toys between rounds of the game. All sessions were videotaped. After an initial warm-up with the play center, E then engaged the child in a warm-up with the buckets and toys. The child sat on the floor (or on a parent's lap) facing E, with the buckets placed between them. E opened each bucket one at a time and showed the child how each toy worked. The child was allowed to play with each toy. This warm-up allowed E to "set the script" of the game, so the child understood that E would open each bucket and then give the toy to the child.

After all five toys had been introduced, E and the child moved to the distractor toys so that O could set up the first round of the hiding game. When it was ready, E began with the first round. E began by giving the language model and finding the toma according to the assigned condition (with or without search); for this round the target object was always in the bucket to the far left. After the child was allowed to play with the target toy, it was replaced and covered. E then proceeded to extract each of the other four nontarget objects in this same way but without any reference to the toma. Thus, E said "Let's see what's in here" and opened the second bucket, finding a nontarget object in the same way as the target object had been found (i.e., extracting and holding up the toy with wide eyes and an "Ahhh!" gasp, and then giving the toy to the child). Each object was replaced after the child had had a chance to interact with it briefly. This was repeated until all five toys had been found. Then the buckets were returned to O to be set up for Round 2 while E and the child played with the distractor toys.

The following two rounds were the same as the first, with the exception that a different bucket contained the target toy and thus the target-finding event occurred at different places within the round. In the second round, E began by saying, "Let's see what's in this one," and the toy in the first bucket was extracted excitedly. After this object was replaced, E gave the language model and then found the toma in the second bucket (either with or without search, depending on condition). The remaining objects were then found in this same way. In the third round, E began by saying, "Let's see what's in this one," and the toy in the first bucket was extracted; this same procedure was followed for the toy in the second bucket. E then gave the language model and found the toma in the third bucket (either with or without search, depending on condition). The remaining objects were then found in this same way. At the end of the third round, all five toys were removed from the buckets and the child was allowed to play with them for a few minutes. E and the child then played with the distractor toy while O set up the comprehension test.

Testing procedure. During the training procedure, O recorded any spontaneous child productions of the target word and made contextual notes sufficient to tell what object it was used to indicate. Only if the child was performing the action, about to perform the action, or watching E perform the action when she uttered the target word was it considered an example of correct production. Use of the target word was checked by a second observer using the videotapes with perfect agreement (100%).

Following the completion of Round 3, each child was allowed a few minutes of free play with the toys. Then elicited production and comprehension tests were conducted. First, to elicit production, E held up the toma and asked the child, "What's this?" If the child gave a verbal response, it was recorded by O, and E went on to the comprehension test. If the child did not respond, the probe was repeated (up to three times) to encourage the child to answer. After a few minutes of free play with the toys, E also attempted to test each child's comprehension. All five toys were laid on the floor in front of the child (in random order) and the child was asked, "Can you give me the toma? Give me the toma." In making this request E looked directly at the child, not at the objects. If the child chose one of the five objects, O recorded the choice and the session ended. If the child made no response, the prompt was repeated up to three times (occasionally by the parent) until the child made a choice. A second observer also coded the videotapes for which object the child chose in the comprehension test, with perfect agreement (100%).

Table 4
Number of Children (Out of 15) Who Produced, Comprehended, or Either Produced or Comprehended (Any Learning) the Intentionally Marked Noun as a Function of Learning Condition, Study 4

Variable	Without search	With search
Children producing for intended object	4	6
Children producing for first rejected object	—	1
Average no. of productions for intended object	0.80	0.87
Children comprehending intended object	10	6
Children with any learning of intended object	11	8
Children comprehending as first rejected object	—	1

Note. Dashes indicate not applicable.

Results

Children produced the novel noun in similar ways in both conditions. As can be seen in Table 4, there were 4 and 6 children in the without-search and with-search conditions, respectively, who produced the target word (chi-square was nonsignificant). There was also no difference in the average number of child productions (t test was nonsignificant). Also in Table 4 are comprehension scores for subjects in the two experimental conditions. Subjects comprehended target words equally well in the two conditions: Ten of 15 children in the without-search condition and 6 of 15 children in the with-search condition chose the correct toy in the comprehension test (chi-square was nonsignificant). Also in Table 4 is the number of children either comprehending or producing the target word correctly. These values are 11 and 8 for without search and with search, respectively, and they do not differ statistically.

It is important as a comparison that subjects in the with-search condition who did not choose the toma did not have sys-

tematic preferences for any of the distractor objects. In particular it is important that almost none of the children in that condition thought that E intended to name the first object extracted from the first bucket. One child both produced the target word for an erroneous object and in the comprehension test chose the first object E had touched in the model. This means that there were 8 subjects who learned the word for the target object and 1 who learned the word for the distractor object first touched but rejected by E. This is unlikely to happen by chance alone ($p < .05$, binomial test). There was 1 other subject (not reflected in these numbers) who chose the first touched nontarget object in the comprehension but produced the target word correctly. If this subject is credited with being correct, the finding is strengthened (9 vs. 1). If he is credited with being incorrect, the numbers change to 8 versus 2, a marginally significant difference ($p = .055$, binomial test).

As an additional control for the comprehension test, tapes were rescored for the amount of time each subject spent interacting with (i.e., touching) each object on each of the three rounds (subjects were given each object to play with briefly on each round). There were no differences (t test was nonsignificant): Children interacted with the toma an average of 70.31 s ($SD = 53.96$) and the other toys an average of 68.36 s ($SD = 47.79$). This indicates that before comprehension testing, subjects indicated no preference for the target object.

Discussion

Young children do not learn nouns in the impending action context by means of any simple rule such as "map the new word onto the next object touched by the speaker." They will ignore an immediately encountered object and learn the new word for a later encountered object if the first object is marked as one that is not intended and the later object is marked as the one intended. Children in this situation clearly monitor the actions of E, discriminate accidental from intentional actions, and learn the new word for the intended object only. This means that the process of learning a novel object label in this situation is not something passive and associative, depending on temporal contiguity of word and referent, but rather it is a much more active process in which the child is continuously seeking to understand the adult's intentions within the immediate action and discourse context.

As in Study 3, there was a complex set of behaviors that indicated the adult's intentionality in this study. Rejecting objects as not intended involved negative facial expressions and replacing rather than interacting with the object. Finding the intended object was marked by a joyful facial expression, eye contact with the child, and the giving of the object to the child. Again we would argue that these patterns go together in the natural social world and to tease them apart experimentally would lead to some very unnatural behaviors (e.g., expressing glee at finding the object but replacing it and continuing the search). This means, however, that for the present we do not know precisely which cues the children used. But we do know that children can use different cues in this same situation. Baldwin (1993a) found that children chose an object an adult looked at while uttering a novel word rather than the first one they encountered after having heard the word. In combination with the present findings, this implies that children are capable of using multiple cues to determine the object of adult acts of reference and that all of these cues are more powerful than temporal contiguity. This overdetermination of cues argues that young children are not learning such things as gaze direction simply as specific discriminative stimuli for correct word–world mappings, but rather as they interact with adults they are learning various cues in an attempt to understand something of the reasons—the underlying intentions—that make the behavior of others comprehensible. This accords with other research showing that children in their second year of life make attempts in other ways to tune into adult intentions through, for example, attempts at social referencing and imitation (Tomasello, 1992b).

General Discussion

The present studies explored 2-year-old children's abilities to learn words in nonostensive contexts. More specifically, the focus of all four studies was on the impending action–object context in which an adult announces her intention to perform an action or find an object. In the case of object labels, it was found that young children do not need to be able to anticipate the object to be found in order to learn a name for it; they can "hold the word in mind" until the adult fulfills her finding intention. They can even do this when it means skipping over an intervening object, provided that the adult provides clear behavioral cues that this is not the intended object (and other nonverbal cues that positively indicate the intended object). In the case of verbs, a similar result was found. Young children do not need to be able to anticipate the action that is impending in order to learn a new word for it. Again they can hold the new word in mind until the action comes, and again they can do this even when it means skipping over an intervening action, provided that that action is behaviorally marked as accidental (and the intended action is marked as intended).

One possible confound in the comprehension tests of all four studies was as follows. In the comprehension test the child's task was to choose from among a number of objects or actions, all of which he or she had experienced previously. It is thus possible that a given child might choose a target object or activity because it had been treated in a certain way previously, despite our attempts to treat all experimental objects and actions equivalently. We do not think this interpretation is plausible for the following reasons. First, in addition to the controls implicit in the experimental designs, each of the four studies included additional control measurements (differing slightly in each of the four cases) to see if children had an object or action preference. In each case children showed no preference for the object or action that was named for them before the comprehension test. Second, the production measures in all cases corroborated the results of the comprehension tests. In the first two studies, approximately one half of the children produced the target word correctly, and in the third and fourth studies, approximately one third of the children produced the target word correctly. The probability of producing a new word for its appropriate referent (required in all cases) is presumably quite low given the number of objects and actions available to the child, especially if the productions were spontaneous, which they were in the vast majority of cases in these studies. Finally, we should at least

mention that it is our strong impression that when 24-month-old children know what an adult is asking for they often, but not always, comply. When they do not know what the adult is asking for, however, they do not then look for something made salient by the adult, but rather ignore the request and do what they feel like doing at the moment. For these reasons, we believe that the children in our studies did indeed learn the experimental words.

The learning contexts investigated in these studies are probably not the only nonostensive contexts in which young children can acquire new words. For example, one interesting possibility for verb learning is that children might "skip over" some intermediate acts if they were seen as clearly preparatory to some consummatory act, in a situation in which the child understands the relation between the two acts. Thus, a child might learn a novel verb for an action on a novel toy in the impending action context even if the adult first performs the act of arranging a toy's parts so that the intended action is made possible. In the case of noun learning, there are any number of social-pragmatic cues that might override temporal contiguity in the child's attempt to understand a novel adult word. For example, in addition to eye contact with the object as in Baldwin (1991) and glee in a finding game as in the present Study 4, one can imagine situations in which adults announce all sorts of intentions ("I'll put some dax on your cake," "I need the dax," "I'll use the dax," "I'll go get the dax," "The dax goes on here," etc.) and then proceed to an intended object with intermediate objects intervening in all sorts of ways. In both the case of verbs and nouns, one can also imagine cues to past actions and objects (many of them verbal) that would indicate skipping over the last object or action seen, for example, "I put the toma on there," "You daxed down the slide," "Where's that peri you had?," "I chammed that book," and so forth. Whether children use any or all of these cues to override temporal contiguity is a question for future research.

It is important to emphasize that the nonostensive contexts used in the present studies—and many others we might envision—are not strange or infrequent learning contexts in the lives of beginning language learners. Indeed, we would argue that it is the ostensive learning context that is strange and infrequent; it is frequent only in psychological experiments and the homes of Western middle-class parents who are concerned that their children have large vocabularies early in development. Children in many cultures, and many children in Western cultures (e.g., working-class children and children growing up in multichild homes), experience their language in a variety of pragmatic contexts, not just those in which someone points and names things for them (cf. Barton & Tomasello, 1994; Schieffelin & Ochs, 1986). A strong argument could in fact be made that acquiring words within the natural flow of human action and intercourse, rather than within minilinguistics lessons, is the norm for all children in some cultures, for some children in all cultures, and, perhaps, for all children in all cultures when learning some types of words (e.g., verbs). And we should also emphasize that, in our interpretation, even the ostensive context requires the child to engage in an active process of intention cue detection such as the ones they engaged in here in nonostensive contexts. It is just that learning in the ostensive context is based on pragmatic cues such as gaze direction and pointing that are so basic and transparent that we do not notice them, despite the well-known and trenchant analyses of Wittgenstein (1953) and Quine (1960) on the social-pragmatic dimensions of the ostensive context.

It is also important to note that none of the well-known lexical constraints posited by theorists such as Markman (1989) could have guided the children in our studies to the adult's intended referent. In Study 4, for example, all five of the novel objects would have fallen under the same set of constraints. They were all whole objects, and thus the *whole-object* constraint would not have helped to choose among them. Also, children did not have a name for any of the objects, and thus the *mutual exclusivity* constraint would not have helped either. Nor were there any pertinent linguistic cues that might have helped the child to choose. In Study 3, a similar situation held for verb acquisition. All four actions might have been the referent of the novel verb. These did not differ in any way in terms of their salience as actions (in fact, they were counterbalanced across conditions), and the child did not have a name for any of them. A whole-action constraint, if there were such a thing, therefore could not help, nor could mutual exclusivity. It is also important to emphasize that the syntactic context in which the verb was used—emphasized by Gleitman (1990) in her theory of syntactic bootstrapping—could not have been of any help to children in this study because the syntactic context was identical across conditions (cf. Rispoli, in press; Tomasello, in press). Thus, although it may be the case that constraints in some form are helpful to beginning language learners (perhaps conceived of as more general knowledge about the world or knowledge of human action; see Clark, 1988; Markman, 1992), our results strongly suggest that they are not sufficient, at least not for 2-year-olds in the kinds of situations we tested here. Two-year-olds learned to pick out referents in our studies by relying mainly on social-pragmatic information about the adult's intentional state.

It is quite possible that lexical constraints and social-pragmatic cues provide complementary sources of information in the word learning process, with constraints helping the child to determine "what kind of referent" and social-pragmatic information helping him or her to decide "which referent." Another possibility, however, is that children's understanding of the pragmatics of social interaction—made possible by their considerable skills of social cognition and social learning and their many months of experience with social interaction by the time language acquisition begins—underlies their determinations of what kind of referent as well (Tomasello, 1992b, in press). If this were the case, lexical acquisition might be seen as one instance of a social-cognitive process that we have previously referred to as cultural learning (Tomasello, Kruger, & Ratner, 1993). At this developmental level, cultural learning refers to children's abilities to understand the intentions of others and to imitatively learn from them the appropriate behavioral strategies for expressing those same intentions themselves. Children at this age demonstrate these abilities not only in their acquisition of linguistic structures, but also in their skills of nonverbal joint attention, social referencing, and imitation of actions on objects, not to mention their emerging skills at manipulating the intentional states of others through nonlinguistic communication. Our main argument is thus that lexical acquisition uses the

same basic abilities of social cognition and cultural learning that children display in other social-cognitive domains at this same age, and that these abilities are exercised in the process of lexical acquisition in much more powerful ways than has hitherto been recognized. How these abilities interact with other processes of lexical acquisition, such as potential lexical and syntactic constraints, is a question for future research.

References

Akhtar, N., Dunham, F., & Dunham, P. (1991). Directive interactions and early vocabulary development: The role of joint attentional focus. *Journal of Child Language, 18,* 41–50.

Baldwin, D. (1991). Infants' contribution to the achievement of joint reference. *Child Development, 62,* 875–890.

Baldwin, D. (1993a). Early referential understanding: Young children's ability to recognize referential acts for what they are. *Developmental Psychology, 29,* 1–12.

Baldwin, D. (1993b). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language, 2,* 395–418.

Barton, M., & Tomasello, M. (1994). The rest of the family: The role of fathers and siblings in early language development. In C. Gallaway & B. Richards (Eds.), *Input and interaction in language acquisition* (pp. 109–134). Cambridge, England: Cambridge University Press.

Brown, R. (1958). How shall a thing be called? *Psychological Review, 65,* 14–21.

Bruner, J. (1983). *Child's talk.* New York: Norton.

Clark, E. (1988). On the logic of contrast. *Journal of Child Language, 15,* 317–336.

Gleitman, L. (1990). The structural sources of verb meaning. *Language Acquisition, 1,* 3–55.

Kuczaj, S., Carter, A., Sherman, J., & Borys, R. (1989, April). *On children's use of experience to determine the meanings of words.* Paper presented to the biennial meeting of the Society for Research in Child Development, Kansas City.

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

Markman, E. (1992). Constraints on word learning: Speculations about their nature, origins, and word specificity. In M. Gunnar & M. Maratsos (Eds.), *Modularity and constraints in language and cognition* (pp. 59–101). Hillsdale, NJ: Erlbaum.

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

Rispoli, M. (in press). Missing arguments and the acquisition of predicates. In M. Tomasello & W. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs.* Cambridge, England: Cambridge University Press.

Schieffelin, B., & Ochs, E. (1986). *Language socialization across cultures.* Cambridge, England: Cambridge University Press.

Tomasello, M. (1988). The role of joint attention in early language development. *Language Sciences, 11,* 69–88.

Tomasello, M. (1992a). *First verbs: A case study of early grammatical development.* Cambridge, England: Cambridge University Press.

Tomasello, M. (1992b). The social bases of language acquisition. *Social Development, 1,* 67–87.

Tomasello, M. (in press). Pragmatic contexts for early verb learning. In M. Tomasello & W. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs.* Cambridge, England: Cambridge University Press.

Tomasello, M., & Kruger, A. (1992). Acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language, 19,* 311–333.

Tomasello, M., Kruger, A., & Ratner, H. (1993). Cultural learning. *Behavioral and Brain Sciences, 16,* 495–552.

Whitehurst, G., Kedesdy, J., & White, T. (1982). A functional analysis of meaning. In S. Kuczaj (Ed.), *Language development: Vol. 1. Syntax and semantics* (pp. 397–428). Hillsdale, NJ: Erlbaum.

Wittgenstein, L. (1953). *Philosophical investigations.* New York: MacMillan.

Received October 27, 1992
Revision received February 18, 1994
Accepted February 18, 1994 ■