Infants’ Learning about Words and Sounds in Relation to Objects

Amanda L. Woodward and Karen L. Hoyne

In acquiring language, babies learn not only that people can communicate about objects and events, but also that they typically use a particular kind of act as the communicative signal. The current studies asked whether 1-year-olds’ learning of names during joint attention is guided by the expectation that names will be in the form of spoken words. In the first study, 13-month-olds were introduced to either a novel word or a novel sound-producing action (using a small noisemaker). Both the word and the sound were produced by a researcher as she showed the baby a new toy during a joint attention episode. The baby’s memory for the link between the word or sound and the object was tested in a multiple choice procedure. Thirteen-month-olds learned both the word–object and sound–object correspondences, as evidenced by their choosing the target reliably in response to hearing the word or sound on test trials, but not on control trials when no word or sound was present. In the second study, 13-month-olds, but not 20-month-olds, learned a new sound–object correspondence. These results indicate that infants initially accept a broad range of signals in communicative contexts and narrow the range with development.

INTRODUCTION

Word learning can be thought of as a feat of social reasoning. The learner’s task is to interpret the speaker’s behavior—the words he or she produces as well as aspects of his or her actions that are relevant to what he or she means to say. For babies, part of this task is learning that different actions have different kinds of communicative value. In particular, babies must acquire the understanding that some communicative behaviors serve as names for things. At a theoretical level, names are defined by what they do. Names are symbols that are used to communicate about their referents. They extend to kinds of entities, rather than being associated with clumps of experience. Names are also independent in form from their referents. Their particular form is arbitrary, set by convention, rather than being determined by the natural order of things. Despite this arbitrariness, within a language, there is a regularity in the form of names that might be useful to learners. In spoken languages, for example, names occur as units of speech, not as other behaviors or sounds. Other actions may provide information relevant to objects—for example, about the typical uses of objects (holding a telephone to one’s ear) or about games that can be played with them (making a “whoosh” sound as one moves a toy airplane through the air)—but these actions are not names.

If babies understood this perceptual regularity, it could inform their learning. That is, if they could identify the kinds of perceptual units that were likely to be names, then learning how names relate to referents would be more tractable. In addition, hearing a new spoken word, as opposed to encountering some other new action, could signal that it was appropriate to search for a referent. In the studies presented here, we asked whether babies’ learning about signals that occur during communicative episodes is informed by the expectation that names will be in the form of spoken words. To ask this question, we presented babies with novel sound-producing behaviors in a context that is typically used to teach new names—during a joint attention episode in which an adult showed the baby a new toy. We then measured babies’ acceptance of the sounds in this context by testing whether they remembered the link between the sound and the toy.

Parents in western cultures often introduce new labels when they and the baby are jointly attending to a new item (Bridges, 1986; Collis, 1977; Fernald & Morikawa, 1993), and a number of studies have found that babies in their second year readily acquire new word–object mappings in this context (e.g., Baldwin, 1991; Oviatt, 1980; Schwartz & Leonard, 1980; Tomasello & Farrar, 1986; Woodward, Markman, & Fitzsimmons, 1994). In one of these, Woodward, Markman, and Fitzsimmons (1994) introduced 13- and 18-month-olds to a new object label given during a joint attention interaction. Babies saw a new object and were given a novel label for it just nine times in the course of a few minutes. Then, babies’ learning of the link between the label and the object was assessed using a multiple-choice procedure. When the testing conditions were simplified to limit babies’ information processing load, even the youngest babies systematically chose the correct object (see Schafer & Plunkett, 1998, for similar findings using a different methodology). Thirteen-month-olds performed at above chance rates.
even when a 24-hr delay was imposed between training and testing. Although these findings indicate that infants can form mappings between words and objects during joint attention, they do not provide information about the specificity of these mappings. Would babies link any signal given during joint attention with an object? Or instead is their learning informed by the expectation that some kinds of signals, spoken words, function in labeling contexts whereas others do not?

Two recent studies bear on this question. In one, Namy and Waxman (1998) found that babies who are acquiring spoken language sometimes accept novel gestures as labels for objects. Namy and Waxman introduced 18- and 26-month-old babies to a gesture that was given as the label for an object. The researcher showed the baby a toy, for example, a toy apple, and said, “We call this—” and then produced a novel gesture with her hand. Then the researcher showed the baby a second toy apple and an unrelated toy, for example, a pig, and asked, “Can you get the [gesture]?” Eighteen-month-olds chose the apple, suggesting that they had accepted the gesture as a label for the toy. Twenty-six-month-olds, in contrast, did not choose the correct item reliably. With more extensive training, Namy and Waxman were able to get 27-month-olds to accept the gestural labels. Thus, this study suggests that 18-month-olds accept as labels behaviors that older babies do not readily accept.

In contrast, a second recent study indicates that 12- and 18-month-olds respond differently to spoken words and other actions in a telling way. Baldwin, Bill, and Ontai (1996) introduced babies to a new word or to a nonlinguistic verbal sound, a sigh. They were interested in babies’ propensity to follow an adult’s gaze, a behavior critical to determining a speaker’s referential intent that has been shown to play a role in word learning for 18-month-olds (Baldwin, 1991). They found that babies at both 12 and 18 months were more likely to follow an adult’s gaze after she had uttered a novel word than after she had produced a sigh. This suggests that by 12 months, babies expect that some sounds will be used referentially whereas others will not.

In each of these studies, babies’ history with the medium of the signal could have contributed to their willingness to treat the signal as communicative. In the case of sighs, babies have most likely heard such sounds in the past and may have learned that they are not ordinarily used to communicate about objects. In the case of gestures, babies may have learned just the opposite. Adult speakers accompany their utterances with gestures, and these gestures often convey unique information (Alibali & Goldin-Meadow, 1993; McNeill, 1992). Moreover, unlike adult speakers, babies often use isolated gestures to communicate (Acredolo & Goodwyn, 1988; Bates, Benigni, Bretherton, Camaioli, & Volterra, 1979; Goodwyn & Acredolo, 1993). Gestures convey meanings differently than words do (for discussions, see Goldin-Meadow, McNeill, & Singleton, 1996; Pettito, 1988). Nevertheless, the fact that gestures are used to convey information may facilitate babies’ acceptance of gestural labels.

Thus, the findings of these two studies may reflect babies’ learning about particular kinds of signals. This learning aside, we wondered whether babies have further refined their expectations so that they would accept only words (or signals that have a history of communicative use) in labeling contexts. To test this, we presented babies with signals with which they did not have previous experience. Babies saw an experimenter produce an unfamiliar sound using a small artifact, for example, an electronic beeper or a whistle. We embedded the sounds in joint attention episodes like the ones in which babies often learn new words: The experimenter established joint attention on a new object with the baby and then intentionally produced the novel sound.

We began by testing 13-month-olds, because by this age, babies seem to know some word–object and word–event correspondences (Bates, Bretherton, & Snyder, 1988; Benedict, 1979; Huttenlocher, 1974), and, as discussed above, they are able to learn new word–object mappings relatively quickly. In our first study, half the babies were introduced to a new object label. This condition closely approximated Woodward et al.’s (1994) procedure. Given prior findings, we expected that 13-month-olds would learn the word–object mapping based on this training. The other half of the babies were introduced to a novel nonlinguistic sound that was produced by the experimenter. If babies have determined that words, but not other sounds, function as labels, we would expect them not to learn this mapping as easily.

The measure of learning was a multiple-choice test. Babies were shown a tray containing the previously labeled (or “sounded”) object and a distracter object and were asked to choose one as they heard the label or the sound. A concern discussed in Woodward et al. (1994) is that the experience of training (i.e., having someone comment on an object while engaged in joint attention), might lead infants to prefer the labeled object. Then, babies might choose the target on test trials not because they remember the link between the word and the object, but because they like the previously labeled object better than other objects. In those studies, this potential problem was dealt with in two ways: (1) to make the distracter object equally inter-

---

**References**

esting to babies, it was also introduced during joint attention, and (2) to test for possible item preferences, there were no-label control trials on which babies were presented with the target object and a distracter and asked to choose one. The first of these procedures was successful as measured by the second: Babies did not choose the previously labeled object systematically on control trials. Thus, for those studies preferences could be eliminated as the source of systematic choices on experimental trials. We opted to follow both of these procedures in the current studies. That way, we could be sure that systematic choices on experimental trials indicated that babies had learned the link between the sound or word and the object and not that babies had a preference for the target object.

In summary, we tested babies’ expectations about the forms of names by presenting non-word signals in a learning context that is typical for names and then assessing whether babies learned about these signals. That is, we tested whether babies would treat words and sounds alike when they occurred in the same communicative context. This approach does not in itself address the question of whether infants understand the necessary characteristics of names (that they are symbols that are used to communicate about things and events, that they extend to members of a kind, and so forth). We will return to this question at the end of the paper.

STUDY 1

Method

Participants

Sixty-four healthy, term babies from the Chicago area participated in Study 1. They had a mean age of 13 months 16 days and ranged in age from 12 months 18 days to 14 months 7 days. All infants were acquiring English as their first language. Parents had been contacted by mailings and advertisements and were offered a small travel reimbursement. An additional 23 infants were tested but were not included in the final sample because they did not complete all trials (11), because of experimenter error (8), or because the parent said the experimental word (4). In the final sample, there were 32 infants in the word condition (mean age = 13 months 15 days) and 32 in the sound condition (mean age = 13 months 18 days).

Design and Procedure

The experimental procedure had two parts: a training phase in which infants were introduced to the correspondence between a new word or sound and a new object, and a testing phase in which their learning of this correspondence was assessed using a multiple-choice procedure. One group of infants was introduced to a novel word–object correspondence and another group was introduced to a novel sound–object correspondence. Within each of these groups, half of the infants received experimental test trials, which assessed infants’ learning of the word– or sound–object link, and the other half received preference control trials, which assessed whether infants had a preference for the object that had been paired with a word or sound. The training and testing phases of the study were videotaped.

Training procedure. Babies heard the word or sound nine times as they attended to the target object. The trainer introduced the word or sound in the course of a game in which the baby took toys out of a plastic bucket. The trainer was careful to only produce the word or sound when she was sure that the baby was attending to the object and to what she was saying. That is, she first established joint attention on the object with the baby. The novel word was introduced in common labeling frames, for example, “Look, it’s a toma. See? A toma. That’s the toma.” The trainer produced the sounds using a small hand-held noisemaker that was hidden inside a gray bag. She placed the bag near to the target object as she produced the sound. She played the sound just after drawing the baby’s attention to the toy verbally, for example, “Look at this. [Squeak]. Yeah, see it? [Squeak]. Wow, look! [Squeak].” The sounds or words were given in groups of three. Once the trainer had produced the word or sound three times, she took the object away and gave the baby a familiar filler toy to play with for a moment. After taking this toy away, she then introduced a second novel object, which served as the distracter during the test phase. She did not label this object, but talked about it and drew the baby’s attention to it so that it would be familiar during the test. This process was repeated three times. For half of the babies in each condition, the trainer began by introducing the distracter.

There were two novel words, toma and gombie, and two novel sounds, an electronic beeper and a squeak.
produced by squeezing a rubber toy. Approximately half of the infants in the word condition heard the word *toma* as the training word and half heard the word *gombie*. Half of the infants in the sound condition heard the squeak during training and half heard the beeper. The two novel objects were a plastic coaster with a cork center and a piece of rubber tubing molded to fit the end of a bicycle handlebar. Approximately half of the infants in each training condition were assigned to have the coaster as the target object and the other half were assigned to have the tube as the target object. There were two exemplars of each type of object, differing from each other only in color. The handles were pink and white and the coasters were purple and yellow. Each infant was randomly assigned one token of each object kind as the training pair.

The trainer attempted to equate the pace of training in the word and sound conditions. After-the-fact coding of the session videotapes confirmed that the word and sound conditions did not differ in the amount of time that babies saw the target object or in the overall length of the training session.

**Testing procedure.** Following training, a multiple-choice procedure was administered as a measure of learning. The multiple-choice test was conducted by a second experimenter who did not know which of the two objects had been assigned as the target for the baby. This prevented the tester from providing any inadvertent cues to the baby. To keep babies attentive and to minimize the effects of preferences for particular toys, we used the procedure developed by Woodward et al. (1994). That is, the test trials were embedded in play activities in which babies used the objects to complete activities such as filling a box or sliding objects down a chute, babies were allowed to take frequent breaks, and the tester only began a trial when the baby was not distracted by other toys (for details see Woodward et al., 1994). Before any test trials were given, each baby was given two warm-up trials in which the tester asked him or her to take a familiar object from the tray. The tester showed the parent a set of small familiar objects (including a shoe, a toy duck, a cup, a toy car, a toy bunny, a toy train, a toy bottle, and a toy cow) and asked which two the baby would most likely recognize by name. She then put these two in the tray and asked the baby to choose one of them twice in a row. The tester gave the baby feedback on these trials, saying, for example, “Yes! You got the duck!” if the baby chose correctly and “Oh, there’s the shoe, but did you see the duck?” if the baby chose incorrectly. The tester did not reinforce the baby’s choices on test trials.

Half of the infants in the sound and word conditions were given six experimental test trials—that is, they were presented with a tray containing target and distracter objects and asked to chose while hearing the word or sound. In the word condition, the tester asked, “Can you get the *toma*?” In the sound condition, she asked, “Can you get one of these?” and then produced the sound holding the noisemaker in its bag near to the tray. The tester repeated the label or sound at least four times before the baby was given a chance to take an object from the tray. To test the possibility that training engendered a preference for the target object, the other half of the infants in each training condition were given six preference control trials after training. The tester presented the tray with the target and distracter and asked the baby to “get one” of the toys. If a baby did not respond or chose both objects simultaneously, the trial was run again later in the session.

In both test conditions, three of the test trials involved the same two objects seen during training (the training pair), while the other three involved new exemplars of the two kinds of objects, differing from the training set in color (the generalization pair). The order of these trials was randomly assigned. The target was on the left side of the tray for half the trials and on the right for the other half. As in training, the tester attempted to pace the questions in the word and sound conditions similarly. Follow-up coding of the session videotapes confirmed that the tester produced the sounds and words the same number of times during testing and that the two conditions did not differ in the overall length of the test session.

The baby’s choices were scored from the videotape by coders who did not know which object had been assigned as the target. The baby’s choice was coded as the object that he or she removed from the tray. The tester did not know which pair of objects had been assigned as the training versus generalization pair or which object had been assigned as the target. The trainer did not know the testing condition to which the infant had been assigned. The identity of the target object (coaster versus handle), label or sound used (*toma* versus *gombie* or beeper versus squeak), and sex of infant were counterbalanced within each training and testing condition with two exceptions: In the word control group, nine babies heard the word *toma* and seven heard the word *gombie*, and in the sound control group, nine babies had the handle as target and seven had the coaster as target.

**Results and Discussion**

For each infant, the total number of target choices (out of a possible six) was counted. Figure 1 summarizes these scores. Preliminary analyses revealed no
sex differences in performance and that the assignment of target object (handle versus coaster), particular word or sound (toma versus gombie or beeper versus squeak), and item introduced first during training (target versus distracter) did not affect babies’ choices on test trials. Therefore, the subsequent analyses collapsed across these dimensions. We first asked whether babies had learned the links between the word or sound and the object. Then, we tested whether babies generalized learning to the new exemplar of the object.

Did Babies Learn the Link Between the Word or Sound and the Toy?

The primary measure of learning is whether infants in the experimental condition chose the target object more often than chance would predict. Since the measure of comprehension was a two-item forced-choice test, if babies were choosing randomly they would select the target for half of the trials on average (that is, on three trials out of six). Babies in the word-experimental group chose the target object more often than chance would predict, mean = 3.9, t(15) = 2.91, p < .01, one-tailed. Babies in the sound-experimental group also chose the target object more often than chance would predict on experimental trials, mean = 3.7, t(15) = 1.90, p < .05, one-tailed. In the word and sound control groups, in contrast, babies chose randomly, mean (word-control) = 3.0, mean (sound-control) = 2.9, ts(15) < 1. Thus, the experience of training did not lead babies to prefer the target object to the distracter in either condition.

A second measure of learning is whether infants were more likely to choose the target object on experimental trials than on control trials. An analysis of variance with stimulus condition (sound versus word) and test condition (experimental versus control) as between-participants factors and test pair type (training versus generalization) as the withinsubjects factor revealed a main effect of test condition, F(1, 60) = 3.90, p = .05, reflecting the fact that infants selected the target object more often in the experimental condition than in the control condition. This pattern did not vary as a function of whether babies heard words or sounds. There was not a reliable effect of stimulus condition nor was there a reliable stimulus condition by test condition interaction, both Fs < 1.

As an alternative approach to the data, we next examined individual patterns of responding. We counted the number of babies in each stimulus and test condition who chose the target on more than three or fewer than three test trials (see Table 1). The scores of babies in the word-experimental group fell reliably above chance, p < .05 by binomial test, and the scores of babies in the word-control group did not differ from chance. The scores of babies in these two groups differed marginally, p = .12 by Fisher’s exact test. In the sound-experimental group, babies’ scores were only marginally above chance, p = .11 by binomial test. In the sound-control group the scores did not differ from chance. The scores of the sound-experimental and sound-control groups did not differ reliably, p = .42 by Fisher’s exact test. Thus, in contrast to the parametric tests, individual response patterns did not provide strong support for the conclusion that babies learned the sound-object links. Study 2 was conducted, in part, to seek converging evidence for this phenomenon.

The finding that babies may have learned the link between the sound and object is surprising given the unnatural feel of this condition from an adult standpoint. Moreover, although it is likely that 13-month-

Table 1  Number of Infants (of 16) in Each Condition Who Chose the Target on 0–2 or 4–6 Test Trials in Study 1

<table>
<thead>
<tr>
<th>Condition</th>
<th>Number of Target Choices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0–2</td>
</tr>
<tr>
<td>Word-experimental</td>
<td>3</td>
</tr>
<tr>
<td>Word-control</td>
<td>8</td>
</tr>
<tr>
<td>Sound-experimental</td>
<td>3</td>
</tr>
<tr>
<td>Sound-control</td>
<td>6</td>
</tr>
</tbody>
</table>
olds have encountered new words in joint attention labeling contexts, it seems less likely that they have encountered novel sounds in such contexts. As an indication of babies’ reactions to the sound condition, we coded the number of test trials on which the baby refused to respond and the amount of time between the beginning of the trial and the baby’s response. If babies found the sound condition unusual, they might be slower to respond and more likely to refuse to choose an object from the tray when they were presented with a sound than when they were presented with a word. However, the data indicate that this was not the case. Because the refusal scores were highly skewed, we tested for differences between the word and sound groups by means of a nonparametric test. A Mann-Whitney test revealed that babies in the sound-experimental group were no more likely to refuse to choose an object from the tray than were babies in the word-experimental group, \( z = 1.15, p = .22 \). Moreover, babies were as likely to fail to complete all trials following training with a word \( (n = 5) \) as they were following training with a sound \( (n = 6) \). The response latencies from babies in the experimental groups were entered into an analysis of variance with stimulus condition (word versus sound) as a between-participants factor and test trial type (training versus generalization) as a within-participants factor. The only reliable effect yielded by this analysis was a main effect of stimulus condition, reflecting the fact that babies in the sound-experimental group responded more quickly than did babies in the word-experimental group, \( \text{mean (sound)} = 10.0 \text{ s, mean (word)} = 12.8 \text{ s, F(1, 30)} = 7.96, p < .01 \). If babies found the sound condition odd, then we did not see evidence of it in their willingness to respond to our questions.

Did Babies Generalize Learning?

Next, we considered the question of whether babies generalized learning. Recall that in testing, babies were given three trials with the pair of objects they had seen during training (the training pair) and three trials with new exemplars of each of these objects (the generalization pair). The analysis of variance reported above revealed no significant effects of test pair type. However, there were two nonsignificant trends, one of test pair type, \( F(1, 60) = 2.72, p = .10 \), and one of training pair type by stimulus condition interaction, \( F(1, 60) = 3.49, p = .07 \). Inspection of the means suggests that these reflect a trend for babies to choose the target more often for generalization pair trials than for training pair trials and a trend for this pattern to be stronger in the sound condition than in the word condition (see Table 2). These trends did not vary as a function of whether babies were given control or experimental test trials, all other Fs < 1. Comparisons against chance for babies in the word- and sound-experimental groups supported the conclusion that babies chose the target somewhat more often on generalization trials than on training pair trials. Babies in the word-experimental group chose the target object 63% of the time on training pair trials, \( \text{mean} = 1.9, t(15) = 1.48, p = .08 \), one-tailed, and 67% of the time on generalization trials, \( \text{mean} = 2.0, t(15) = 2.74, p < .01 \), one-tailed. Babies in the sound-experimental group chose the target only 50% of the time on training pair trials, \( \text{mean} = 1.5 \), and 73% of the time on generalization trials, \( \text{mean} = 2.2, t(15) = 3.30, p < .005 \), one-tailed. Thus, babies in both the word-experimental and sound-experimental groups clearly generalized. It is not clear, however, why they performed less well on test trials involving the training pair.

In a final analysis, we explored one potential explanation for babies’ performing slightly better on generalization trials than on training pair trials: perhaps infants were bored with the training pair and more attentive on trials that involved a new pair of objects. The coding of infants’ willingness to respond on experimental trials did not provide strong evidence in support of this possibility. A Wilcoxon signed rank test revealed no difference in the number of refusals to respond on training pair versus generalization pair trials, \( z < 1 \). The analysis of response latencies yielded no reliable effects of test pair type, although there was a nonsignificant trend for infants to respond more quickly on generalization trials than on training pair trials, \( \text{mean (generalization pair)} = 10.6 \text{ s, mean (training pair)} = 12.2 \text{ s, F(1, 30)} = 2.71, p = .11 \).

Summary

To summarize, the results of this study support two conclusions. First, 13-month-olds learned the link

<table>
<thead>
<tr>
<th>Condition</th>
<th>Training</th>
<th>Generalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word-experimental</td>
<td>1.9 (.3)</td>
<td>2.0 (.2)</td>
</tr>
<tr>
<td>Word-control</td>
<td>1.6 (.3)</td>
<td>1.4 (.3)</td>
</tr>
<tr>
<td>Sound-experimental</td>
<td>1.5 (.2)</td>
<td>2.2 (.2)</td>
</tr>
<tr>
<td>Sound-control</td>
<td>1.3 (.3)</td>
<td>1.6 (.3)</td>
</tr>
</tbody>
</table>

*Marginally above chance, \( p = .08 \). Above chance, \( p < .01 \).
between the novel word and the object. Hearing the word that had been paired with the target object led infants to select that object at above chance rates, and in the absence of the word, babies chose randomly. Thus, babies’ choices on experimental trials did not result from training making the target object more interesting than the distracter. Rather, babies remembered the specific link between the word and the target object. This finding replicates those of Woodward et al. (1994) and Schafer and Plunkett (1998).

Second, the findings from the sound condition indicate that babies at this age may also be able to link novel sounds with objects during joint attention. Babies who heard a sound paired with the object chose that object at above chance rates during the test when they heard the sound and did not do so on control trials. The analysis of variance revealed no differences in the performance of the babies who heard words and the babies who heard sounds. After-the-fact coding indicated that babies in the word and sound groups did not differ in their propensity to respond to experimental questions. Because this is a relatively surprising finding and because the analyses of individual response patterns did not clearly support the conclusion that babies learned in this condition, we sought further evidence for this learning in Study 2.

In Study 2, we tested whether the sound learning seen in Study 1 would vary as a function of the age of the learner by testing older babies, 20-month-olds, as well as 13-month-olds. Perhaps the experimental context in which the sounds were introduced would lead any learner to select the object that had been paired with the sound. We gave babies strong pragmatic cues indicating that the sounds were intended to be communicative. The trainer established joint attention on the object with the baby, called attention to the object verbally, and then intentionally produced the sound as the baby watched. Similarly, the tester produced the sound after she had shown that she meant for the baby to listen to her when choosing objects from the tray. Babies may have assumed we meant to communicate and did the best they could to accommodate us. On this account, we might expect that older babies, who are presumably more pragmatically skilled, would also pick the target in the sound condition. On the other hand, if older babies have refined their expectations about which actions can serve as labels, they might be less willing to learn the sound-object links.

With these issues in mind, in Study 2 we compared 13- and 20-month-olds’ ability to learn sound-object correspondences during joint attention. To test the generality of the findings with 13-month-olds from Study 1, we introduced babies to new nonlinguistic sounds, produced by whistles rather than handheld devices.

STUDY 2
Method
Participants
In Study 2, there were 64 healthy, term infants recruited as in Study 1. All infants were acquiring English as their first language. There were 32 infants at 13 months (mean age = 13 months 18 days, range = 12 months 27 days to 14 months 2 days) and 32 at 20 months (mean age = 20 months 13 days, range = 19 months 9 days to 21 months 6 days). An additional 22 infants began the procedure but were not included in the final sample due to failure to complete all trials (11), parental interference (1), or experimenter error (10).

Design and Procedure
In this study, all infants were introduced to a novel nonlinguistic sound paired with a novel object during joint attention with an experimenter. The novel sound was produced by blowing on a small plastic instrument, which was hidden in the experimenter’s hand. There were two such sounds, a siren-whistle and a harmonica. Each of these was assigned to half the babies at each age. Otherwise, the training and testing procedures were identical to those used in Study 1: After nine training trials, half the babies at each age were given six experimental test trials and the other half were given six preference control trials. As in Study 1, separate experimenters administered training and testing, and these experimenters did not know the details of the portion of the experiment they did not administer. Babies’ choices were scored from videotape by coders who were not aware of which object had been assigned as the target. The object assigned as target (coaster versus handle) and the sound used (siren versus harmonica) were counterbalanced at each age and in each test condition. There were equal numbers of males and females in each age group and condition with the exception that there were nine males and seven females in the 20 month control group.

Results and Discussion
As in Study 1, the number of target choices was totaled for each infant. Figure 2 summarizes these
scores. Preliminary analyses revealed that the pattern of results was similar across the different target objects, sounds and training orders. There were no reliable sex differences in performance at either age. Subsequent analyses collapsed across these dimensions.

Did Babies Learn the Link Between the Sound and the Toy?

First, we asked whether infants selected the target object at above chance rates. In the experimental condition, 13-month-olds chose the target more often than chance would predict, mean = 4.3, t(15) = 4.20, p < .0005, one-tailed, but 20-month-olds did not differ from chance in their selections, mean = 3.3, t(15) < 1. Babies at both ages were at chance in their choices in the control condition, mean (13-month-olds) = 2.6, t(15) = 1.0, p = .83, one-tailed, mean (20-month-olds) = 2.9, t(15) < 1. Thus, once again, there was evidence that training did not engender a preference for the target object.

We then asked whether babies chose the target object more often on experimental trials than on control trials. This was the case overall: an analysis of variance with age group and test condition as between participants factors and test trial type as a within-participants factor revealed a main effect of test condition, F(1, 60) = 6.7, p < .05, reflecting the fact that overall babies chose the target more often on experimental than on control trials. There were no other reliable effects. Despite the difference between 13- and 20-month-olds in the comparisons against chance, the test condition by age group interaction was not reliable, F(1, 60) = 2.7, p = .10.

Analyses of individual patterns of response provided support for the conclusion that babies learned the sound-object link at 13 months, but not at 20 months. Table 3 provides the number of babies at each age who chose the target on fewer than three versus more than three test trials. The scores of 13-month-olds in the experimental condition fell reliably above chance, p < .005 by binomial test, and the scores of 13-month-olds in the control condition did not differ from chance. The response patterns of these two groups differed reliably, p = .002 by Fisher’s exact test. In contrast, neither the 20-month-old experimental group nor the 20-month-old control group differed from chance, and these two groups did not differ from one another, p > .9 by Fisher’s exact test.

As in Study 1, as an indication of babies’ reactions to these novel testing conditions, we coded the sessions of the babies in the experimental condition for the number of trials on which babies refused to respond and the latency between the beginning of the trial and the baby’s response. Thirteen- and 20-month-olds in

Table 3  Number of Infants (of 16) in Each Age Group and Condition Who Chose the Target on 0–2 or 4–6 Test Trials in Study 2

<table>
<thead>
<tr>
<th>Condition</th>
<th>Number of Target Choices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0–2</td>
</tr>
<tr>
<td>13 month-experimental</td>
<td>1</td>
</tr>
<tr>
<td>13 month-control</td>
<td>7</td>
</tr>
<tr>
<td>20 month-experimental</td>
<td>7</td>
</tr>
<tr>
<td>20 month-control</td>
<td>5</td>
</tr>
</tbody>
</table>
the experimental test condition were equally likely to refuse to respond to the tester’s request, Mann-Whitney, \( z < 1 \), and 13-month-olds took longer to respond to the experimenter’s request than did 20-month-olds, \( \text{mean (13 months)} = 9.0 \text{ s}, \text{mean (20 months)} = 7.3 \text{ s}, t(30) = 2.82, p < .01 \). Thus, this coding did not indicate that 20-month-olds were less willing to respond to the test questions than were 13-month-olds. However, differences in willingness to respond might have been obscured by general developmental differences in babies’ ability to coordinate a response in the task.

**Did Babies Generalize Learning?**

Next, we tested whether babies generalized learning. Table 4 summarizes babies’ performance on training pair and generalization pair test trials. The analysis of variance reported above yielded no reliable or marginal effects of test pair type. Thirteen-month-olds selected the target object at above chance rates on both kinds of test trials, choosing the target 77% of the time on training pair trials and 67% of the time on generalization trials, \( \text{mean (training pair)} = 2.3, t(15) = 4.6, p < .0005 \), one-tailed, \( \text{mean (generalization pair)} = 2.0, t(15) = 2.2, p < .025 \), one-tailed. Thus, as was the case for Study 1, 13-month-olds performed well on generalization trials. In this study, 13-month-olds’ performance on training pair trials was less ambiguous than in the first study. Twenty-month-olds in the experimental group were reliably above chance for neither training pair trials, \( \text{mean} = 1.4, t(15) < 1 \), nor generalization trials, \( \text{mean} = 1.8, t(15) = 1.3, p = .11 \), one-tailed.

**Summary**

In summary, 13-month-olds in Study 2, like those in Study 1, systematically selected the target object when they heard the sound on test trials, but did not do so on control trials. Moreover, they generalized what they had learned in training, choosing correctly when presented with new exemplars of the training objects. Twenty-month-olds, in contrast, did not choose the target object at above chance rates. Given the same training as younger babies, 20-month-olds did not show evidence of having learned the link between the sound and the object.

**GENERAL DISCUSSION**

We began by asking whether 13-month-olds could learn about sounds other than words in joint attention contexts. In the first study, we confirmed that 13-month-olds are able to learn a word–object link based on a brief introduction during joint attention. In the first and second studies, we found that 13-month-olds can also learn a sound–object link based on similar training. In both studies, 13-month-olds who had been introduced to a new sound–object pairing selected that object at above chance rates on the multiple choice test. Since babies did not systematically choose the target object on preference control trials, we conclude that babies’ choices reflected their memory for the specific link between the word or sound and the object rather than a preference for the object that had been paired with a novel word or sound. As was the case for word learning, babies generalized in the sound condition, choosing the target object when presented with novel exemplars of the objects they had seen in training.

In the second study, we found a developmental difference in babies’ acceptance of the sound–object pairings. Given the same training as the 13-month-olds, 20-month-old infants did not learn the link between the sound and the object. Although we did not test 20-month-olds’ word learning, a plethora of studies have documented that babies of this age readily acquire new word–object mappings during joint attention (e.g., Baldwin, 1991; Namy & Waxman, 1998; Ross, Nelson, Wetstone, & Tanouye, 1986; Schwartz & Leonard, 1980; Tomasello & Farrar, 1986; Woodward et al., 1994). Thus, 20-month-olds did not learn a sound–object mapping in a context in which they have been found to learn word–object mappings.

Taken together, these findings suggest that, in communicative contexts, younger infants are more open-minded as learners than are older infants. This pattern is in keeping with general patterns seen in other studies. As discussed above, Namy and Waxman (1998) have found that 18-month-olds who were introduced to novel gestural labels for objects accepted those labels more readily than did 26- and 27-month-olds. In addition, Acredolo and Goodwyn (1988; Goodwyn & Acredolo, 1993) found that young 1-year-olds can acquire and use arbitrary gestural labels for objects, but that use of these gestures declines toward the end of the second year of

### Table 4

<table>
<thead>
<tr>
<th>Condition</th>
<th>Training mean (SE)</th>
<th>Generalization mean (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 month-experimental</td>
<td>2.3 (.2)*</td>
<td>2.0 (.2)*</td>
</tr>
<tr>
<td>13 month-control</td>
<td>1.5 (.3)</td>
<td>1.1 (.2)</td>
</tr>
<tr>
<td>20 month-experimental</td>
<td>1.4 (.3)</td>
<td>1.8 (.3)</td>
</tr>
<tr>
<td>20 month-control</td>
<td>1.6 (.3)</td>
<td>1.3 (.2)</td>
</tr>
</tbody>
</table>

* Above chance, \( p < .025 \).
life. Our findings indicate that 13-month-olds can also link novel behaviors, in this case the use of novel noisemakers, with objects in a communicative context, and that toward the end of the second year of life, at 20 months, babies’ ability to learn these links declines.

We believe that our findings are useful chiefly because of the questions they raise. We turn now to the two most central of these: (1) On what basis did the 13-month-olds learn about the new sounds and words? (2) On what basis did the 20-month-olds avoid learning about the sounds?

On What Basis Did 13-Month-Olds Learn the Sound– and Word–Object Links?

The finding that 13-month-olds accept novel sounds in labeling routines is surprising in light of other evidence about infants’ abilities. Before their first birthdays, infants’ speech perception abilities are like those of the adults in their language community (see Goodman & Nusbaum, 1994, for a review). Moreover, evidence from other studies suggests that infants below the age of 12 months sometimes differentiate between novel nonlinguistic sounds and words under other testing conditions, specifically when words and sounds are played over an audio speaker rather than being produced by an interlocutor. Balaban and Waxman (1997), for example, report that recorded words increase 9-month-old infants’ propensity to categorize visual images more effectively than recorded tones do (but see Roberts & Jacob, 1991). Thus, young babies seem to have different expectations about words and sounds in some situations. Nevertheless, when the novel sounds were produced in a joint attention labeling routine, 13-month-olds treated them just as they treated labels.

Despite their ability to discriminate between speech and other sounds, then, 13-month-olds may be readily swayed by pragmatic evidence to accept a nonword signal in a labeling routine. There were a number of pragmatic cues that might have facilitated 13-month-olds’ learning about the correspondence between the word or sound and the object. The tester established joint attention with the baby and indicated the new object as she produced the word or sound. The tester demonstrated that the baby should listen to her when choosing an object from the tray. Although the sounds were novel, the experimenters intentionally produced them as the baby watched, and this behavioral cue may also have facilitated learning. We do not know from the current findings which, if any, of these cues were necessary for 13-month-olds to learn the word–object and sound–object links.

If babies were relying on cues to the experimenters’ communicative intentions, then this raises the question of the range of such cues that 13-month-olds can use to inform learning. We presented the sounds and words in an ostensive labeling routine, but, as Tomasello and colleagues have demonstrated, ostension per se is not necessary for word learning in older toddlers (Tomasello & Akhtar, 1995; Tomasello & Barton, 1994), and babies frequently encounter new words in contexts other than this labeling routine (Tomasello & Kruger, 1992). In learning new words, older toddlers can draw on several sources of pragmatic information, including the speaker’s focus of attention, the novelty of an item in the discourse context, overt cues as to whether an act was accidental or purposeful, and behavioral cues to an intention that is not fulfilled (for reviews see Baldwin, 1995; Baldwin & Tomasello, 1998; and Tomasello, 1995). A critical project for future research is to investigate the range of cues that young 1-year-olds draw on in learning communicative signals.

An alternative possibility is that 13-month-olds did not make use of behavioral cues to the experimenters’ communicative intentions in learning the word– and sound–object correspondences. In establishing joint attention, we ensured that babies were looking at the target object as they heard the new word or sound. It is possible that 13-month-olds learned the word– and sound–object correspondences based on this contiguity alone. In fact, essentially this model has often been proposed to account for babies’ first learning about words (see e.g., Lock, 1980; Piaget, 1962). The argument is that initially babies do not think of words as parts of communicative acts, but rather learn words as classically conditioned associates of situations and events.

A number of studies have ruled out this model as accounting for word learning in somewhat older babies, 18-month-olds. These studies have provided compelling evidence that by the time babies are 18 months of age, they understand that words are referential. When learning a new word, babies at this age seek out and use information relevant to the speaker’s communicative intentions, and they do not map words onto referents when such cues are absent. They can do this even when the speaker’s focus of attention differs from their own (Baldwin, 1991, 1993a, 1993b, 1995) and when it requires discriminating between accidental and purposeful behaviors (Tomasello, 1995; Tomasello & Barton, 1994).

Less is known about this kind of knowledge in babies younger than 18 months, but what is known suggests that even 12-month-olds use and understand words as communicative signals, not as associates.
For one, the associative model cannot easily explain patterns of language production at this age. In their analyses of word production in young 1-year-olds, Huttenlocher and Smiley (1987) explored the claim that babies’ first word meanings are complexive—
that is, that words are general associates of situations and events rather than being referential. They found that babies’ first uses of object labels were nearly always extensions of the label to objects within a basic level or superordinate category, rather than uses in the presence of items typically associated with the object. When children did use a word in the absence of its referent, there was contextual evidence that they were not using the word complexively but instead were requesting an absent object or commenting on a relation between objects.

In addition, recent research reported by Baldwin and Tomasello (1998) indicates that 12-month-old infants understand something about the communicative nature of words in that, like older babies, they respond to new words by checking the speaker’s line of regard and do this more often when an utterance is ambiguous than when the speaker’s intended referent is clear. Moreover, before they have words, babies show signs of understanding aspects of reference and communication (Bates et al., 1979; Bretherton, 1991). It is likely that babies recruit this knowledge as they begin to make sense of words. Nevertheless, there are many unanswered questions concerning young 1-year-olds’ understanding of the distinct role played by words in acts of communication. It is possible that infants at this age understand some aspects of name–referent relations, but not others.

These considerations raise a fundamental question. Is it correct to conclude from our findings that 13-month-olds construed the novel sounds as names? This question rests on the issue of whether 13-month-olds understand the specific ways that names relate to their referents, that is, that they are symbols used to refer to objects and events. Beyond the suggestive evidence just discussed, little is known about this knowledge in 13-month-old infants. For this reason, we believe it is premature to conclude that the 13-month-olds construed the sounds as names, in the adult sense of the word. Our findings indicate that as we investigate young 1-year-olds’ understanding of name–referent relations, we should also investigate their understanding of the relations between non-word signals and objects and events. As babies’ understanding of referential symbols develops, their understanding of the category of signals that serves this function may also be undergoing change.

On What Basis Did the 20-Month-Olds Avoid Learning about the Sounds?

Despite the pragmatic cues that seem to have influenced 13-month-olds, 20-month-olds did not learn the link between the novel sounds and the objects. What accounts for the differences in performance between the 13- and 20-month-olds? While our findings do not provide conclusive evidence on this question, when considered in combination with the findings of other studies they do provide some clues. To begin, we think it is unlikely that 20-month-olds have lost their sensitivity to the pragmatic cues that were present in the sound-learning situation. In fact, as discussed above, babies at this age show pragmatic skill when learning new words. Instead, 20-month-olds’ resistance to learning the sound–object links likely reflects an awareness of some of the ways in which the sounds differed from names. The question, then, is which difference or differences led to this resistance.

One possibility is that the perceptual difference was key. Perhaps by 20 months, babies have acquired the strong expectation that spoken words serve as names, and thus resist learning about signals which do not have this form, even when other evidence is consistent with their functioning as names. Namy and Waxman’s (1998) findings support a similar conclusion for somewhat older babies. The 26- and 27-month-olds in their studies resisted learning gestural labels even given pragmatic and syntactic cues that the gestures were being used as labels.

There is a second factor that may have played a role. The sounds in these studies lacked a critical feature of names in that they were not treated as grammatical units. Instead, they occurred outside of utterance boundaries. Perhaps 20-month-olds, who are on the verge of producing two-word utterances and who seem to understand some aspects of syntax (Hirsch-Pasek & Golinkoff, 1996), expect that names will occur as parts of sentences. This possibility could account for a difference in the findings of the current studies and those of Namy and Waxman (1998). In their studies, gestural labels were embedded in a sentence, for example “We call this [gesture].” Given this training, 18-month-olds learned the new gestural label, whereas babies of nearly the same age, 20-month-olds, failed to learn the sound–object link in our second study. Thus, 20-month-olds might accept a novel sound signal if it were embedded in a sentence. This possibility requires further investigation.

Even older language users can be convinced that non-word signals are names. Although a typical adult English speaker will not often encounter manual signs, labels involving phonemic clicks, or unpro-
nounceable names like the one adopted by the artist formerly known as Prince, he or she could be persuaded that those things are names. Namy and Waxman (1998) found that older babies are similarly flexible. Although the 27-month-olds in their studies did not readily accept gestural labels based on a brief training session, they accepted them after extensive training and practice with the use of a gestural label before they were presented with the experimental items. This familiarization included two manipulations that provided evidence that the gestures were communicative—babies were encouraged to produce as well as respond to the gestures, and the gestures were presented in an extended dialogue between the experimenter and a puppet. In addition, babies were rewarded for choosing the correct item during familiarization. A direction for future studies is to explore the kinds of experiences that are important in getting older babies to accept non-word signals.

In sum, the available evidence suggests that as babies get older they require that possible names have more of the features that are typical of words from a mature standpoint. Given greater contextual support, babies who are initially resistant can be convinced to accept a non-speech signal as a name. The manipulations that may be most powerful in getting learners to accept novel signals as names are those that support the interpretation that the signal is intended to serve as a communicative tool.

Conclusions

In conclusion, our findings indicate that 13-month-olds accept a wider range of signals in communicative contexts than do 20-month-olds. The fact that 13-month-olds learned the links between novel sounds and objects may seem surprising given that by the time they are 12 months of age, babies have considerable knowledge about the perceptual class of spoken words. We suggest that infants may initially base their developing category “name” on features of behavior relevant to communicative intent, rather than on a particular perceptual class. This makes sense when we consider that not all babies learn spoken languages, and those that learn sign languages show no delay in attaining language milestones (Newport & Meier, 1985). Moreover, the features of action that are relevant to communicative intent are closer to the heart of what it means for a signal to be a name than are the particular channels by which meaning is conveyed.

ACKNOWLEDGMENTS

This research was supported by grants from the John Merck Fund and the Robert R. McCormick Tribune Foundation to the first author. Pilot work was supported by a National Institutes of Health postdoctoral fellowship to the first author at Cornell University. We thank Dare Baldwin, Laura Namy, and two anonymous reviewers for their insightful comments on an earlier version of the manuscript and the parents and infants who volunteered their participation. Portions of the data were presented at the 1997 biennial meetings of the Society for Research in Child Development in Washington, DC.

REFERENCES


