At 6–9 months, human infants know the meanings of many common nouns

Elika Bergelson[a,b,1] and Daniel Swingley[a,b]

[a]Department of Psychology and [b]Institute for Research in Cognitive Science, University of Pennsylvania, Philadelphia, PA 19104

Edited by Willem J. M. Levelt, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, and approved December 8, 2011 (received for review August 17, 2011)

It is widely accepted that infants begin learning their native language not by learning words, but by discovering features of the speech signal: consonants, vowels, and combinations of these sounds. Learning to understand words, as opposed to just perceiving their sounds, is said to come later, between 9 and 15 mo of age, when infants develop a capacity for interpreting others' goals and intentions. Here, we demonstrate that this consensus about the developmental sequence of human language learning is flawed: in fact, infants already know the meanings of several common words from the age of 6 mo onward. We presented 6- to 9-mo-old infants with sets of pictures to view while their parent named a picture in each set. Over this entire age range, infants directed their gaze to the named pictures, indicating their understanding of spoken words. Because the words were not trained in the laboratory, the results show that even young infants learn ordinary words through daily experience with language. This surprising accomplishment indicates that, contrary to prevailing beliefs, either infants can already grasp the referential intentions of adults at 6 mo or infants can learn words before this ability emerges. The precocious discovery of word meanings suggests a perspective in which learning vocabulary and learning the sound structure of spoken language go hand in hand as language acquisition begins.

word learning | cognitive development | infant cognition

Most children do not say their first words until around their first birthday. Nonetheless, infants know some aspects of their language's sound structure by 6–12 mo: they learn to perceive their native language's consonant and vowel categories (1–4), they recognize the auditory form of frequent words (5, 6), and they employ these stored word forms to draw generalizations about the sound patterns of their language (7, 8), using cognitive capacities for pattern finding (9, 10). Although this learning about regularities in speech reveals impressive perceptual and analytical skill, it is generally accepted that young infants do not know the meanings of common words. Indeed, although some experimental work has shown that young infants can associate syllables with individual objects after laboratory training (11), prior experimental tests have failed to detect understanding of common native-language words before around 12 mo (12).

Infants are, on the whole, proficient and precocious learners in other domains (13), so why would learning word meanings be difficult for them? The most prominent hypothesis is that true word learning is possible only when infants can grasp a speaker's referential intentions and understand language as a motivated, communicative activity (14–17). Evidence that infants begin to understand other humans as intentional agents only at around 9–10.5 mo has been argued to explain the earliest emergence of word learning shortly thereafter (17). Understanding reference is said to be necessary for word learning because the natural conditions of language use do not support the simple associations that underlie, for example, a trained dog's ability to fetch specific toys on command (18). The statistical connection between instances of words and the details of infants' observations is tenuous: parents do not reliably say "doll" in the exclusive presence of dolls, and they say "Hi, I'm home!" more often than "Daddy is moving through the doorway!" (19). Furthermore, words (excepting proper names) refer to categories, not individuals, and the learner must discover each category and its boundaries. Thus, although infants can link "mommy" with films of their mother, these labels do not indicate that infants have induced the relevant category (20). Because of these complexities inherent in language understanding, the predominant view is that word learning is possible only when children can surmise the intentions of others enough to constrain the infinite range of possible word meanings, a skill believed to develop gradually after 9 mo (17). Until that age, infants' native language learning is held to be restricted to speech signal analysis (21).

In the present research, we examined young infants' knowledge of word meaning using a variant of a task called "language-guided looking" or "looking-while-listening" (22, 23). In this method, infants' fixations to named pictures are used to measure word understanding. Infants are presented with visual displays, usually of two discrete images, one of which is labeled in a spoken sentence such as "Look at the apple" (24, 25). In our variant, the parent uttered each sentence, prompted over headphones with a prerecorded sentence, ensuring that infants (n = 33) heard the words pronounced by the familiar voice of their parent. Each infant experienced two kinds of trial: trials with two discrete images (paired-picture trials) and trials with a single complex scene (scene trials) (*Materials and Methods*; Fig. 1; Fig. S1; and Table 1).

Two word categories were tested: food-related words and body-part words. Paired-picture trials (n = 32) presented one image from each category (e.g., apple–mouth), and scene trials presented one image (n = 16) depicting several category members together (e.g., a full-length picture of a boy, a close-up of a face, or a table with food-related items on it). All pairs and scenes occurred in multiple instantiations within and between infants (e.g., there were two different "apple" photos and two different "full-body" photos) (Fig. S1).

Results

Children who understood a word were expected to fixate on the target picture more upon hearing it named. To evaluate this, the two trial types were analyzed separately because their demands are distinct and the ideal analytical methods are different, particularly in how to best correct for infants' preferences for individual pictures. (An analysis of both trial types using the same dependent measure is given in *SI Text* and in Table S1.)

For both analyses, the posttarget analysis window extended from 367 to 3,500 ms after the onset of the spoken target word

Author contributions: E.B. and D.S. designed research; E.B. performed research; E.B. analyzed data; and E.B. and D.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: elikab@psych.upenn.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1113380109/-/DCSupplemental.

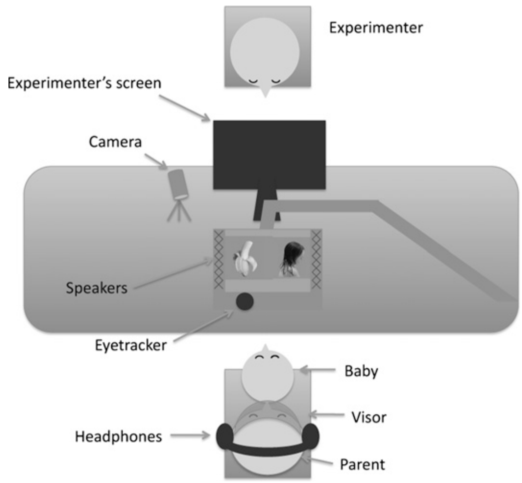

Fig. 1. Experimental setup. The child sat on her parent's lap and was presented with images and sounds from a computer equipped with an eye tracker and speakers. The experimenter sat behind a screen and was not visible to the infant. The experimenter controlled presentation of stimuli and monitored the child on a live-feed camera. A backup video recording of the session was made to allow for confirmation of the validity of gross characteristics of the eye-tracking data stream. The figure shows an example of images presented on a paired-picture trial testing "banana" or "hair."

(Fig. S2). The 367-ms starting time is the standard in the field and allows for the time required to initiate an eye movement in response to the speech signal; earlier fixation responses are unlikely to be reactions to the signal. The 3,500-ms window offset is later than the 2,000-ms offset that is typically used with older children. It was implemented here because, in previous research testing the 12- to 24-mo age range, children were discovered to be faster with increasing age and experience with words ; thus we assumed that younger children would require more time to demonstrate recognition.

For paired-picture trials, word recognition performance was operationalized as a difference of fixation proportions: for paired pictures A and B, the fixation to picture A relative to B when A was the target, minus the fixation to A when A was the distracter.* For example, given the pair of images hair–banana, a child's performance was given as how much she looked at "hair" when it was named as the target, relative to her looking at "hair" when "banana" was the named target. Positive difference scores are consistent with word understanding. This pair-based analysis corrects for infants' picture preferences without relying on infants looking during the portion of the trial before the mother speaks. A total of 26 of the 33 6- to 9-mo-olds (M = 7.44 mo, SD = 1.26) showed a positive mean difference score (all 33 subjects: M = 0.074, P = 0.0005, Wilcoxon test; P = 0.001, binomial test). Children showed positive performance on six of eight item pairs (M = 0.065, P = 0.020, Wilcoxon test). Fig. 2 illustrates these results showing the 6- to 7-mo-olds and 8- to 9-mo-olds separately.

For scene trials, word recognition performance was operationalized as the proportion of target looking upon hearing the target word (367–3,500 ms post target onset), minus the proportion of target looking before hearing the word (from when pictures were displayed until just before target onset) (Fig. S2). This analysis corrects for fixation preferences within portions of the scene, preventing an advantage for targets that occupy more of the scene. A total of 22 of 33 infants showed a positive proportion of target looking; performance was statistically significant over subject means (M = 0.042, P = 0.020, Wilcoxon test) (Fig. 2). Infants showed positive performance on 12 of 16 items; performance over item means fell short of significance (M = 0.023, P = 0.058, Wilcoxon test) (Fig. 2).

Infants of 8–9 mo are known to be capable of learning the sound forms of words and retaining them over long intervals (6), whereas infants of 6–7 mo have thus far shown much more limited word-form knowledge (26). It is therefore of interest to determine whether the present findings are due only to the older children in the sample.

This was not the case. Considering the subject-means data in Fig. 2, it is clear that, for both types of trial, at both ages most children scored above zero. On the paired trials, performance was significantly above chance levels in each age group (6–7 mo: M = 0.058, P = 0.027; 8–9 mo: M = 0.082, P = 0.0052). In the scene trials, evidence of recognition was strong in 6- to 7-mo-olds (M = 0.068, P = 0.015) but less strong in 8- to 9-mo-olds (M = 0.013, P = 0.27) although these age groups were not significantly different from one another (paired-picture trials: M = 0.036, P = 0.37; scene trials: M = –0.067, P = 0.093). The apparently inferior performance of the 8- to 9-mo-olds on the scene trials may be traced to their tendency to fixate the "eyes" and "face" regions before the mother named any pictures (Fig. 2*H*). This tendency, which may have its origins in previously observed developmental changes in infants' attention to social stimuli (27), did not interfere with infants demonstrating recognition in the paired-picture context, but impeded accurate measurement of 8- to 9-mo-olds' word recognition in the scenes containing faces.

A correlational analysis over the 6- to 9-mo range indicated that performance on paired-picture trials was not correlated with age (τ = 0.042, P = 0.75). Performance on scene trials was negatively correlated with age (τ = –0.25, P = 0.039); however, excluding the two words "eyes" and "face" (or just "eyes" or just "face"), the correlation of performance with age was negligible (τ = 0.015, P = 0.91). The lack of a positive correlation with age, and the consistently strong performance of the 6- to 7-mo-olds, confirm that the word recognition performance of the 6- to 9-mo-old sample cannot be attributed to the older children alone.

The time course of infants' picture fixation is shown in Fig. 3, which presents data from the 33 6- to 9-mo-olds, as well as results from three older groups of children tested in the same procedure (*SI Text*). Children initially fixated the target and distracter equally (averaging over items); then, upon hearing the target word, they shifted gaze to the named picture, thenceforth remaining above chance levels of target looking over most of the trial. Although most infants showed knowledge of the meanings of most items, target fixation performance at 6–9 mo and even at 10–13 mo was below levels shown by slightly older children (Fig. 3). The data suggest a discontinuity in performance at around 14 mo: performance was stable with respect to age before 14 mo and was substantially better afterward. We speculate that this phenomenon reflects the acquisition of linguistic knowledge and the development of social or other communicative skills, a topic that we return to in the *Discussion*. A more detailed analysis of the developmental pattern of results is given in *SI Text*.

Two additional measures of 6- to 9-mo-old infants' word knowledge were obtained from their parents: the MacArthur–Bates Communicative Development Inventory (CDI), which is a vocabulary checklist originally intended for children 8 mo and older (28), and an item exposure survey asking parents to estimate how often their child heard our target items on a scale from "never" to "several times a day." The modal response from parents on the CDI was that their child did not know any of the 395 words on the inventory; furthermore, no parent reported that his or her child was producing any of the words tested in our experiment (Table 2). Parental ratings of item exposure did not

*Because the pairs are yoked, for each pair A–B, the values of this measure for A and for B are arithmetically redundant (the value for A is necessarily the complement of the value for B). Thus, the item results are presented item pair by item pair (Fig. 2).

Table 1. Target items

Body items	Ear, eyes, face, foot, feet, hair, hand, hands, leg, legs, mouth, nose	
Food items	Apple, banana, bottle, cookie, juice, milk, spoon, yogurt	
Scene body pictures	Face: eyes, hair, ear, mouth	Body: face, legs, hands, feet
Scene food pictures	Tabletop 1: milk, juice, spoon, banana	Tabletop 2: apple, cookie, yogurt, bottle
Paired-picture pairs	Apple–mouth, banana–hair, bottle–leg, cookie–eyes, juice–nose, milk–foot, spoon–ear, yogurt–hand	

The upper half of the table lists all food and body-part target words that were tested. The lower half of the table lists the target words for each trial type and indicates the image (scene trials) or yoked pair (paired-picture trials) within which each target word's referent picture occurred.

correlate significantly with scene or paired-picture trial performance (scene: $\tau = -0.022$, $P = 0.65$; paired-picture: $\tau = 0.0079$, $P = 0.83$, Kendall's correlation test). Thus, there was no indication that the knowledge that infants revealed in the experiment was apparent to the infants' parents.

Discussion

The present findings provide an important contribution to our understanding of language acquisition, showing that, by 6–9 mo, infants have already begun to link words with their referents over a range of food and body-part terms. The two trial types indicated complementary abilities. Success on paired-picture trials showed that infants could understand words whose referents were presented in extremely stripped-down contexts: for example, a nose without the eyes or mouth. By contrast, performance on the scene trials showed that infants could differentiate at least some of the tested words from related alternatives: for example, an infant who heard "banana" and then looked at the banana in a tabletop scene containing several objects from typical mealtime contexts provided evidence of distinguishing the "banana" from semantically related objects. On both trial types, infants demonstrated abstraction from their experience, in that the pictures that we selected were not adapted to children's individual experiences in any way. Each word was tested on three trials (twice in picture pairs and once in a scene). In each case, the image was different and the spoken words, produced "live" by the parent, were never exactly the same. Infants therefore showed generalization in the way that language normally demands: common nouns refer to categories, not a specific instance, and spoken words are consistent phonologically (in their sequence of consonants and vowels) but not acoustically (e.g., in the details of each word's pitch and duration).

Several features of these results merit further exploration. The present study showed that 6- to 7-mo-olds understand something of the meaning of at least some words for foods and body parts. The results do not establish the size of infants' vocabularies, nor do they prove exactly which of the tested words each child knew. The item-wise histograms in Fig. 2 show that mean performance was greater than zero and that there was roughly normally distributed variation around that mean. Some variation is undoubtedly due to chance, but it is also quite likely that there were some words that some children did not know. After all, children's experience is various, and the items tested were not calibrated to each child's history. Confirming knowledge of particular words would require a design devoting more test trials to a smaller number of different words. The fact that the present results are statistically significant in analytical models that include subjects and items as random factors (*SI Text* and Table S1) ensures that the conclusions are not based on just a few words or a few children; however, the present design does not support strong conclusions about individual words or individual children.

We have focused here on the youngest children in the sample because the assumption that young infants learn about speech sounds but not about word meaning has predominated for at least 25 y. However, we also found a surprising developmental pattern when considering a much wider age range, from 6 to 20 mo. Previous studies have not explored this age range using consistent materials and procedures. Doing so allowed us to document little change in performance from 6 to 13 mo and then a substantial improvement at around 14 mo, with some gains observed thereafter. This pattern raises two questions: Why was there no apparent improvement in the youngest half of the sample? And what caused the change in performance at around 14 mo?

One explanation of the seeming lack of developmental change from 6 to 13 mo is that it is artifactual. First, informally speaking, infants of about 9–12 mo are often more difficult to evaluate in experimental procedures than younger or older children, seeming more distractible and more likely to become fussy. This might be a product of infants' eagerness to exercise their rapidly changing motor skills and a consequent lack of attention to the experimental materials, thus masking underlying developments in linguistic knowledge or ability. Of course, we might also speculate that a lack of attention to language among some 8- to 12-mo-olds who are prioritizing some other cognitive domain indicates something true about their mental life outside the laboratory as well (29).

An alternative account for the apparent lack of improvement in performance from 6 to 9 mo is tied to an explanation for the material elevation in performance around 14 mo. We speculate that younger and older children might learn words, interpret sentences, and conceptualize the experimental situation in quite different ways. In the domain of word learning, infants may be restricted to relatively inefficient learning strategies, a point we return to below. Infants are also likely to be much less sophisticated in their interpretation of the sentences in which our target words were embedded. Knowledge of English sentences and English syntax make a sentence such as "Can you find the juice?" interpretable, not just as a string of syllables followed by a familiar word, but as a hierarchy of syntactic phrases unfolding in somewhat predictable ways. Understanding the sentence is likely to make the target words easier to grasp. It is also possible that at around 14 mo many infants begin to catch on to the nature of the experiment, regarding it as a repetitive game of object searching, and that this helps them to focus their attention on the task. These factors could explain the apparent lack of change in performance from 6 to 9 mo. The relatively low levels of target looking at the younger ages may have been possible with basic knowledge of word meaning but not a richer understanding of the words, a better grasp of the sentences, or a helpful conceptualization of the task. Of course, these comments are speculative: although developmental improvements along these lines are to be expected at a gross level, the present experiment was not intended to evaluate these possibilities.

How do 6- to 7-mo-olds learn words? The key finding of this study is that infants recognized words and demonstrated through their behavior that they knew something of the meaning of those words. Because the study involved no training, the result implies that 6- to 7-mo-olds learned the words through their daily experience. The predominant account of word learning holds that intention reading is a fundamental prerequisite. One interpretation of our results is that the relevant social-cognitive skills are available earlier in development than previously assumed. This is consistent with recent evidence of young infants' early

PSYCHOLOGICAL AND COGNITIVE SCIENCES

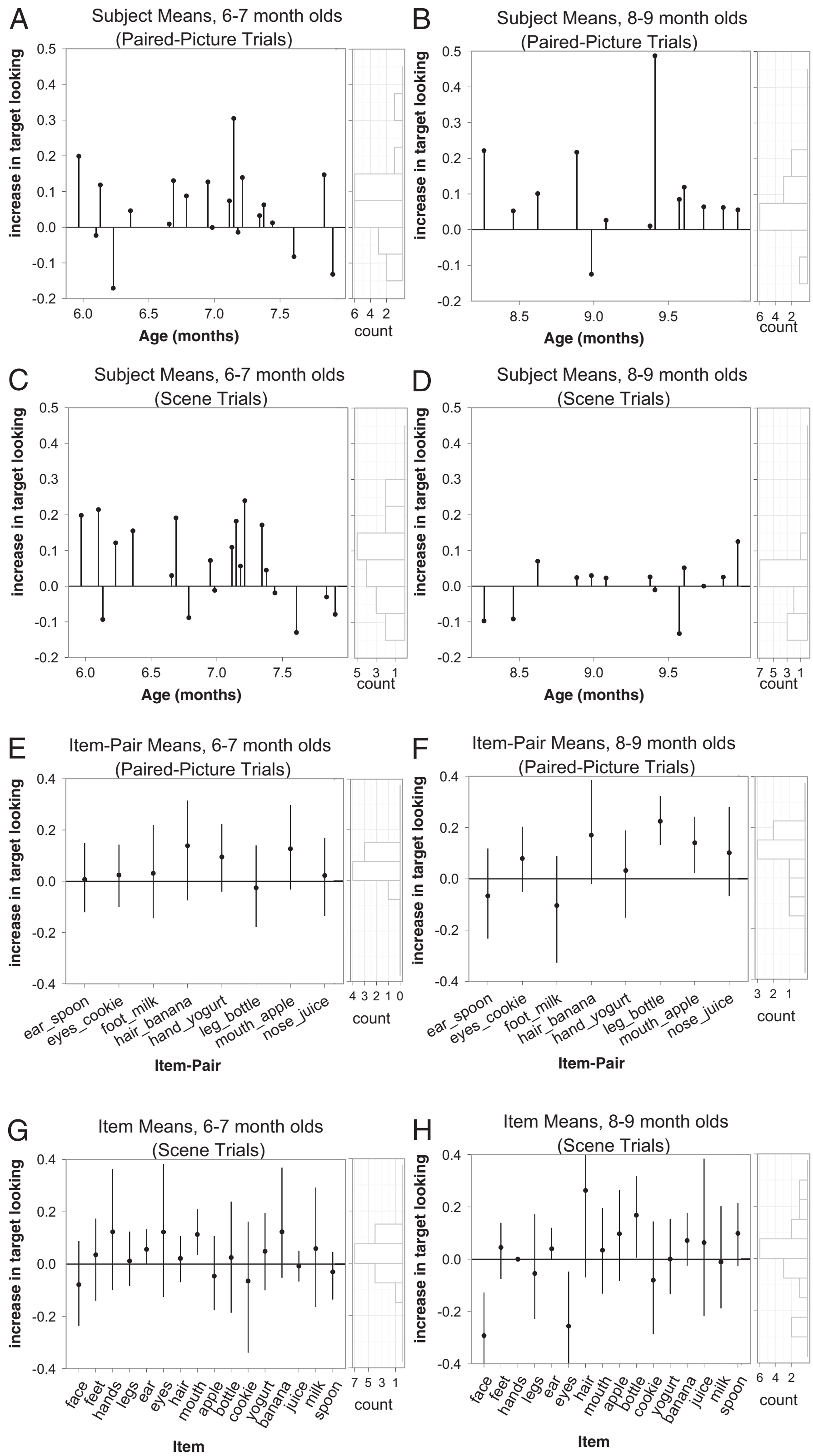

Fig. 2. Subject and item-pair means for 6- to 7- and 8- to 9-mo-olds. All data (*A–H*) were calculated over a window from 367 to 3,500 ms post target word onset. Subject mean difference scores are shown for paired-picture trials for 6- to 7-mo-olds (*A*) and for 8- to 9-mo-olds (*B*). Subject mean increases in target looking, corrected for baseline looking, are displayed for scene trials for 6- to 7-mo-olds (*C*) and for 8- to 9-mo-olds (*D*). Item-pair mean difference scores are shown for paired-picture trials for 6- to 7-mo-olds (*E*) and for 8- to 9-mo-olds (*F*). Item mean increases in target looking, corrected for baseline looking, are given for scene trials for 6- to 7-mo-olds (*G*) and for 8- to 9-mo-olds (*H*). (*E–H*) Error bars represent bootstrapped nonparametric 95% confidence intervals. On the right of each subplot is a histogram of the responses in the main plot; all histograms show more positive than negative responses for each subset of subjects and of item pairs.

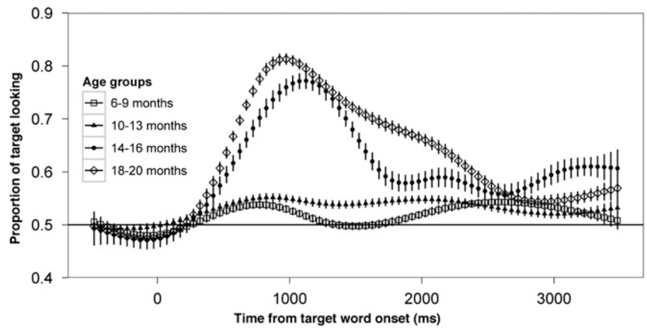

Fig. 3. Time course of infants' picture fixation on paired-picture trials, averaged over infants in four age groups. The ordinate shows the mean proportion of infants who were looking at the named (target) picture at each moment in time. Error bars indicate SEMs, with means computed over subjects in each age range. At all four ages, target fixation rose from about 0.50 (chance) shortly after the onset of the spoken word. Overall, accuracy in fixating the named picture increased with age across the age groups. See *SI Text* for further details.

sophistication in social cognition (30, 31). For example, 6-mo-olds follow an adult's gaze when the adult signals an intent to communicate, but not otherwise (32). Infants' appreciation of their parents' referential intentions, although perhaps limited at 6 mo, may narrow the range of potential word referents, making word learning possible.

Some theorists argue, by contrast, that domain-general cognitive capacities suffice for word learning, without invoking any understanding of the referential nature of words or of others' intentional states (33, 34). This association-based kind of account would explain our results by appealing to infants' ability to track consistent features of their physical environment when hearing words, progressively hypothesis testing (with or without an active intent to learn) until the referents of the words were isolated. In considering this hypothesis, we note that many of our target words did not refer to distinct, bounded objects, which have been suggested as good defaults for such hypotheses (35). Infants here performed well across an array of items containing well-delineated objects (e.g., cookie, bottle), amorphous substances (e.g., milk, juice), and unclearly bounded body parts (e.g., nose, hands) (see Fig. S1 for visual stimuli and regions of interest). This does not rule out associationist theories of word learning, but it raises the stakes; an association-based account (or indeed any account) cannot rely strongly on a bounded-object bias to limit its search space.

Our results do not imply that infants have an understanding of words that is comparable to that of adults, or even to older children's. Although infants did generalize from their experience to the particular photographs that we happened to choose, their categories of each object may nevertheless be different from those of adults, possibly by being based more strongly on perceptual attributes than on functional ones, for example. It is also not clear that our target words, which grammatically were all nouns of English, are interpreted by infants using a linguistic system that includes "noun" as a category. Finally, although there is some evidence that young infants' knowledge of the sound forms of words is accurate [e.g., infants recognize words more readily when the words are pronounced correctly than when their forms are altered (26)], the present study did not test the details of infants' speech perception.

Table 2. Summary statistics for CDI completed by 30 of 33 caregivers

Measure (no. of items)	Range	Mean	Median	Mode (n = 30)
CDI understood (n = 395)	0–71	19	7.5	0 (n = 7)
CDI said (n = 395)	0–7	0.83	0	0 (n = 20)
Test items from CDI understood (n = 16)	0–10	2.10	1	0 (n = 13)
Test items from CDI said (n = 16)	0–0	0	0	0 (n = 30)

Our findings indicate that native-language learning in the second half of the first year goes beyond the acquisition of sound structure. The fact that even 6- to 7-mo-olds learn words suggests that conceptual and linguistic categories may influence one another in development from the beginning (36) and that aspects of meaning are available to guide other linguistic inferences currently thought to depend only on distributional analysis of phonological regularities (37, 38). Understanding word meaning could also support the acquisition of syntax by guiding infants' inferences about how nouns and words from other word classes are placed in sentences. Precocious word learning also helps explain why hearing-impaired infants identified for fitting with cochlear implants before 6 mo reveal better language skills at 2 y than children identified just a few months later: 6-mo-olds who can hear are already learning words (39).

More generally, these results address one of the central mysteries of language acquisition: how children demonstrate proficiency in their native language so rapidly, typically speaking hundreds of words by the age of 2 y. Part of the solution, it appears, is that learning begins very early in life, hidden from view; even before they begin to babble, infants understand some of what we tell them.

Materials and Methods

Participants. Subjects were 33 6–9 mo-old infants (M = 7.45 mo, R = 5.99–10.00 mo, 19 female). Infants were recruited from the Philadelphia area by mail, e-mail, phone, and in person. All children were healthy, carried full-term, and heard 75% English or more in the home. None had a history of chronic ear infections. A second set of 50 children ranging in age from 10 to 20 mo (M = 14.1 mo, R = 10.13–20.85 mo, 27 female) were also recruited, using the same methods and criteria, to participate in the developmental portion of the study reported in *SI Text* and Fig. 3.

Materials. On each trial, parents spoke a single sentence to their child. To do this, they repeated verbatim a prerecorded sentence that they heard over headphones. These prerecorded sentences were produced by a native English-speaking woman and recorded in a sound-treated room. Each sentence that was presented to parents followed one of four different formats: "Can you find the *X*?", "Where's the *X*?", "Do you see the *X*?", and "Look at the *X*!", where *X* stands for the target word (only one sentence format was used per item) (Fig. S2). Sentence formats varied across trials pseudorandomly. The sentences were uttered at a slow speed, about four syllables per second, with slightly exaggerated intonation, which parents were asked to emulate. The recorded sentences were 1–1.5 s in duration and were presented to parents at loudness levels of 31.5–33.75 dB. Pretesting determined that speech presented at this volume over the closed-ear headphones was audible only to the parent.

Visual stimuli were displayed on a 34.7- × 26.0-cm LCD 75 dpi screen. On paired-picture trials, two 16.9- × 12.7-cm photos were displayed on the right and left side of the screen on a gray background; the side of presentation was counterbalanced across trials and trial orders. There were 32 such photos, namely two instances each of 16 items (Table 1; Fig. 1; and Fig. S1). Photos were edited so that their relative size and brightness were approximately equivalent.

On scene trials one photo was displayed in the center of the screen on a gray background. The photos were of people (whole body, clothed), faces, and tabletops with four food items on them (Table 1 and Fig. S1). There were two photos of faces (widths of 21.84 and 25.40 cm), two photos of bodies (widths of 16.08 and 20.93 cm) and two photos of tabletops: one with milk, juice, a spoon, and a banana on it and another with a cookie, an apple, a bottle, and yogurt on it (all widths 34.67 cm). Infants saw only one of these images on a given scene trial.

In addition, on every eighth trial, a 2-s movie—featuring colorful shapes and smiley faces flitting around the screen accompanied by a whistling sound—was played to maintain infants' interest.

Apparatus and Procedure. Infant visual fixation data were collected using an Eyelink CL computer (SR Research), which provides an average accuracy of 0.5°, sampling from one eye at 500 Hz. It operates using an eye-tracking camera at the bottom of the computer screen; no equipment is mounted on the child's head, except a small sticker with a high-contrast pattern on it for aiding the eyetracker in keeping the infant's position.

Before the experiment began, the procedure was explained to parents, informed consent was obtained, and a vocabulary checklist and word-exposure survey were completed. The child and parent were then led to the dimly lit testing room where the infant sat on his or her parent's lap facing a computer display (Fig. 1). Parents wore a visor that prevented them from seeing the screen and headphones over which they were prompted with the target sentence. The prerecorded sentence was then followed by a tone that indicated to the parent that she should begin repeating the sentence that she had just heard (Fig. S2).

Infants were presented with 48 test trials under two interspersed conditions: 32 paired-picture trials and 16 scene trials. There were eight foods and eight body-part items under each condition. During paired-picture trials, infants saw two images on the screen: one from the food category and one from the body-part category. On scene trials, infants saw a single complex image of a body, face, or of one of two tabletops with four food items on it (Fig. S1). Thus, paired-picture trials presented targets across the domains of food and body parts whereas the scene trials presented targets within one of these domains. Images were shown for 3.5 or 4 s after target onset (paired-picture and scene trials, respectively) (Fig. S2); the length of time before the parent said the target varied from trial to trial, averaging ∼3–4 s. All subjects saw both trial types, and subjects were randomly assigned to one of two pseudorandomized trial orders, which counterbalanced side, picture instance, and ordering of images and target items. The experiment lasted ∼15–20 min, after which families were compensated with a choice of $20 or two children's books. The entire visit lasted ∼45 min.

ACKNOWLEDGMENTS. We thank students and faculty at Penn's Institute for Research in Cognitive Science for useful discussions of these findings, especially those of Josef Fruehwald, and we thank Richard Aslin and Lila Gleitman for comments. This work was funded by National Science Foundation Graduate Research Fellowship Program and National Science Foundation Integrative Graduate Education and Research Traineeship grants (to E.B.), by National Science Foundation Grant HSD-0433567 (to Delphine Dahan and D.S.), and by National Institutes of Health Grant R01-HD049681 (to D.S.).

1. Polka L, Werker JF (1994) Developmental changes in perception of nonnative vowel contrasts. *J Exp Psychol Hum Percept Perform* 20:421–435.
2. Werker JF, Tees RC (1984) Cross-language speech-perception: Evidence for perceptual reorganization during the 1st year of life. *Infant Behav Dev* 7:49–63.
3. Kuhl PK, Williams KA, Lacerda F, Stevens KN, Lindblom B (1992) Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255:606–608.
4. Gervain J, Mehler J (2010) Speech perception and language acquisition in the first year of life. *Annu Rev Psychol* 61:191–218.
5. Jusczyk PW, Aslin RN (1995) Infants' detection of the sound patterns of words in fluent speech. *Cognit Psychol* 29:1–23.
6. Jusczyk PW, Hohne EA (1997) Infants' memory for spoken words. *Science* 277: 1984–1986.
7. Mattys SL, Jusczyk PW, Luce PA, Morgan JL (1999) Phonotactic and prosodic effects on word segmentation in infants. *Cognit Psychol* 38:465–494.
8. Swingley D (2005) Statistical clustering and the contents of the infant vocabulary. *Cognit Psychol* 50:86–132.
9. Marcus GF (2000) Pabiku and Ga TiGa: Two mechanisms infants use to learn about the world. *Curr Dir Psychol Sci* 9:145–147.
10. Saffran JR, Aslin RN, Newport EL (1996) Statistical learning by 8-month-old infants. *Science* 274:1926–1928.
11. Shukla M, White KS, Aslin RN (2011) Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proc Natl Acad Sci USA* 108: 6038–6043.
12. Thomas DG, Campos JJ, Shucard DW, Ramsay DS, Shucard J (1981) Semantic comprehension in infancy: A signal detection analysis. *Child Dev* 52:798–903.
13. Baillargeon R (2008) Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspect Psychol Sci* 3:2–13.
14. Waxman SR, Gelman SA (2009) Early word-learning entails reference, not merely associations. *Trends Cogn Sci* 13:258–263.
15. Bloom P (2000) *How Children Learn the Meanings of Words* (MIT Press, Cambridge, MA), pp xii, 300 pp.
16. Tomasello M (2001) Could we please lose the mapping metaphor, please? *Behav Brain Sci* 24:1119–1120
17. Carpenter M, Nagell K, Tomasello M, Butterworth G, Moore C (1998) Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr Soc Res Child Dev* 63(4):i–vi, 1–143.
18. Pilley JW, Reid AK (2011) Border collie comprehends object names as verbal referents. *Behav Processes* 86:184–195.
19. Gleitman L (1990) The structural sources of verb meanings. *Lang Acquis* 1:3–55.
20. Tincoff R, Jusczyk PW (1999) Some beginnings of word comprehension in 6-month-olds. *Psychol Sci* 10:172–175.
21. Swingley D (2009) Contributions of infant word learning to language development. *Philos Trans R Soc Lond B Biol Sci* 364:3617–3632.
22. Fernald A, Zangl R, Portillo AL, Marchman VA (2008) Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. *Developmental Psycholinguistics: On-line Methods in Children's Language Processing*, eds Sekerina I, Fernandez E, Clahsen H (John Benjamins, Amsterdam), pp 97–135.
23. Swingley D (2011) The looking-while-listening procedure. *Guide to Research Methods in Child Language*, ed Hoff E (Wiley-Blackwell, Hoboken, NJ).
24. Hirsh-Pasek K, Golinkoff RM (1996) The preferential looking paradigm reveals emerging language comprehension. *Methods for Assessing Children's Syntax*, eds McDaniel D, McKee C, Cairns H (MIT Press, Cambridge, MA), pp 105–124.
25. Swingley D, Pinto JP, Fernald A (1999) Continuous processing in word recognition at 24 months. *Cognition* 71:73–108.
26. Bortfeld H, Morgan JL, Golinkoff RM, Rathbun K (2005) Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychol Sci* 16:298–304.
27. Meltzoff AN, Brooks R (2007) Eyes wide shut: The importance of eyes in infant gaze following and understanding other minds. *Gaze Following: Its Development and Significance*, eds Flom R, Lee K, Muir D (Erlbaum, Mahwah, NJ), pp 217–241.
28. Fenson L, et al. (1994) Variability in early communicative development. *Monogr Soc Res Child Dev* 59(5):1–173; discussion 174–185.
29. Bloom L (1995) *The Transition from Infancy to Language: Acquiring the Power of Expression* (Cambridge University Press, Cambridge, UK).
30. Hamlin JK, Wynn K, Bloom P (2007) Social evaluation by preverbal infants. *Nature* 450:557–559.
31. Kovács AM, Téglás E, Endress AD (2010) The social sense: Susceptibility to others' beliefs in human infants and adults. *Science* 330:1830–1834.
32. Senju A, Csibra G (2008) Gaze following in human infants depends on communicative signals. *Curr Biol* 18:668–671.
33. Smith L, Yu C (2008) Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition* 106:1558–1568.
34. Sloutsky VM, Robinson CW (2008) The role of words and sounds in infants' visual processing: From overshadowing to attentional tuning. *Cogn Sci* 32:342–365.
35. Markman EM (1990) Constraints children place on word meanings. *Cogn Sci* 14:57–77.
36. Yeung HH, Werker JF (2009) Learning words' sounds before learning how words sound: 9-month-olds use distinct objects as cues to categorize speech information. *Cognition* 113:234–243.
37. Chemla E, Mintz TH, Bernal S, Christophe A (2009) Categorizing words using 'frequent frames': What cross-linguistic analyses reveal about distributional acquisition strategies. *Dev Sci* 12:396–406.
38. Gervain J, Nespor M, Mazuka R, Horie R, Mehler J (2008) Bootstrapping word order in prelexical infants: A Japanese-Italian cross-linguistic study. *Cognit Psychol* 57:56–74.
39. Yoshinaga-Itano C, Sedey AL, Coulter DK, Mehl AL (1998) Language of early- and later-identified children with hearing loss. *Pediatrics* 102:1161–1171.