



# Nature and origins of the lexicon in 6-mo-olds

Elika Bergelson<sup>a,b,1</sup> and Richard N. Aslin<sup>b,c</sup>

<sup>a</sup>Psychology & Neuroscience, Duke University, Durham, NC 27708; <sup>b</sup>Brain & Cognitive Sciences, University of Rochester, Rochester, NY 14627; and <sup>c</sup>Haskins Laboratories, New Haven, CT 06511

Edited by Patricia K. Kuhl, University of Washington, Seattle, WA, and approved September 28, 2017 (received for review July 22, 2017)

**Recent research reported the surprising finding that even 6-month-olds understand common nouns [Bergelson E, Swingle D (2012) *Proc Natl Acad Sci USA* 109:3253–3258]. However, is their early lexicon structured and acquired like older learners? We test 6-month-olds for a hallmark of the mature lexicon: cross-word relations. We also examine whether properties of the home environment that have been linked with lexical knowledge in older children are detectable in the initial stage of comprehension. We use a new dataset, which includes in-lab comprehension and home measures from the same infants. We find evidence for cross-word structure: On seeing two images of common nouns, infants looked significantly more at named target images when the competitor images were semantically unrelated (e.g., milk and foot) than when they were related (e.g., milk and juice), just as older learners do. We further find initial evidence for home-lab links: common noun “copresence” (i.e., whether words’ referents were present and attended to in home recordings) correlated with in-lab comprehension. These findings suggest that, even in neophyte word learners, cross-word relations are formed early and the home learning environment measurably helps shape the lexicon from the outset.**

word learning | lexicon | cognitive development | language acquisition | environmental effects

To learn words, infants integrate their linguistic experiences with word forms and the conceptual categories to which they refer. They do this fast: A growing literature demonstrates that, by around 6 mo, infants have begun understanding nouns (1–5), suggesting they form word-referent links from their environment in the first half-year.

The speech–sound learning trajectory in year one is relatively well-established (6): Infants’ language-specific sensitivity emerges around 6 mo for vowels, and 12 mo for consonants (7, 8). Indeed, by 12 mo, infants reveal robust phonetic representations for common words (9–11), and fine-grained knowledge of native language speech–sound combinatorics (12). Before this, their sensitivity to phonemic and talker-specific differences can be fragile (3, 13).

In contrast, early meaning is understudied: It’s not clear what makes the first words infants understand learnable, or what aspects of meaning infants initially represent. This is partly because meaning components are not straightforward. While phonetic features (e.g., voicing) let us readily quantify speech–sound differences, characterizing meaning is harder; consider describing or comparing how “dog” and “log” sound versus what they mean. While toddlers are sensitive to visual similarity, shape, and semantic category (14–17), little is known about nascent semantic representations.

Regarding early semantics, Arias-Trejo and Plunkett (18) find that both visual similarity and category membership contribute to semantic competition: For toddlers, understanding “shoe” in the context of a boot and a shoe was harder than when shoe appeared with a hat or bin instead. Thus, even in seasoned word learners, certain visual contexts make it harder to ascertain a spoken word’s referent.

Bergelson and Aslin (19) provide further data on early meanings. They find that, over 12 mo to 20 mo, infants’ semantic specificity increases: Although younger infants looked at a named

target to similar degrees whether hearing an appropriate or a related label (e.g., “cookie” or “banana” to label a cookie), older infants did so less. This suggests that 1-y-olds have immature extensions for words they know something about (i.e., “banana” could refer to a cookie), but leaves open whether younger infants are already sensitive to words’ relatedness (i.e., banana’s and cookie’s shared meaning).

Despite the relative dearth of infant work, a large literature on adults’ single-word representations (20) reveals sensitivity to context and meaning. Adults consider semantic and perceptual relations among words in both the visual world paradigm (21, 22) and lexical decision tasks (20). Taken together, previous work with children and adults suggests that knowledge of how words are related goes hand in hand with knowledge of what words mean. Here we ask whether this is true for infants’ earliest words, or whether initial words are more like “islands,” unrelated to other emerging lexical entries.

Although lab studies provide controlled assessments of children’s knowledge, their natural habitats are far more complex. Corpus research has been crucial for establishing what children may learn from in their daily environments. Such data have been used to unpack both linguistic and nonlinguistic aspects of the input (12, 23–27); they may also prove critical for understanding early lexical development.

However, there are exceedingly few available corpora of young infants, fewer yet with video, and none linking to comprehension measures in those same children. Here we begin to address these gaps by gathering real-time processing data for words within and across semantic categories, and investigating how infants’ home experiences with concrete nouns may influence their overall early comprehension. We ask the following: (i) Does semantic relatedness between visually available referents influence word

## Significance

**Infants start understanding words at 6 mo, when they also excel at subtle speech–sound distinctions and simple multimodal associations, but don’t yet talk, walk, or point. However, true word learning requires integrating the speech stream with the world and learning how words interrelate. Using eye tracking, we show that neophyte word learners already represent the semantic relations between words. We further show that these same infants’ word learning has ties to their environment: The more they hear labels for what they’re looking at and attending to, the stronger their overall comprehension. These results provide an integrative approach for investigating home environment effects on early language and suggest that language delays could be detected in early infancy for possible remediation.**

Author contributions: E.B. and R.N.A. designed research; E.B. performed research; E.B. analyzed data; E.B. wrote the paper; and E.B. and R.N.A. contributed to the revisions of the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

<sup>1</sup>To whom correspondence should be addressed. Email: elika.bergelson@gmail.com.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1712966114/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1712966114/-DCSupplemental).

comprehension in novice word learners? (ii) Do readily measurable aspects of infants' home life account for those same infants' variability in word comprehension?

We answer these questions through in-lab eye tracking and home recordings. The eye-tracking experiment addressed question *i*: We presented infants with image pairs that were either semantically related or unrelated (e.g., car–stroller or car–juice); then, one image was named aloud (e.g., “car”). By hypothesis, if young infants are influenced by semantic relatedness (as toddlers and adults are), we predict better performance in the unrelated trials. That is, in the context of two semantically unrelated images, we predict stronger comprehension than in the context of two related images.

Our home environment analysis addressed question *ii*: We gathered daylong audio and hour-long video recordings from infants in their homes, and examined just those time slices when concrete nouns were directed to the infant. We then derived measures of quantity, talker variability, utterance type, and situational context, and explored how they might be related to performance on the eye-tracking task.

Previous research with older infants suggests that hearing more words strengthens the early vocabulary whether in daily interactions (24), or in shared reading (28, 29). In the lab, talker variability aids word learning (30). For utterance type, words said in short phrases (i.e., alone or at utterance edges) are proposed to be learned more readily (31, 32). Relatedly, syntactic diversity (i.e., hearing a word across more sentence structures) has been shown to aid word learning (25).

Finally, we examined referential transparency (i.e., whether parents talk about observable referents), which is suggested to facilitate language learning (33–35). Notably, previous research examining why infants learn common nonnouns (e.g., “hi,” “eat”) later than common nouns found that visual copresence varied across word class (36). Nonnouns were more likely to be said when the referent event was not occurring (i.e., “hi” said with no one entering the scene) vis-à-vis nouns, which were generally proximally present when named. If such referential transparency were a basic feature that boosts learnability, then here too we would expect that the degree of “object copresence” in infants' experience would map onto early comprehension, which has not previously been shown within infants, or within nouns.

## Results

Data processing and annotation information, and in-house scripts are available on our OSF lab wiki and github repository (<https://github.com/SeedlingsBabylab/>). Raw home-recording data (audio and video) are available through HomeBank and Databrary; see details in *Materials and Methods*; clips are available in *Supporting Information*.

**Eye-Tracking Results.** Eye-tracking data were processed in R 3.3.1 to determine where the child was looking for each 20-ms bin during the test trials: the target or distractor interest areas (an invisible 620 × 620 pixel rectangle around each image), or neither. Eye movement data were time-aligned to parents' target word utterances (noted by experimenter key press).

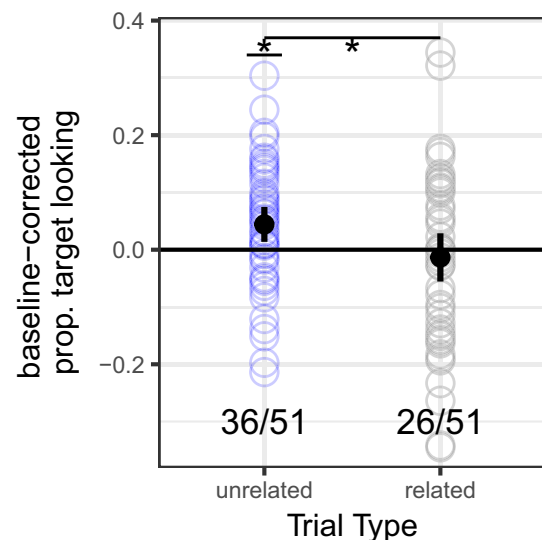
These data were then aggregated across two time windows: a pretarget baseline from trial start to target word onset and a posttarget window from 367 ms to trial end, i.e., 5,000 ms after target onset. Given the longitudinal home-and-lab design of this data collection, we exclude at the trial level rather than the infant level, where possible. Trials were excluded if infants did not look at either image for at least 1/3 of the target window (367 ms to 5,000 ms), if no looking was recording in the pretarget baseline window, or if the trial was never displayed due to experiment termination for infant fussiness; see *Supporting Information*.

We used the standard baseline-corrected target looking metric, which calculates the proportion of target looking [target/

(target+distractor)] in the posttarget window, and subtracts this same proportion from the baseline window. Visual inspection of subject means revealed one outlier (>5 SD above the mean). Once this outlier was removed, this outcome measure did not differ from a normal distribution (Shapiro–Wilk test,  $P = 0.93$ ).

We predicted that, if infants' comprehension is affected by semantic relatedness, performance on related trials would be worse than on unrelated trials. Indeed, performance was significantly above chance on unrelated trials [ $t(50) = 2.91$ ,  $P = 0.005$ , by two-tailed one-sample  $t$  test], at chance on related trials [ $t(50) = -0.64$ ,  $P = 0.524$ , by two-tailed one-sample  $t$  test], and significantly different between the two [ $t(50) = 2.22$ ,  $P = 0.031$ , by two-tailed paired  $t$  test]. That is, infants looked more at the labeled image on unrelated trials than on related trials. Over subjects, 36/51 infants attained positive subject means for the unrelated trial type ( $M = 0.044$ ,  $SD = 0.108$ ,  $P = 0.005$  by binomial test), while only 26/51 did so for related trials ( $M = -0.013$ ,  $SD = 0.15$ ,  $P = 1$  by binomial test). Over items, infants showed the same numeric pattern as over subjects, i.e., the item mean was higher for unrelated items than related items ( $M = 0.035$ ,  $SD = 0.081$  and  $M = 0.003$ ,  $SD = 0.067$ , respectively), and item means were positive for most items in the unrelated condition, but not the related condition; with only 16 items, item-level effects were not different from chance ( $P > 0.05$ ). Summarily, in the related condition but not in the unrelated condition, performance was positive over most infants and most items; see Fig. 1 and Figs. S1 and S2.

**Home-Recording Results.** After preprocessing (see *Materials and Methods*), annotators marked each object word (i.e., concrete noun) in the recordings along with three properties: utterance type, object copresence, and speaker. Each word's utterance type was classified by its syntactic and prosodic features into seven categories: declaratives, questions, imperatives, short phrases, reading, singing, and unclear. Object copresence was coded “yes,” “no,” and “unclear” based on whether the annotator felt that the object corresponding to the word being annotated was present and attended to by the child. For videos, this was generally visually appreciable; for audio files, annotators used their impression from the context (e.g., generally, in “here's your spoon!” spoon



**Fig. 1.** Comprehension by trial type. Dots represent each infant's baseline-corrected proportion of target looking, per trial type (unrelated, related). Mean and 95% CIs are in black. Asterisk indicates  $P < 0.05$  for the unrelated and related trial types; the fraction indicates the proportion of infants with positive trial-type means.

**Table 1. Home-recording descriptive statistics**

Rec type	No. tokens		No. types		No. speakers		Prop mat input		Prop obj.cop.		Utt type ent		Prop short phrase	
	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)
Aud	74 to 1,619	714.5 (355.7)	39 to 375	184.5 (64.9)	2 to 16	6.5 (3.4)	0.2 to 1	0.6 (0.2)	0.1 to 0.7	0.5 (0.2)	1.3 to 2.4	1.9 (0.3)	0 to 0.2	0.1 (0.05)
Vid	30 to 635	171.5 (120.9)	13 to 169	58.3 (34.4)	1 to 6	2.9 (1.2)	0 to 1	0.7 (0.3)	0.1 to 0.9	0.6 (0.2)	1.2 to 2.5	1.9 (0.3)	0 to 0.5	0.1 (0.1)

Aud, audio; ent, entropy; mat, maternal; obj.cop., object copresence; Prop, proportion; Rec, recording; Utt, utterance; Vid, video.

was coded “yes” for object copresence, while in “twinkle little star,” star was coded “no”). Interrater reliability was high [computed for 10% of annotations for utterance type: 89% agreement, Cohen’s (*K*) = 0.8; object copresence: 85% agreement, Cohen’s (*K*) = 0.7]. We then derived token counts for the various forms a word occurred in (e.g., “tooth,” “tootheroo,” “teeth”), and type counts for the lemma (e.g., “tooth”).

While input data varied by child, there was relatively high consistency in our proportional measures (Table 1). We operationalized input quantity as the number of types and tokens of each lemma. Given that the video recordings were all ~1 h, while the audio recordings varied based on the nap time of the child, we report daily audio rates and hourly video rates.

In our daylong audio recordings, infants heard ~710 object word tokens of ~180 word types, from seven speakers, on average. Sixty percent of this input came from infants’ mothers, and 50% of the time that infants heard an object word, the corresponding referent was visible and attended to (i.e., 50% object copresence). In our hour-long video recordings, infants heard ~170 tokens, from ~60 types, from three speakers. In the videos, 70% of the input came from infants’ mothers, with 60% object copresence. For both audio and video recordings, the average entropy across utterance types (i.e., the variability proportions of each utterance type in “bits”) was 1.9; short phrases were 10% of the input (Fig. 2).

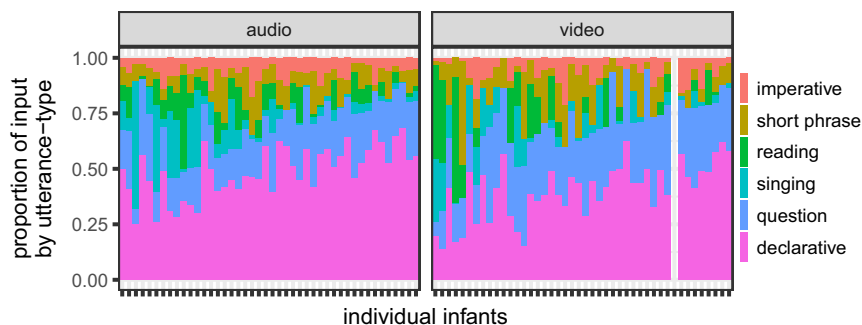
Comparing the daylong audio and hour-long video recordings, we found that, although infants heard more object word input in the audio recordings in an absolute sense, they heard relatively more input in the videos. That is, infants heard only 25 to 50% fewer word tokens, word types, and speakers in the hour-long video than in the daylong audio recording (from a different day), even though the latter was ~10–11× longer than the former; we explore this in ongoing work.

**Questionnaires.** Vocabulary questionnaires [MacArthur-Bates Communicative Development Inventory (MCDI)] showed that parents felt their infants understood few of our 16 tested words [*M* = 1.96 (3.98); *R*: 0 to 15; mode: 0], whereas, on Word Exposure Surveys, parents indicated that their child heard these words

daily on average (4 on a 1 to 5 scale from “never” to “several times a day”). According to parental report, 71% of infants were not yet babbling, all but one were not yet hands-and-knees crawling, and 29% were exclusively breast-fed; see [Supporting Information](#).

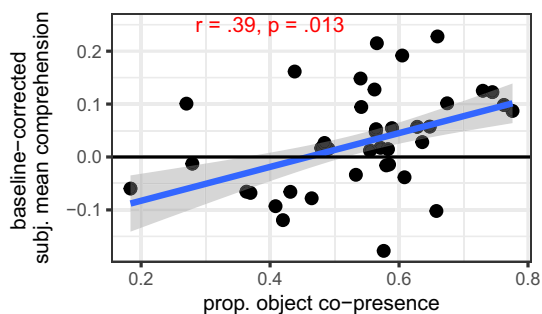
**Home and Lab Linkages.** We next examined data from the infants who provided both home and in-lab data (video and eye tracking: *n* = 40; audio and eye tracking: *n* = 41). We modeled infants’ subject means from our eye-tracking experiment as a function of the properties we had annotated (noun input, talker, utterance types, and object copresence). Given the relatively small sample size, the large number of ways one might aggregate the home data, and both predicted and unanticipated collinearity among these four preidentified properties of interest, we opted for a simple analysis approach. Namely, we tested directionally specified correlations between home environment measures found to predict lexical knowledge in previous research and infants’ in-lab comprehension. That is, if measures that predict lexical knowledge in older infants hold at 6 mo at levels our home measures can detect, we would see positive correlations between in-lab comprehension and noun input, talker variability, utterance type diversity, and object copresence. When analyzing counts, we examined audio and video data separately, since length varied substantially by recording type. When examining proportions, we averaged audio and video data. We first conducted Shapiro–Wilk Normality tests; if both variables were normally distributed, we used Pearson correlations; if one or both were not, we used Kendall correlations.

As described above, the eye-tracking results suggest that neophyte word learners are sensitive to semantic similarity. However, given that the home environment is not split into experiences relevant for our two experimental trial types, we examine how overall lab performance correlates with home measures. Notably, aggregating across trial types for each infant, performance overall was not above chance, given the relatively weak performance on related trials [*t*(40) = 1.59, *P* = 0.12]. [This *t* test was run on the subset of infants reported above for whom there is



**Fig. 2.** Distribution of object word utterances, by infant. Color shows utterance type [daylong audio (Left) and hour-long video (Right)]. One video was lost; one word with unclear utterance type was removed. The x axes are identically ordered by each child’s overall proportion of object words in declaratives and questions.





**Fig. 3.** In-lab comprehension by proportion of object copresence. Each point indicates a given infant's average proportion of object copresence in the audio and video home recordings (x axis) by that same infant's subject mean in the eye-tracking experiment (y axis). Line indicates robust linear fit, with 95% CI in gray.

also home-recording data ( $n = 41$ ); the same pattern holds over all infants included in the initial analysis.]

We find a significant correlation between the proportion of object copresence and infants' overall in-lab comprehension ( $r = 0.39$ ,  $P = 0.013$ ) (Fig. 3). This result is consistent with a broader role for referential transparency in word learning, even among (early-learned) nouns; we return to this in *Discussion*.

We next considered four measures of object word quantity from the literature: number of types, tokens, and words read, along with type–token ratio. These variables tend to be highly correlated with each other, and, indeed, we find each of them significantly correlated with the others in both audio and video data ( $|r|$  between 0.23 and 0.82;  $P < 0.05$ ). (As expected, type–token ratio negatively correlated with quantity; all other  $r$ 's were positive.)

Given predicted correlations among quantity measures, we asked how these measures correlated with lab performance. No video quantity measures correlated significantly with in-lab comprehension. However, several quantity metrics from the audio recordings (number of types, tokens, and object word–containing reading utterances) were close to the  $P < 0.05$  significance threshold [ $r = 0.32$ ,  $P = 0.047$ ;  $r = 0.29$ ,  $P = 0.074$ ; and  $(\tau) = 0.19$ ,  $P = 0.088$ , respectively]; this leaves open the possibility that an expanded home and lab sample, or other measures of word knowledge and/or input quantity, would render clearer results.

Looking at talker variability, we again did not find significant correlations with in-lab comprehension in audio or video recordings, or in the proportion of input from infants' mothers (all  $P > 0.05$ ). Finally, looking at utterance type variability (calculated as each infant's average entropy across utterance types), we found no significant correlations with in-lab comprehension ( $P = 0.58$ ), likely due to highly similar utterance type distributions across participants (Fig. 2). We replicated previous work reporting that infants hear  $\sim 10\%$  of words in isolation (31), but this property too was not associated with in-lab comprehension ( $P = 0.43$ ).

Thus, across the four aspects of home environment we predicted would positively correlate with in-lab comprehension, only object copresence clearly did so, although it is premature to conclude that this variable is a differentially better predictor than

the others. Given that our analyses were predicated upon directional predictions, but served as initial exploratory steps, we evaluated this correlation further by calculating bootstrapped CIs with 1,000 iterations. The 95% CI did not include 0 (0.15 to 0.62). (Audio and video object copresence were marginally correlated with each other,  $r = 0.29$  and  $P = 0.064$ .)

## Discussion

Consistent with research on adults and children, we find that 6-mo-olds understand words more readily when shown two semantically unrelated referents than when shown two related ones. We further find initial evidence that, in these same infants, in-lab word comprehension is linked with referential transparency in the home, but not with measures of talker or utterance-type, and only marginally with input quantity. These findings enrich our understanding of infants' real-time word comprehension, and the longer-scale learning environment that fuels it.

Our eye-tracking results suggest that even first words are not unconnected islands of meaning; they already contain semantic structure. Still, there are various interpretations of infants' relatively strong performance on unrelated trials (e.g., car–juice), versus their poor performance on related trials (e.g., car–stroller). One possibility is that infants know (something about) the tested words, but cannot overcome semantic competition on related trials, i.e., hearing “car” leads to car looking, but also activates related words, e.g., stroller, to a similar (or indistinguishable) degree.

Alternatively, infants' word knowledge may be underspecified: They may know enough about a word's meaning to tell it apart from the unrelated referent but not the related one (which, by design, has distributional, conceptual, and/or visual overlap). That is, perhaps infants know “car” cannot refer to juice, but not whether stroller is in the “car” category.

Furthermore, these options may intertwine, and, indeed, our time course data are compatible with both (Fig. S2): Infants consistently looked at the labeled target image in related trials, but shifted between the images in unrelated trials, consistent with underspecification or competition.

In studies of children and adults, these possibilities are disambiguated by both “cleaner” (i.e., less noisy, more accurate) eye movements, which reveal transient looks at the semantically related distractor, and overt (touch or click) target selection, which 6-mo-olds (who do not even point yet) cannot do. Additional infant measures (e.g., neural recordings or reaching tasks) may be promising future directions.

These results also complement previous early word comprehension research (1, 2, 4, 5). Here too, further work may elucidate how presentation method influences infants' looking behavior (e.g., video vs. photo; two-image displays vs. scenes).

Turning to the corpus results, home–lab links were relatively limited. Specifically, we did not find a relation between how infants' object word input was distributed across speakers and utterance types, and infants' comprehension of common nouns. Especially for utterance type, this may reflect the limits of analyzing only object word utterances. Similarly, input quantity measures (which are tied to toddlers' vocabulary, e.g., ref. 24) were only marginally correlated with comprehension. While these variables may matter for subsequent lexical knowledge, they did



**Fig. 4.** Item pairs in eye-tracking study. Infants saw each image pair twice. There were 16 trials in each trial type (related, unrelated), 32 trials total.

not unequivocally do so here. Given our young and relatively homogenous sample, we may have had limited variability therein.

Intriguingly, object copresence significantly correlated with in-lab comprehension. Of course, infants do not learn words and referents they have not experienced. Furthermore, parents' "focusing in" on infants' attention likely provides higher-quality learning instances (35). Our results also suggest an expansion of Bergelson and Swingley (36), who found that common nouns are more referentially transparent than common nonnouns. Here we suggest that referential transparency for a given infant may map onto that same child's comprehension, within the already referentially transparent noun class.

These results are first steps toward understanding how the initial lexicon is organized and acquired from experience by 6 mo. We find that real-time comprehension is influenced by links among words. We further find promising results tying infants' experiences with referential transparency to early word knowledge. Finally, these results suggest a combined lab-home (LH) approach can begin to reveal the range and dynamics of learner-by-environment interaction, as infants' initial understanding of words starts to give way to robust knowledge of their language.

## Materials and Methods

**Participants.** The final eye-tracking experiment sample was 51 6-mo-olds ( $M = 6.1$  mo,  $r = 5.6$  mo to 6.7 mo, 23 female). Families elected to participate in a one-time lab-only (LO) study ( $n = 12$ ; LO group), or additionally enroll in a larger yearlong study with home visits ( $n = 44$ ; LH group). Four further infants who participated in the eye-tracking study were excluded for fussiness or calibration failure resulting in 0 trials with sufficient data for analysis, in one or both trial types. One additional excluded infant performed  $>5$  SD above the mean. See [Supporting Information](#) for further details on per-trial exclusion criteria.

Infants were recruited from a database in Rochester, NY. All children were healthy, had no hearing or vision problems, were carried full term ( $40 \pm 3$  wk), and heard  $\geq 75\%$  English at home. Families received \$10 and a small gift for participating in the lab study, and a further \$5 if they also completed the home visits (audio and video). Families who completed our optional demographics questionnaire (98%) reported that infants were largely white and middle class (LH: 95% white, 75% of mothers received a B.A. or higher; LO: 79% and 57%, respectively).

Given unknown effect sizes for home-lab links in young infants, the target sample size was 48, i.e.,  $3 \times$  the standard minimum sample size ( $n = 16$ ). We enrolled and retained 44 infants over an 8-mo enrollment window for the LH group. Once LH enrollment ended, LO did as well.

**Lab Visit Procedure.** First, staff explained the study to families and obtained consent for either the yearlong study (including home recordings) or one-time eye tracking, as relevant (approved by the University of Rochester IRB). Parents then completed surveys about their child; see [Supporting Information](#).

Next, the parent sat with the infant in their lap in a dimly lit testing room, in front of an EyeLink 1000+ Eyetracker (SR Research), which was in head-free mode, and sampled monocularly at 500 Hz ( $<0.5^\circ$  average accuracy). A small sticker on the infant's forehead tracked head movements. The experiment was run from a computer that was back-to-back with the testing monitor, allowing for adjustment if the child moved out of eye-tracking range.

The experiment began with four "warm-ups," in which a single image was labeled by a sentence played over speakers (e.g., "Look at the apple!"). Parents were then given a visor that blocked the screen (or closed their eyes,  $n = 3$ ), and over-ear headphones. The experiment was also recorded by camcorder to ensure compliance and monitor the child's state.

Next came 32 test trials, in which infants saw two images on a gray background, and heard a sentence labeling one image (Fig. 4). An attention getter was shown as needed. On each test trial, parents spoke a single sentence aloud to their child, which labeled one of the images on the screen; they first heard a prerecorded sentence over headphones that they then repeated aloud (1). Images were shown for 5 s after target word onset; the length of time before the parent said the target word after the images appeared varied across trials, averaging  $\sim 3$  s to 4 s.

Each infant saw both trial types (16 related and 16 unrelated trials, interspersed pseudorandomly). Infants were alternately assigned to two trial orders, which counterbalanced side and ordering of images, target items, and trial type.

**Stimuli.** Sixteen common concrete nouns were chosen as target words, based on corpora and prior research; see [Supporting Information](#) and Fig. 4. Each word was part of two item pairs, one in each trial type (related and unrelated, e.g., dog-baby and spoon-baby;  $n = 16$  item pairs). Pairings maximized semantic overlap within related pairs, and minimized it in unrelated pairs. Semantic network analyses confirmed that related pairs were more similar than unrelated pairs; see [Supporting Information](#). Within trial type, items were paired to minimize phonetic overlap (one related pair unavoidably had the same initial consonant: book-ball).

Audio stimuli were sentences recorded by a female using infant-directed speech prosody in a sound booth; they were 1.1 s to 1.8 s, normalized to 72 db (i.e., a volume that allowed only parents to hear them when they were played over the parent headphones). Infants only heard the sentences from their parents (see [Lab Visit Procedure](#)). Each sentence occurred in one of four carrier phrases: "Can you find the X?" "Where's the X?" "Do you see the X?" and "Look at the X!" where X is the target word (only one sentence frame was used per item pair).

Visual stimuli were photos of each target and warm-up word, edited onto a gray background, displayed at  $500 \times 500$  pixels on a 27.4- by 34-cm LCD 96 pixels per inch screen, at a viewing distance of 55 cm to 60 cm. Warm-up images ( $n = 4$ ) were displayed centrally. For test trials ( $n = 32$ ), each image was centered within the left and right half of the screen (counterbalanced across trials). Each of 16 test photos occurred four times: once as target and once as distractor in each trial type (related and unrelated; see Fig. 4).

**Home-Recording Procedure.** Home recordings captured infants' typical environment through an hour-long video recording and, on a separate day, a daylong audio recording. Before recording, parents were given a release form, which allowed up to three levels of sharing; see [Supporting Information](#). Recordings that parents opted to share with authorized researchers can be found on Databrary and Homebank ( $n = 43$ ); access is available to researchers who complete ethics certification and membership agreements through these repositories. See sample clips in [Supporting Information](#).

Video recordings took place at infants' homes. Research staff put a specialized hat on the child, and a small Looxcie camera ( $8.4 \times 1.7 \times 1.3$  cm, 22 g) was affixed above each ear with Velcro (one pointed slightly up and one slightly down, to better capture the child's visual field). Staff also put a camcorder on a tripod in the corner. Parents were given an information sheet, asked to move the tripod if they changed rooms, and given staff contacts. Staff departed and returned after 1 h.

At either the lab or home visit, parents were given a LENA (Language Environment Analysis) audio recorder ( $8.6 \times 5.6 \times 1.3$  cm, 57 g) and clothing with a LENA pocket. (LENA Foundation). Parents were asked to turn the recorder on when the child awoke, and off at the end of the day or if they wanted to stop recording. Recording time ranged from 10.66 h to 16.00 h ( $M = 14.37$ ).

**ACKNOWLEDGMENTS.** We thank SEEDLingS staff: Amatuni, Dailey, Koorathota, Schneider, Tor; research assistants at University of Rochester and Duke University; and National Institutes of Health Grants T32 DC000035 and DP5-OD019812 (to E.B.) and HD-037082 (to R.N.A.).

- Bergelson E, Swingley D (2012) At 6-9 months, human infants know the meanings of many common nouns. *Proc Natl Acad Sci USA* 109:3253-3258.
- Bergelson E, Swingley D (2015) Early word comprehension in infants: Replication and extension. *Lang Learn Dev* 11:369-380.
- Parise E, Csibra G (2012) Electrophysiological evidence for the understanding of maternal speech by 9-month-old infants. *Psychol Sci* 23:728-733.
- Tincoff R, Jusczyk PW (1999) Some beginnings of word comprehension in 6-month-olds. *Psychol Sci* 10:172-175.
- Tincoff R, Jusczyk PW (2012) Six-month-olds comprehend words that refer to parts of the body. *Infancy* 17:432-444.

- Gervain J, Mehler J (2010) Speech perception and language acquisition in the first year of life. *Annu Rev Psychol* 61:191-218.
- Werker JF, Tees RC (1984) Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behav Dev* 7:49-63.
- Polka L, Werker JF (1994) Developmental changes in perception of nonnative vowel contrasts. *J Exp Psychol Hum perception Perform* 20:421-435.
- Swingley D, Aslin RN (2002) Lexical Neighborhoods and the word-form representations of 14-month-olds. *Psychol Sci* 13:480-484.
- Mani N, Plunkett K (2010) Twelve-month-olds know their cups from their keps and tups. *Infancy* 15:445-470.

11. Vihman MM, Nakai S, DePaolis RA, Hallé PA (2004) The role of accentual pattern in early lexical representation. *J Mem Lang* 50:336–353.
12. Swingle D (2009) Contributions of infant word learning to language development. *Philos Trans R Soc Lond Ser B Biol Sci* 364:3617–3632.
13. Bergelson E, Swingle D (June 22, 2017) Young infants' word comprehension given an unfamiliar talker or altered pronunciations. *Child Dev*, 10.1111/cdev.12888.
14. Smith LB, Jones S, Landau B, Gershkoff-Stowe L, Samuelson L (2002) Object name learning provides on-the-job training for attention. *Psychol Sci* 13:13–19.
15. Wojcik EH, Saffran JR (2013) The ontogeny of lexical networks: Toddlers encode the relationships among referents when learning novel words. *Psychol Sci* 24:1898–1905.
16. Borovsky A, Ellis E, Evans J, Elman J (2016) Semantic structure in vocabulary knowledge interacts with lexical and sentence processing in infancy. *Child Dev* 87:1893–1908.
17. Luche CD, Durrant S, Floccia C, Plunkett K (2014) Implicit meaning in 18-month-old toddlers. *Dev Sci* 17:1–8.
18. Arias-Trejo N, Plunkett K (2010) The effects of perceptual similarity and category membership on early word-referent identification. *J Exp Child Psychol* 105:63–80.
19. Bergelson E, Aslin R (2017) Semantic specificity in one-year-olds' word comprehension. *Lang Learn Dev* 13:481–501.
20. Neely JH (1991) Semantic priming effects in visual word recognition: A selective review of current findings and theories. *Basic Processes in Reading: Visual Word Recognition*, eds Besner D, Humphreys G (Lawrence Erlbaum Assoc, Hillsdale, NJ), pp 264–336.
21. Huettig F, Altmann GTM (2005) Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition* 96:23–32.
22. Dahan D, Tanenhaus MK (2005) Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychon Bull Rev* 12:453–459.
23. Hart B, Risley TR (1992) American parenting of language-learning children: Persisting differences in family-child interactions observed in natural home environments. *Dev Psychol* 28:1096–1105.
24. Weisleder A, Fernald A (2013) Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychol Sci* 24:2143–2152.
25. Hoff E, Naigles LG (2002) How children use input to acquire a lexicon. *Child Dev* 73:418–433.
26. Shneidman L, Goldin-Meadow S (2012) Language input and acquisition in a mayan village: How important is directed speech? *Dev Sci* 15:659–673.
27. Lieven E (1994) Crosslinguistic and crosscultural aspects of language addressed to children. *Input and Interaction in Language Acquisition*, eds Gallaway C, Richards BJ (Cambridge Univ Press, Cambridge, UK), pp 56–73.
28. Debaryshe BD (1993) Joint picture-book reading correlates of early oral language skill. *J Child Lang* 20:455–461.
29. Montag JL, Jones MN, Smith LB (2015) The words children Hear: Picture books and the statistics for language learning. *Psychol Sci* 26:1489–1496.
30. Rost GC, McMurray B (2009) Speaker variability augments phonological processing in early word learning. *Dev Sci* 12:339–349.
31. Brent MR, Siskind JM (2001) The role of exposure to isolated words in early vocabulary development. *Cognition* 81:B33–B44.
32. Seidl A, Johnson EK (2006) Infant word segmentation revisited: Edge alignment facilitates target extraction. *Dev Sci* 9:565–573.
33. Medina TN, Snedeker J, Trueswell JC, Gleitman LR (2011) How words can and cannot be learned by observation. *Proc Natl Acad Sci USA* 108:9014–9019.
34. Yurovsky D, Smith LB, Yu C (2013) Statistical word learning at scale: The baby's view is better. *Dev Sci* 16:959–966.
35. McGillion ML, et al. (2013) Supporting early vocabulary development: What sort of responsiveness matters? *IEEE Trans Auton Ment Dev* 5:240–248.
36. Bergelson E, Swingle D (2013) The acquisition of abstract words by young infants. *Cognition* 127:391–397.
37. MacWhinney B (2000) *The CHILDES Project: Tools for Analyzing Talk, Volume II: The Database* (Lawrence Erlbaum, Mahwah, NJ), 3rd Ed.
38. Dale PS, Fenson L (1996) Lexical development norms for young children. *Behav Res Methods Instrum Comput* 28:125–127.
39. Libertus K, Landa RJ (2013) The Early Motor Questionnaire (EMQ): A parental report measure of early motor development. *Infant Behav Dev* 36:833–842.
40. Walle EA, Campos JJ (2014) Infant language development is related to the acquisition of walking. *Dev Psychol* 50:336–348.