Supporting Information

Bergelson and Aslin 10.1073/pnas.1712966114

SI Materials and Methods

Eye-Tracking Target Word Selection and Relatedness Quantification. The 16 target words tested in the eye-tracking experiment were selected as high-frequency concrete nouns commonly heard by infants. These words occurred on average 613 times (R: 46 to 2,197), from 14/16 mothers (R: 6 to 16) in the Brent Corpus (an audio corpus of mothers and 9- to 15-mo-olds) (31). Many of the words were also used in previous word comprehension studies of 6- to 9-mo-olds (1–4).

To quantify the relatedness among the related and unrelated item pairs used in the eye-tracking experiment, we used a semantic network model (word2vec) over the North American English portion of CHILDES (Child Language Data Exchange System) (37), and computed the similarity (i.e., cosine between two word vectors) for each unrelated and related pair. The unrelated pairs had significantly lower similarity ($M = 0.18$) than the related pairs [0.47; $T(12.19) = 0.009$] by Welch two-sample t test. Code for this analysis is available on github: https://github.com/SeedlingsBabylab/w2v_cosines/.

Home-Recording Data Processing and Sharing. For the hourlong home video recordings, the head camera and camcorder feeds were merged into a single stream using Sony Vegas and exported as .mp4 files for annotation in Datavyu. The daylong home audio recordings were processed using LENA's proprietary algorithm (which provides machine-generated utterance segmentation); the raw audio file (.wav) and LENA output were exported for further processing. We converted the LENA output into a CLAN (Child Language Analysis)-compatible file (37), for annotation. This provided a "skeleton" with each row's time stamp corresponding to a LENA "utterance." We then used Audacity's Silence Finder algorithm to demark long periods of silence (corresponding to naptime or silent car rides) in the annotation files. Omitting these silent stretches left 7.25 h to 16.00 h per file ($M = 11.17$), which were then manually annotated as described in the main text.

Families completed an audio–video release form in which they could elect for their recordings to be shared at several levels: sharing with the lab only, sharing with other authorized researchers (e.g., HomeBank and Databrary), and sharing short excerpts for demonstration purposes in publications and/or research talks. Families were also informed that they could elect not to share a section of any recording, they could stop the recording for any reason, and they should inform anyone who had more than incidental contact with the child that they were being recorded, and obtain their permission. Each release form was collected after the audio and video recordings for a given family were completed.

Additional Questionnaire Description and Subanalysis. As summarized in the main text, parents completed a series of questionnaires about their infants. They completed the MCDI Words & Gestures Form (38), and two motor surveys: the gross-motor section of the Early Motor Questionnaire (39) and four locomotion questions (40). Parents were also orally asked our Word Exposure Survey, which asked how often they believed their child heard our test words on a five-point scale (1 = "Never", 5 = "Several times a day"). Parents also described infants' vocalizations, to ascertain whether canonical babbling had begun, and were asked whether their child was breast-fed or bottle-fed (since "milk" and "bottle" are two of our tested words).

Parents generally did not believe that their infant understood very many words: The modal number of words infants were reported to understand on the MCDI is 0, with one notable outlier of 162 (Table S1). This pattern of MCDI results vis-a-vis in-lab data was also found in Bergelson and Swingley (36). We believe this reflects the inherent difficulty in determining whether young infants, who lack overt behavioral cues like pointing or saying words, understand the words around them. Thus, the in-lab results we report underscore the utility of eye-tracking measures with infants of this age; subtle eye movements may provide researchers with a way to measure comprehension that caretakers themselves may miss or find hard to assess.

The questionnaire data, in principle, open up the possibility of further subanalyses. Given that the majority of infants (in our relatively small sample of 51 6-mo-olds) were not yet crawling or babbling, we did not conduct further analyses of these abilities in relation to in-lab or home language data. For the feeding data, given that ∼30% of infants were exclusively breast-fed, we reanalyzed the in-lab data, excluding the item "milk" for these children, under the rationale that our image of milk may not have comported with the majority of their milk experiences. All patterns of significance reported in the main text remained the same. This may reflect that infants have experiences with milk other than during their own feeding (e.g., seeing older children and adults drinking milk), or that infants saw breast milk in bottles (i.e., parents may report the child is "breast-fed exclusively" in reference to the source of the milk rather than the delivery method). It may also reflect that removing one item does not change the robust patterns across conditions we report here, especially given that feeding survey results were not reported for nearly half of infants, as this questionnaire was added a few months after data collection began (Table S1).

Data Exclusion. As stated in the main text, we opted to exclude data at the trial level, rather than the infant level, to retain infants in the home and lab analyses. Given our relatively large number of trials per infant ($n = 32$), we were able to proceed with analysis using trials of comparable and sufficient data quality. Noting that practices vary across labs and studies, we provide here the exclusion rates from language comprehension studies with young infants, to help situate our trial exclusion rate. In two studies with 6-mo-olds, Tincoff and Jusczyk (4, 5) presented infants with a single test trial, and excluded 20 to 33% of the sample for failure to complete the trial due to inattention or fussiness. Parise and Csibra (3) excluded 43% of 9-mo-olds in a word comprehension EEG study. In the present eye-tracking study, we retain approximately half of the trials, from most participants. Given that the lab-and-home sample ($n = 44$) was part of a longitudinal set of recordings and in-lab studies future research may assess whether fussiness and inattentive behavior at 6 mo is predictive of subsequent behavior in eye-tracking experiments, and in measures of language development.

Figs. S1 and S2 supplement the analyses provided in the main text.

SI Multimedia

The clips linked in *Supporting Information* show three sample audio clips and two sample video clips from the corpus. Parents provided permission for these to be shared with researchers. As explained in the main text, each object word was tagged for utterance type, talker, and object copresence. We provide clips with a range of object copresence for clarity on how this was determined in our audio and video files.
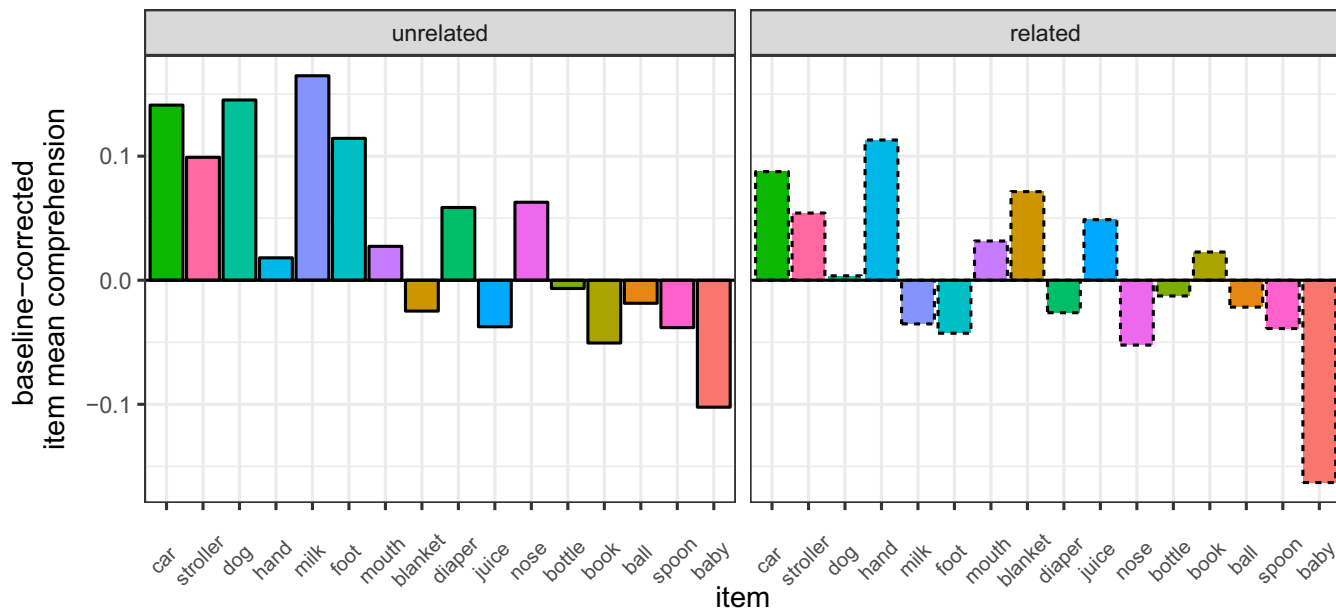
Fig. S1. Item-level performance, across subjects, in each condition. Bars are ordered by cross-condition average for each item.
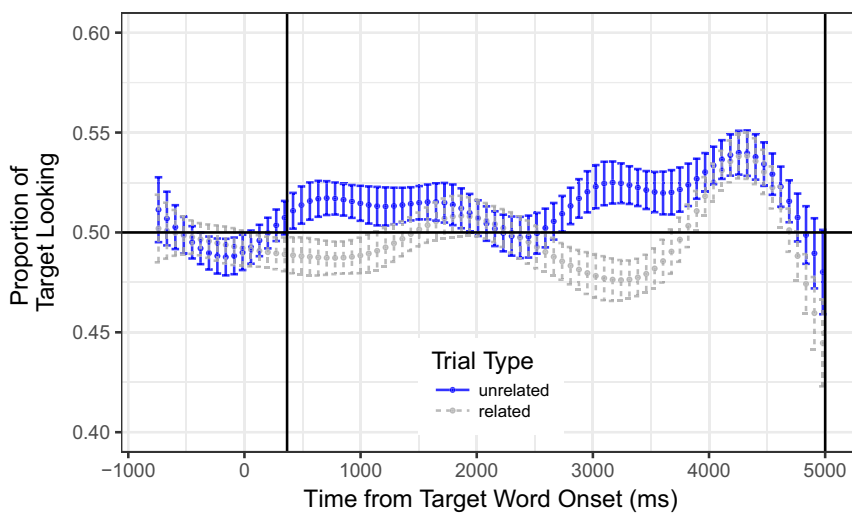

Fig. S2. Time course of infant gaze by trial type. Shown is the proportion of fixation to the target image over time, for each trial type (blue, unrelated; gray, related). We smooth over each 20-ms bin, averaging over subjects and trials, and add 95% bootstrapped CIs. Black vertical lines demark the target window of analysis (367 ms to 5,000 ms). The baseline window is all looking before target onset, i.e., time < 0.

Table S1. Questionnaire results

Measure	Lab-and-home	Lab-only	All infants	Missing
MCDI: tested words	$M = 1.52$ (2.51), R: 0 to 12, mode = 0	$M = 3.46$ (4.63), R: 0 to 15, mode = 0	$M = 1.96$ (3.18), R: 0 to 15, mode = 0	1 LH
MCDI: all words	$M = 9.43$ (16.57), R: 0 to 75, mode = 0	$M = 23.85$ (43.41), R: 0 to 162, mode = 0	$M = 12.72$ (25.53), R: 0 to 162, mode = 0	1 LH
Word exposure	$M = 3.92$ (0.67), R: 2.09 to 4.93, mode = 4	$M = 3.92$ (0.6), R: 2.62 to 5, mode = 4	$M = 3.92$ (0.64), R: 2.21 to 4.95, mode = 4	0 LH, 1 LO
Motor status (a)	2.27%	0%	1.79%	1 LH, 1 LO
Feeding status (b)	31.03%	16.67%	28.57%	15 LH, 8 LO
Production status (c)	75%	58.33%	71.43%	0 LH, 2 LO

Infant categories are as follows: a, infants who had begun hands-and-knees crawling; b, infants who were exclusively breast-fed; and c, infants who are not yet babbling. The "Missing" column reflects how many infants in the lab-and-home (LH) and lab-only (LO) samples did not fill out the questionnaire.
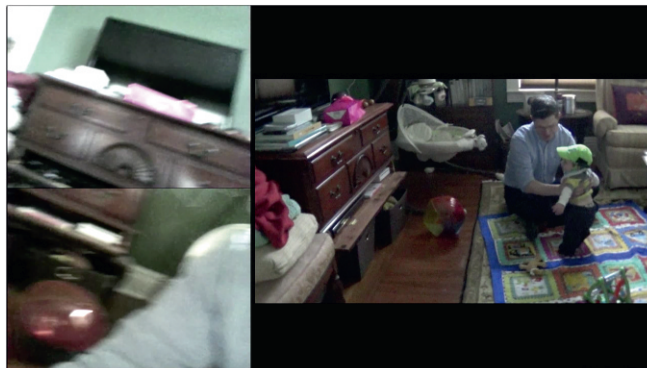
Movie S1. In this clip, the father is singing and dancing with his child. None of the words in the song (e.g., "chickie," "egg") are copresent in the scene.

Movie S1
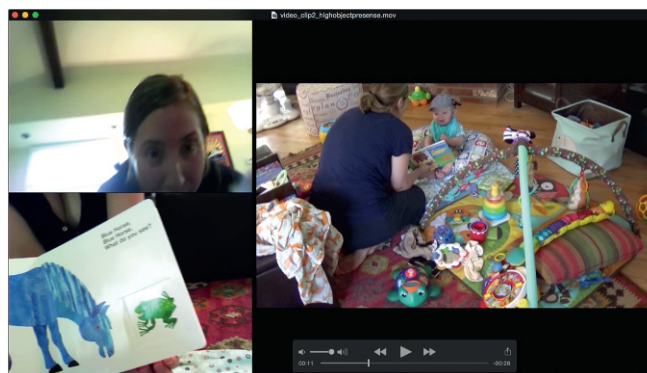
Movie S2. In this clip, the mother is reading a picture-book to her child. We see from his head-mounted cameras and the camcorder that the images in the book go with the words that the mother is saying, and that the child is attending to them, i.e., they are copresent.

Movie S2

Audio File S1.   In this clip, the infant's father discusses his missing phone with his 6-mo-old. This is an example of low object copresence, since the phone is not present as the father discusses it.

[Audio File S1](#)

Audio File S2.   In this clip, the infant's mother and father are reading a book to their child. We hear the pages being turned, suggesting the book and the items it describes are present and attended to as they are mentioned.

[Audio File S2](#)

Audio File S3.   In this clip, the mother is nursing, and mentions a future eating-at-the-beach event. We can infer from context that the baby food and sand discussed are not copresent in the scene. In contrast, the sweet potato she says is in the child's eye was tagged as copresent.

[Audio File S3](#)