



Published in final edited form as:

Lang Learn Dev. 2015 ; 11(4): 369–380. doi:10.1080/15475441.2014.979387.

Early Word Comprehension in Infants: Replication and Extension

Elika Bergelson^{a,1} and Daniel Swingley^b

^a Brain and Cognitive Sciences, University of Rochester Center for Language Sciences Meliora Hall Rochester, NY, 14627 U.S.A.

^b Dept. of Psychology, University of Pennsylvania Institute for Research in Cognition Science 3401 Walnut St., 400a Philadelphia, PA, 19104 U.S.A.

Abstract

A handful of recent experimental reports have shown that infants of 6 to 9 months know the meanings of some common words. Here, we replicate and extend these findings. With a new set of items, we show that when young infants (age 6-16 months, $n=49$) are presented with side-by-side video clips depicting various common early words, and one clip is named in a sentence, they look at the named video at above-chance rates. We demonstrate anew that infants understand common words by 6-9 months, and that performance increases substantially around 14 months. The results imply that 6-9 month olds' failure to understand words not referring to objects (verbs, adjectives, performatives) in a similar prior study is not attributable to the use of dynamic video depictions. Thus, 6-9 month olds' experience of spoken language includes some understanding of common words for concrete objects, but relatively impoverished comprehension of other words.

Introduction

Recent studies have demonstrated that infants begin to understand common words in the second six months of life (Bergelson & Swingley, 2012, 2013b; Parise & Csibra, 2012; Tincoff & Jusczyk, 1999, 2012). While this is in keeping with some parental diary-based accounts (Bates, 1993; Benedict, 1979), these findings contrast with the generally accepted developmental timeline which purports that the first year of life involves mainly phonological development, with word-referent links occurring around infants' first birthday (Bloom, 2001; Kuhl, 2011; Tomasello, 2001). Replicating and extending this work is critical for ascertaining how robust these findings are across groups of infants, words, and methodological variations. The longstanding earlier view that infants only begin learning word meanings at 9 months or later was based on a substantial body of evidence--largely drawn from study of children's language production. This in turn has left quite open what kind of word-referent linkages infants may have acquired before beginning to say words themselves. The present replication provides further evidence in support of infants' early word learning capabilities. Moreover, the findings reported here clarify whether a previously

¹Corresponding Author: Elika Bergelson Erika.bergelson@gmail.com.
swingley@psych.upenn.edu

observed contrast between infants' success in understanding nouns when tested with static images, and failure in understanding non-nouns when tested with dynamic images, should be attributed to the nature of the images, or to the nature of the words.

Bergelson and Swingley (2012) demonstrated that by 6-9 months, infants understand a set of nouns for foods and body parts (e.g. "banana," "mouth"). Infants were shown photos of these items either in pairs on a plain background ("paired-picture trials") or in large images featuring several test items ("scene trials"). The paired-picture trials tested infants' knowledge across the food and body-part categories (the test displays always showed one of each), while the scene trials featured within-category items. The conclusions of that work were threefold: infants understand words by 6-9 months, performance remains stable from 6 months to about 13 months, and performance thereafter increases robustly.

In an extension to EEG, Parise and Csibra (2012) demonstrated that 9-month-olds show an N400 mismatch effect when a revealed object (shown as a still image) does not match the word just uttered by infants' mothers. The items were common objects familiar to children, e.g. *banana*, *bunny*, *socks*. This work contributes an important confirmation that using a different method, infants demonstrate word comprehension at 9 months. However, younger infants were not tested, and surprisingly, even at 9 months infants showed no recognition response when hearing a voice other than their mother's.

The present work is methodologically closest to two studies by Tincoff and Jusczyk (1999, 2012). These authors demonstrated that 6-month-olds understand two pairs of words: "hand" and "feet" (2012) and the proper names "Mommy" and "Daddy" (1999). In these experiments, infants were shown a single pair of videos depicting these items. The "Mommy" and "Daddy" were presented live on screen, and the "hand" and "feet" moved a bit in the image. While one video was named, infants' eyegaze was measured; infants looked more at the matching video.

Another recent study tested early comprehension of an abstract set of non-nouns (e.g. "eat" and "uh-oh"; Bergelson & Swingley, 2013b). This work used paired video-clips as stimuli. Here, 10-13-month-olds showed above-chance levels of performance, but 6-9-month-olds did not. As Bergelson and Swingley (2012) had found, comprehension increased substantially around 14 months, suggesting this developmental improvement is unrelated to the type of word being tested.

Thus, prior studies using paired video stimuli have tested only one word-pair (Tincoff & Jusczyk, 1999, 2012), have tested substantially older children (Golinkoff et al., 1987; Houston-Price et al., 2005), or have actually failed to reveal word comprehension in children under 10 months (Bergelson & Swingley, 2013b). These results leave open whether young infants can link many common words and referents when those referents are depicted in naturalistic, moving videos. That is, on the one hand, 6-9 month olds have only been shown to understand common nouns depicted as still images, excepting Tincoff & Jusczyk (2012), which only tested a single word pair, *hand* and *foot*. On the other hand, 6-9 month-olds failed to demonstrate word comprehension using paired video stimuli testing a more heterogeneous set of non-nouns. In sum, the ensemble of prior studies leads to some doubt

about the robustness of extremely early word comprehension, and also some question about the real nature of the previously observed difference between concrete nouns and non-nouns. The present study was intended to help resolve these problems.

One methodological concern with paired video stimuli is that one video might capture infants' attention more strongly than a static image would, reducing the directive impact of the language stimulus. Another is that the background and other context typical of a live-action video might make the target referent more difficult for infants to isolate and identify, limiting the power of the method to uncover still-fragile language competence. On the other hand, video stimuli may provide a better proxy for infants' real day-to-day interactions, given that their word-referent experiences with their caregivers happen in a dynamic context.

The present research examined early comprehension of words that were instantiated as competing video stimuli. We presented 6–16-month-old infants with a broad set of common words, such as “dog” and “car” (see Table 1). The words we tested had some overlap with previous work, but only one pair of nouns (hand-feet) had been tested before with young infants, as either still images or video clips (Tincoff & Jusczyk, 2012). We also included two pairs of verbs, one of which was tested in Bergelson & Swingley (2013b). Thus, this study set out to replicate the basic finding of previous work, i.e. that young infants understand common nouns, while also extending the range of the words tested, and their format of presentation. The present study also extends previous findings in its methods, utilizing a varied series of video stimuli over twelve trials.

Methods

Participants

Infants were recruited from the Philadelphia area. 49 were retained in the final sample (M=12.0 mo., R=6.7-16.4 mo., 26 girls). An additional 35 infants were excluded due to fussiness or parental influence leading to fewer than half the trials being usable¹ (n=26), technical problems (n=6), or failure to meet language criteria (n=3). Infants were recruited by mail, email, phone, internet, and in person. All were healthy, had been carried full-term, heard >75% English at home, and had no history of chronic ear infections. Infants came from a range of socioeconomic statuses. Mother's education was distributed as follows: high school or less (n=9), some college (n=14), college degree (n=12), advanced degree (n=12). Of those subjects in the final sample who provided ethnicity information (n=48), 46% identified as White, 35% as African American, and 19% as multi-racial or other. Of those providing information about whether or not they are Hispanic (n=44), 10% identified as Hispanic, 90% as not Hispanic.

¹Whether or not to remove a child's data from analyses was determined through the following criteria. First, we removed trials in which the parent failed to say the target sentence, said it wrong, or pointed, as well as trials in which the experimenter's notes indicated that infants were crying or yelling. We also removed trials in which infants had failed to look at both videos at some point during the trial, and trials on which they had looked at the videos for less than 1/3 of the window of interest (here, 1,211ms). Many of these criteria were redundant, and eliminated the same trials. Based on this trial-level vetting, subjects were removed entirely if they failed to contribute data to 3 or more of the 6 pairs of items. This led to the removal of 26 subjects. These infants' age distribution (M=11.2mo., R=5.9-13.5mo., 12 girls), did not differ from the distribution of infants included in the final sample (D=.27, p=.18 by Kolmogorov-Smirnoff test)

While age is clearly a continuous variable, for ease of exposition and comparison to previous work we present analyses that bin infants into three age groups: 6-9 months ($n=12$; $R=6.74-9.56$ mo., $M=8.34$), 10-13 months ($n=21$; $R=10.15-13.99$ mo., $M=12.05$), and 14-16 months ($n=18$; $R=14.09-16.43$ mo., $M=15.22$); see figure 1. The same general pattern holds if we divide the age-range into three equal-sized bins, or three bins with equal numbers of infants in them. We also present results from a loess smoother that includes age as a continuous predictor in Figure 1.

Materials

Infants were presented with twelve 10s videos organized into six yoked pairs: *bottle-car*, *dog-bear*, *drink-kiss*, *hand-feet*, *hug-eat*, and *juice-ball*. Videos (15.2 x 12.7 cm) were displayed side-by-side on a 34.7x26.0-cm LCD screen. Items were selected based on frequency of occurrence in an infant-directed speech corpus including 16 mothers interacting with their infants (Brent & Siskind, 2001), and based on a database of parental reports indicating which words parents believed their children understood or said (MCDI; Dale & Fenson, 1996). Each word appeared in 75-100% of the 16 Brent & Siskind mothers' speech, and had a corpus \log_{10} frequency ranging from 2.09-3.02. Each word was reportedly understood by 46-96% of 16 month olds in the MCDI database (Dale & Fenson, 1996). Additionally, infants have been shown to understand a subset of these items (*hand*, *feet*, *juice*, and *bottle*) in word comprehension research (Bergelson & Swingley, 2012, 2013b; Tincoff & Jusczyk, 2012). All videos were recorded in a room with muted colors, on a table, or in front of a blank screen, and are available upon request. Videos of objects included a baby bottle ("bottle"), a sippy cup of juice ("juice"), a small toy car ("car"), and a teddy bear. These videos did not include whole people or other objects, but were zoomed-in clips of a hand moving these objects around with lateral movements, shaking, and looming motions commonly used with infants. The hand and feet videos featured these body parts moving slowly about, and no other body parts or objects (with cropping to limit the shot to the arm/legs). The dog video featured a live dog turning his head a bit. The hug and kiss videos featured a research assistant hugging or kissing a doll. The drink and eat videos featured a research assistant drinking out of a cup, or eating cheerios, both while facing a bit to the side. These individual videos were put into pairs based on intrinsic differences in degree of movement in the videos, to ensure non-overlapping target onsets and codas, and for compatibility of carrier phrase across the pair. See Table 1.

Apparatus and Procedure

Visual fixation data were collected using an Eyelink CL computer (SR Research), with a reported accuracy of .5, sampling monocularly at 500Hz. The eyetracker operated using a camera below the computer screen, and required no head-restraint. A sticker with a high-contrast pattern, which aided the eyetracking mechanism, was placed on the infant's forehead. All stimuli were presented on the eyetracker's 17-inch monitor.

Before the experiment began, the procedure was explained to parents, who gave informed consent. Parents completed the full form of the MacArthur-Bates Communication and Development Inventory (MCDI), a vocabulary survey instrument. The majority of parents completed the Words and Gestures version of this inventory².

Parents also completed a word exposure survey, which provided, for each child, a parental estimate of how often their child normally heard our test words in daily life, ranging from “never” to “several times a day”. When the forms were completed, parent and child were led to the dimly-lit testing room where the infant sat on the parent's lap facing a computer display. Parents wore an opaque visor preventing them from seeing the screen, and headphones over which they were prompted with the target sentence. The experimenter controlled the experiment using Eyelink's Experiment Builder software, from a second computer positioned back-to-back with the first.

Each of the six yoked item-pairs was presented for one block consisting of a familiarization phase and test phase. Each such block was preceded by a spinning star attention-getter, to draw infants' attention to the screen's center; this disappeared once infants fixated it (or after 10s). Then, the familiarization phase began. One video of the pair played twice through on one side of the screen, followed by the other video playing twice through on the other side (40s total). Then both videos played simultaneously twice more (20s). This phase was accompanied by music, and familiarized infants with the videos and their locations.

Then the test phase of the block began. Here, the videos played through four more times simultaneously (40s). In the first test trial for the pair, parents spoke a single sentence to their child, repeating a prerecorded utterance that they heard over headphones. This utterance named one of the two videos in the pair. During the second presentation of the pair, parents again heard and repeated this sentence. Then, parents did the same for the other video in the pair during the pair's third and fourth presentation in the block. The utterances are listed in Table 1. At the end of this test phase, the block for the next pair began.

The sentences parents heard and repeated had been recorded by a native English-speaking woman talking at a moderate speed, with slightly exaggerated intonation. The recorded sentences were 1.6-2.8s long and were audible only to the parent (~34 dB). The exact timing of parental sentences varied across trials, but the onset of the target word was recorded by the experimenter by key press; the videos played for ~6s after the parent's utterance ended. The onset of the target word was later used to align fixations with respect to this time point.

Each participant was randomly assigned to one of four pseudorandomized trial orders, which varied in the ordering of the blocks and in which side each video was on. To investigate the role of order regularity, children were also assigned to one of two conditions. In the *fixed* condition (n=23), the mother always named the video on the left for the first test trial, and then the video on the right for the second test trial (thereby making up a block with a Left-Right target structure that was repeated over all 12 trials). In the *random* condition (n=26), whether the first named target was on the right or on the left varied quasirandomly across item-pairs. All children were tested on all 12 items. The experiment lasted about 20 minutes. Families were compensated with a choice of two children's books or \$20. The entire visit lasted about 45 minutes.

²Five subjects in the oldest age-group were given the ‘words and sentences’ version of the CDI, with 680 items rather than 395. The subset of words from this older CDI that are on the other CDI was used for the numbers reported in this table, except the 20 words that do not overlap between the two. All patterns of results remain the same regardless of which vocabulary measure is used for these 5 subjects.

Results

The outcome measure was how much more infants looked at the target video upon hearing it named, computed as in previously published work (Bergelson & Swingley, 2012, 2013b). That is, we calculated the difference in fixation proportions for each item-pair: the proportion of time that infants looked at one video when it was the target, minus the proportion of time they looked at it when it was the distracter. This computation corrects for biases from preferences for one video over the other (Bergelson & Swingley, 2012, 2013b), and yields one score for each item-pair, computed across trials for each infant. For instance, with the pair *juice-ball* an infant's performance was given as how much she looked at *juice* when “juice” was said, relative to her looking at *juice* when “ball” was said. Positive difference scores indicate word comprehension. We measured performance in the window from 367 to 4000ms after the onset of the spoken target word. Fixation responses earlier than 367ms are unlikely to be responses to the speech signal (Swingley, 2009); the 4000ms window offset comports with previous work (Bergelson & Swingley, 2013b). This is a longer window than is typically used with older children, because younger children respond more slowly (Fernald et al., 1998).

Subject and Item Tests, 6-9 month olds

Preliminary analyses indicated no main effect of the order condition (*fixed, random*), no interaction between order and age, and no difference in subject means across order by two-sample Wilcoxon test (estimated difference=-.001, $p=.98$; See Figure 1). All further analyses thus collapse across order. At 6-9 months ($n=12$), performance was significantly above chance over subjects (Mdn=.051, $CI=[.006,.010]$, $p=.034$ by two-tailed Wilcoxon Test; all subsequent subject and item tests are one or two sample two-tailed Wilcoxon Tests, unless indicated otherwise)³. See Figures 1 and 3. This pattern held for 9/12 infants. This pattern also held over most item-pairs (4/6) though performance over item-pair means did not reach significance (Mdn=.048, $p=.31$). This was partly caused by the mixed pattern of performance on the two verb pairs: while infants performed well on *hug-eat*, they performed poorly on *drink-kiss*, due to strong biases to look at the *kiss* video whether it was the target or not.⁴

Subject and Item Tests 10-13 and 14-16 month olds

Performance over subjects and items in both older groups was quite strong. 15/20 10-13 month olds showed positive performance over subject means (Mdn=.081, $CI=[.02, .15]$ $p=.019$). Performance was positive over 5/6 item-pairs, but did not reach significance, again because of the pair *drink-kiss* (Mdn=.091, $CI=[-.028,.18]$, $p=.094$). Similarly, for 14-16 month olds, performance was robust but quantitatively higher: 13/17 subjects attained

³The outcome of one-sample Wilcoxon tests is a distribution-free measure of central tendency, the pseudo-median (i.e. the Hodges-Lehman Estimate); we abbreviate this as Mdn, and provide the 95% confidence interval (CI). For two-sample Wilcoxon Tests (a.k.a. Mann-Whitney Tests), we report the difference estimator between groups; we abbreviate this as estdiff. This is the median of the difference between a sample from group 1 and a sample from group 2.

⁴For all pairs, there was generally a preferred video across infants before any target word was said; this bias was numerically more pronounced for *kiss-drink* than for other pairs. The proportion of *kiss/kiss+drink* looking in the pre-target window was .70; the average bias for the more popular item of all other pairs was .62 (SD across pairs=.020).

positive subject means, (Mdn=.16, CI=[.039,.29], $p=.0066$). Performance for 14-16 month olds was positive over all 6 item-pairs (Mdn=.14, CI=[.087,.23] $p=.031$). See figures 1-3.

Comparison across Ages

A direct two-way comparison across each pair of age-groups over subject means and over items found no statistically significant differences between any groups (all $\text{estdiff} < .08$, all $p > .26$ over subject means and all $\text{estdiff} < .1$, all $p > .13$ over item means). However, an examination of subject means by age suggests that the increase with age is non-linear, and that the lack of group differences may be an artifact of the binning process. Indeed, fitting a loess estimator to the data suggests a sharp increase in performance around 13-14 months (see Figure 1).

Vocabulary Measures

Data from parent reports of their infants' vocabulary is given in Table 2, along with a "usage frequency" of how often parents believed their infants heard our tested words (a 0-4 scale derived from parents responses to the options "never", "less than once per week", "a few times per week", "once a day" and "several times a day"). Vocabulary by all measures increased with age (Kendall's τ range: .20-.50, $p < .05$) but usage frequency did not ($\tau=.038$, $p=.71$), with most parents reporting that infants heard the target words once a day on average.

Considering the entire sample, infants' mean target fixation performance was not correlated with the CDI vocabulary measures or scores on our word-exposure survey (all $p > .23$). However, overall production vocabulary did correlate with performance in the 6-9 month olds ($\tau=.63$, $p=.0078$), and in the two younger groups combined ($\tau=.33$, $p=.01$). See Figure 4. Across the two younger groups, both children not yet saying any of our test words, and children who were saying at least one test word evinced word comprehension (Mdn_{zero_testwords}=.044, CI=[.002,.086], $p=.045$; Mdn_{nonzero_testwords}=.18, CI=[.059,.41], $p=.016$). However the children saying at least one test word showed significantly better performance (EstDiff=.13, CI=[.03,.25], $p=.027$); 17/25 non-word producers, and 7/7 word producers showed positive subject means. Looking at children's production on the entire vocabulary checklist, infants across the two younger groups saying no words (unsurprisingly few in number) did not show comprehension in our experiment (5/10 positive subject means, Mdn_{zero_words}=.01, CI=[-.05,.094], $p=.77$, while those who are saying words showed strong comprehension (19/22 positive subject means, Mdn_{nonzero_words}=.095, CI=[.05,.144], $p=.00050$). Here too, the difference between these groups of infants is significant (EstDiff=.079, CI=[.005,.16], $p=.035$)⁵. This observation, which awaits confirmation in other samples, is consistent with other apparent thresholding effects in language tasks (Waxman & Markow, 1995). The subset of CDI items corresponding to the 12 tested target words also correlated with younger infants' performance, driven by the 10-13 month olds, since no 6-9 month olds were reported to say any of our tested words ($\tau=.42$, $p=.021$).

⁵All of these patterns hold with the 10-13 month old outlier (see Figure 4) removed.

Discussion

These findings indicate that 6-9 month old infants understand words that are common in their environment, and that this knowledge is broader both in its content (i.e. which words infants know) and in the robustness with which infants can demonstrate it (i.e. with somewhat more complex video stimuli) than previous work has demonstrated. While the study was not designed to test the difference between noun and verb acquisition, 6-9-month-olds did perform well on one of the two verb pairs (*hug-eat*), contrary to their poor performance on a broader set of non-nouns in a prior study (Bergelson & Swingley, 2013). Whether infants' apparent success on *hug-eat* represents random item variation, or a first fragile glimmer of infants' understanding of these verbs, is unclear, particularly given infants' failure to demonstrate understanding of *kiss-drink* here.

Performance in young infants was correlated with production vocabulary: though none reportedly produced any *target* words, those producing *any words* showed stronger comprehension. This may be due to some infants' greater linguistic or non-linguistic cognitive capacities. However, this merits further exploration given potential floor-effects in 6-9-month-olds' vocabulary reports.

This study confirms several patterns that have emerged in previous work (Bergelson & Swingley, 2012, 2013b; Parise & Csibra, 2012; Tincoff & Jusczyk, 1999, 2012). Namely, we find further evidence, with a new sample of infants and a new set of item pairs, that 6-9 month olds understand words. Not unexpectedly, this ability increases with age, but perhaps less expectedly the rate with which it increases is non-linear. This is the third study that finds evidence of a boost in performance around 13-14 months, independent of word type (Bergelson & Swingley, 2012, 2013b). We find three possible explanations for this boost viable. First, this is around the time that infants begin talking (Dale & Fenson, 1996), which may increase their metalinguistic awareness, or understanding of words as community-wide symbols in a way that boosts comprehension. Second, this is just before the time that infants begin to show basic syntactic knowledge, and this in turn may help them use the distribution of arguments in the sentence to listen more predictively (Jin & Fisher, 2014). Third, there may be a non-linguistic development (candidates include increased gaze sharing, attention, and the onset of walking), which plays a causal role in this seemingly qualitative shift in lexical development (Carpenter & Call, 2013; Rose et al., 2009; Walle & Campos, 2014). It will be critical for future research to further investigate the nature of this improvement

The present results suggest that the previously observed developmental difference between understanding of nouns (by 6 months) and non-nouns (not until 10 months) was not due to a methodological difference between employing static images and videos; infants succeed with both presentation methods. Instead, the relative delay likely concerns a true difference in learning across word-types (see discussion in Bergelson & Swingley, 2013a, 2013b; Fisher & Gleitman, 2002).

The results presented above are thus an important confirmatory contribution to the literature. They provide convergent data with a growing number of studies indicating early knowledge of nouns and slightly later knowledge of verbs and other common non-nouns. These results

also help underscore a deeper point about language development: word learning is not contingent on infants' command of a solidified native-language receptive phonology. Phonetic perception adapts to the native language to a substantial degree throughout the first year (e.g., Werker & Tees, 1984 Polka and Werker, 1995), and it is increasingly evident that infants learn words over this same developmental period. The present results therefore support a developmental narrative in which word-meaning pairings form in parallel, perhaps synergistically, with infants' growing phonology.

Acknowledgements

We thank students and faculty at Penn's Institute for Research in Cognitive Science for useful discussions of these findings, and the dedicated RAs and staff of the Penn Infant Language Center. Thanks also to our action models: Alba Tuninetti, Allison Britt, and Bodie Snyder-Mackler. This work was funded by NSF GRFP and IGERT grants, and NIH T32 DC000035 (to E.B.), and by NIH Grant R01-HD049681 (to D.S.).

Works Cited

- Bates E. Comprehension and Production in Early Language Development: Comments on Savage-Rumbaugh et al. Monographs of the Society for Research in Child Development. 1993; 58(233)
- Benedict, H. Early lexical development: Comprehension and production.. Journal of Child Language. 1979. Retrieved from <http://journals.cambridge.org/production/action/cjoGetFulltext?fulltextid=1769952>
- Bergelson E, Swingley D. At 6-9 months, human infants know the meanings of many common nouns. Proceedings of the National Academy of Sciences of the United States of America. 2012; 109(9): 3253–8. doi:10.1073/pnas.1113380109. [PubMed: 22331874]
- Bergelson, E.; Swingley, D. Social and Environmental Contributors to Infant Word Learning.. In: Knauff, M.; Pauen, M.; Sebanz, N.; Wachsmuth, I., editors. Proceedings of the 35th Annual Conference of the Cognitive Science Society. Cognitive Science Society; Austin: 2013a. p. 187-192.
- Bergelson E, Swingley D. The acquisition of abstract words by young infants. Cognition. 2013b; 127(3):391–7. doi:10.1016/j.cognition.2013.02.011. [PubMed: 23542412]
- Bloom P. Précis of How children learn the meanings of words. The Behavioral and Brain Sciences. 2001; 24(6):1095–103. discussion 1104–34. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12412326>. [PubMed: 12412326]
- Brent MR, Siskind JM. The role of exposure to isolated words in early vocabulary development. Cognition. 2001; 81(2):B33–B44. doi:10.1016/S0010-0277(01)00122-6. [PubMed: 11376642]
- Carpenter, M.; Call, J. How joint is the joint attention of apes and human infants?. In: Metcalfe, J.; Terrace, H., editors. Agency and joint attention. Oxford University Press; 2013. p. 49-61. Retrieved from http://books.google.com/books?hl=en&lr=&id=IjFpAgAAQBAJ&oi=fnd&pg=PA49&dq=How+joint+is+the+joint+attention+of+apes+and+human+infants%3F&ots=9e4zwRDuoZ&sig=_duIvH5QJxwCf_UXRM1J7cbFmK
- Dale PS, Fenson L. Lexical development norms for young children. Behavior Research Methods, Instruments, & Computers. 1996; 28(1):125–127. doi:10.3758/BF03203646.
- Fernald A, Pinto JP, Swingley D, Weinberg A, McRoberts GW. Rapid Gains in Speed of Verbal Processing by Infants in the 2nd Year. Psychological Science. 1998; 9(3):228–231. doi: 10.1111/1467-9280.00044.
- Fisher, C.; Gleitman, LR. Language Acquisition.. In: Pashler, H.; Gallistel, R., editors. Steven's Handbook of Experimental Psychology, Vol. 3: Learning, motivation, and emotion. 3rd ed.. Vol. 3. Wiley; New York: 2002. p. 445-496.
- Golinkoff RM, Hirsh-Pasek K, Cauley KM, Gordon L. The eyes have it: lexical and syntactic comprehension in a new paradigm. Journal of Child Language. 1987; 14(1):23–45. doi:10.1017/S030500090001271X. [PubMed: 3558524]

- Houston-Price C, Plunkett K, Harris P. "Word-learning wizardry" at 1;6. *Journal of Child Language*. 2005; 32(1):175–189. doi:10.1017/S0305000904006610. [PubMed: 15779882]
- Jin, K.; Fisher, C. Early evidence for syntactic bootstrapping: 15-month-olds use sentence structure in verb learning. In: Orman, W.; Valteau, M., editors. *Boston University Conference on Language Development*. Cascadia Press; Boston: 2014. Retrieved from <http://www.bu.edu/buclid/files/2014/04/jin.pdf>
- Kuhl PK. Who's talking? *Science*. 2011; 333(6042):529–30. doi:10.1126/science.1210277. [PubMed: 21798918]
- Parise E, Csibra G. Electrophysiological evidence for the understanding of maternal speech by 9-month-old infants. *Psychological Science*. 2012; 23(7):728–33. doi:10.1177/0956797612438734. [PubMed: 22692337]
- Rose, S. a; Feldman, JF.; Jankowski, JJ. A cognitive approach to the development of early language. *Child Development*. 2009; 80(1):134–50. doi:10.1111/j.1467-8624.2008.01250.x. [PubMed: 19236397]
- Swingley D. Contributions of infant word learning to language development. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. 2009; 364(1536): 3617–32. doi:10.1098/rstb.2009.0107. [PubMed: 19933136]
- Tincoff R, Jusczyk PW. Some Beginnings of Word Comprehension in 6-Month-Olds. *Psychological Science*. 1999; 10(2):172–175. doi:10.1111/1467-9280.00127.
- Tincoff R, Jusczyk PW. Six-Month-Olds Comprehend Words That Refer to Parts of the Body. *Infancy*. 2012; 17(4):432–444. doi:10.1111/j.1532-7078.2011.00084.x.
- Tomasello M. Could we please lose the mapping metaphor, please? *Behavioral and Brain Sciences*. 2001 doi:10.1017/S0140525X01390131.
- Walle, E. a; Campos, JJ. Infant language development is related to the acquisition of walking. *Developmental Psychology*. 2014; 50(2):336–48. doi:10.1037/a0033238. [PubMed: 23750505]
- Waxman SR, Markow DB. Words as invitations to form categories: evidence from 12- to 13-month-old infants. *Cognitive Psychology*. 1995; 29:257–302. doi:10.1006/cogp.1995.1016. [PubMed: 8556847]

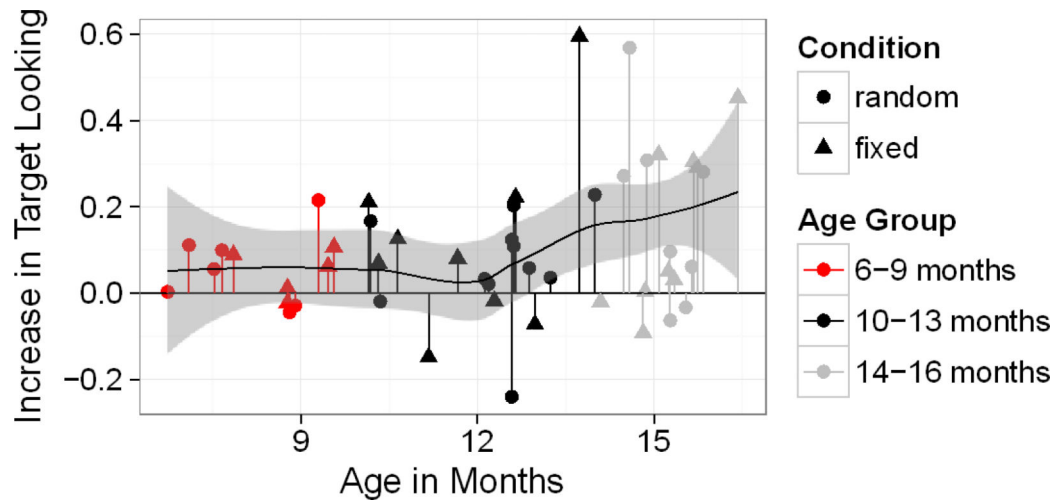


Fig. 1.

Target looking performance in each infant. Data are subject mean difference scores calculated over the 367–4000ms window. These were calculated by averaging the six item-pair mean difference scores for each subject. The shape of the symbol indicates which order manipulation condition infants were in (circles indicate “random”, triangles indicate “fixed”). The color of the lines indicates the age group used for a subset of analyses; see text for details. The black line indicates the output of a loess smoother (a locally-fitted polynomial regression), with the grey zone indicating standard errors of this fit.

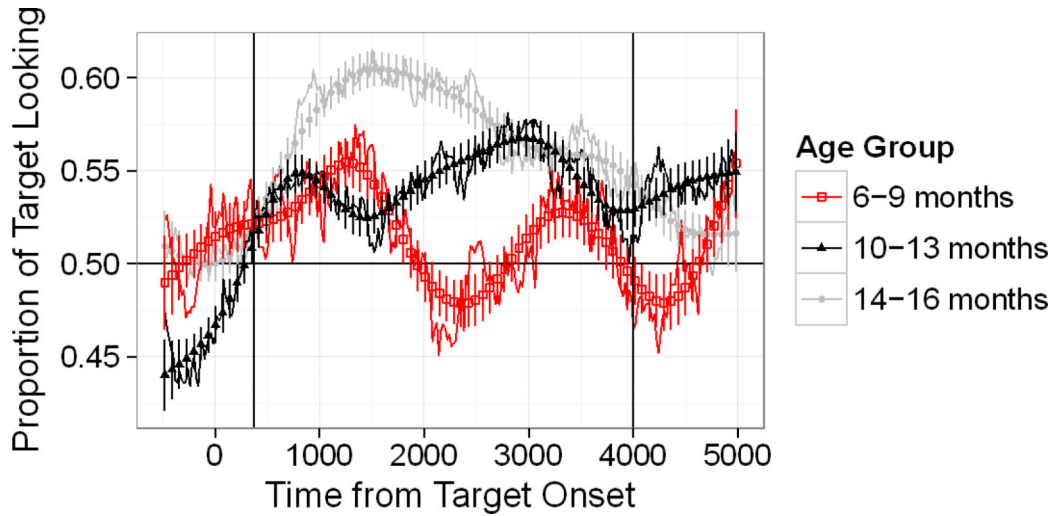


Fig. 2.

Time course of infants' picture fixation over all item-pairs, averaged over infants in three age groups. The ordinate shows the mean proportion of infants who were looking at the named (target) picture at each moment in time. The abscissa shows time from the onset of the target word, determined for each individual trial. Error bars indicate SEMs, with means computed over subjects in each age range, with an overlaid smooth using GAMs (generalized additive models with integrated smoothness estimation). At all three ages, target fixation rose from about 0.50 (chance) shortly after the onset of the spoken word. Overall, target-looking increased with age across the age groups. Black vertical lines indicate the window over which target looking was calculated.

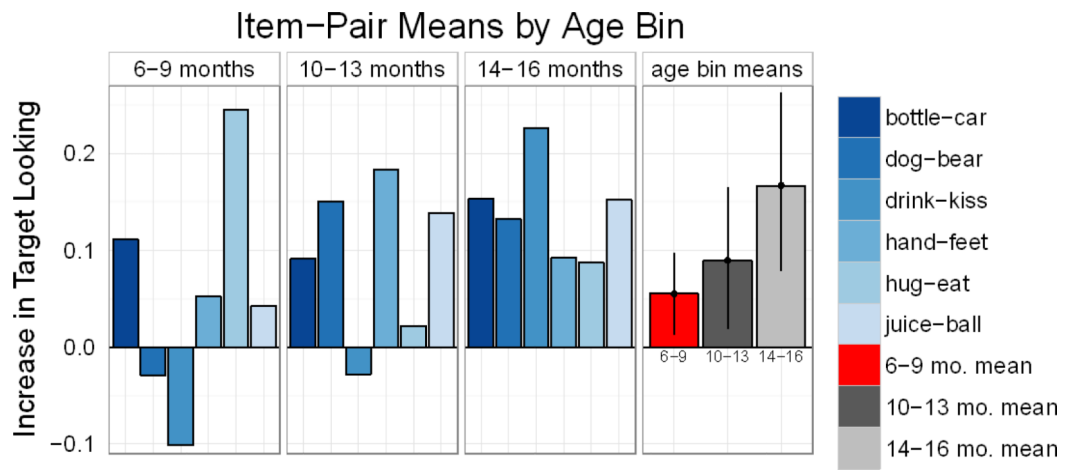


Fig. 3. Item-pair mean difference scores within each age group (three left panels), and subject means by age group (right panel). Error bars for subject means represent bootstrapped nonparametric 95% confidence intervals.



Fig. 4. Relationship between parental reports of infants' total production vocabulary, as measured by the CDI, and subject mean difference scores, by age-bin. Children reported to say zero words are indicated by open symbols; children saying at least one word are indicated by filled symbols. The scale of the abscissa varies among the three plots.

Table 1

Stimulus pairs and spoken utterances.

Pair	Carrier phrases
<i>Bottle-car</i>	"Where's the bottle/car?"
<i>Dog-bear</i>	"Look at the dog/bear!"
<i>Drink-kiss</i>	"Look, she's gonna drink/kiss!"
<i>Hand-feet</i>	"Do you see the hand/feet?"
<i>Hug-eat</i>	"Does she wanna hug/eat?"
<i>Juice-ball</i>	"Can you see the juice/ball?"

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Mean Vocabulary measures from parental reports, by age-bin.

	Overall CDI (n=395)		Tested Words CDI (n=12)		Usage
	Comprehends	Says	Comprehends	Says	Freq. (n=12)
6-9 m.o.	47 (0-194)	2 (0-12)	4 (0-10)	0 (0-0)	2.9 (2.0-3.7)
10-13 m.o.	60 (0-166)	14 (0-126)	5 (0-11)	1 (0-6)	3.0 (1.9-3.7)
14-16 m.o.	99 (0-309)	20 (0-58)	6 (0-10)	2 (0-5)	3.2 (2.7-3.8)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript