Concept-Based Word Learning in Human Infants

Jun Yin[1] and Gergely Csibra[1,2]

[1]Cognitive Development Center, Department of Cognitive Science, Central European University, and
[2]Department of Psychological Sciences, Birkbeck, University of London

Abstract

Whether infants initially learn object labels by mapping them onto similarity-defining perceptual features or onto concepts of object kinds remains under debate. We addressed this question by attempting to teach infants words for behaviorally defined action roles. In a series of experiments, we found that 14-month-olds could rapidly learn a label for the role played by the chaser in a chasing scenario, even when the different instances of chasers did not share perceptual features. Furthermore, when infants could choose, they preferred to interpret a novel label as expressing the agent's role within the observed interaction rather than as being associated with the agent's appearance. These results demonstrate that infants can learn labels as easily (or even more easily) for concepts identified by abstract behavioral characteristics as for objects identified by perceptual features. Thus, at early stages of word learning, infants already expect that novel words express concepts.

People often label objects when talking to preverbal infants. Although such labeling is usually applied to specific objects, infants can extend the meaning of words to other objects that are new instances of the same kind (Bergelson & Swingley, 2012; Parise & Csibra, 2012). The nature of this ability is the subject of controversy. On the one hand, it has been proposed that infants learn object labels by simply mapping them onto perceptual features, such as shapes, that characterize objects belonging to the same category (Landau, Smith, & Jones, 1988; Sloutsky & Fisher, 2012; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002). On the other hand, infants, like adults, may be able to conceive of new labels as names of concepts of object kinds, which would manifest a direct link between linguistic and conceptual development (Csibra & Shamsudheen, 2015; Macnamara, 1982; Waxman & Gelman, 2009).

When probing infants' recognition and generalization of words, researchers usually use either objects that are familiar to infants, such as a shoe or a banana (e.g., Bergelson & Swingley, 2012), or artificially created object categories that share features along one or more dimensions (e.g., Plunkett, Hu, & Cohen, 2008). To respond

correctly, infants in these tasks are expected to extend the label to other objects that look similar to the familiar exemplars, regardless of whether they take the label as an additional feature of the objects (Sloutsky & Fisher, 2012) or as a symbol for a concept (Waxman & Gelman, 2009). Because objects of the same kind tend to have a similar appearance, these tests of word knowledge and word learning would not allow us to disentangle whether infants' interpretation of object labels is based on appearance or concept.

To investigate whether infants apply their conceptual knowledge when they interpret novel words, we examined whether they could learn labels for behaviorally defined concepts that represent situationally identified kinds in dynamic scenes. Unlike members of taxonomic kinds (e.g., shoes, bananas), for which the recognition of members does not necessarily depend on grasping concepts that define these kinds (but could be based on their

Corresponding Author:
Jun Yin, Department of Cognitive Science, Central European University, Nádor u. 9, Budapest 1051, Hungary
E-mail: yinj@ceu.hu

perceptual features), objects belonging to kinds that are defined in relational or behavioral terms can be identified only if they enter into a relation or are engaged in certain behaviors. Evidence shows that infants can interpret the actions of a human or nonhuman agent in terms of the goal it pursues (Gergely, Nádasdy, Csibra, & Bíró, 1995; Woodward, 1998). For example, 1-year-old infants can readily understand the goal of the chaser in a chasing scene (Csibra, Bíró, Koós, & Gergely, 2003; Southgate & Csibra, 2009; Wagner & Carey, 2005). Note that the concept of a chaser is situationally and behaviorally defined within a relational structure, and the perceptual appearance of the agent is not diagnostic for its recognition as a chaser (because "chaser" is a role-governed category rather than a feature-based category; see A. B. Markman & Stilwell, 2001). We exploited this dissociation in testing infants' intuition about the meaning of a novel word applied to chasers. If infants expect that words express concepts, learning a novel label for chasers (a concept they already possess) should be faster and easier for them than mapping words onto objects defined by clusters of perceptual features.

Four experiments were conducted. First, we assessed whether infants could map a word onto agents that acted as a chaser but varied in appearance (Experiment 1) or had a fixed appearance without a definite action role (Experiment 2). In Experiment 3, we offered infants the direct choice of mapping a label onto the action role of chasing or onto the agents' appearance. In Experiment 4, we tested whether the roles of the chaser and the target in a chasing scene enjoyed the same conceptual status in infants' minds (for details, see the Experiment 4 and Fig. S2 in the Supplemental Material available online).

Experiment 1

We used a looking-while-listening procedure (Swingley, 2011) to test whether infants could learn a novel word meaning "chaser." During training trials, infants were first briefly exposed to two animated agents, one chasing the other, and then the chaser was labeled with a nonsense word. Crucially, each trial presented a new pair of agents and a new trajectory of pursuit, but the agent labeled was always the chaser. During the test, infants were presented with two new pairs of agents; one of each pair acted as the chaser and the other acted as the target. The infants were asked to find the object that was the referent of the word they had heard in the training trials (trained word) or of a different word (untrained word). If the infants interpreted the trained word as referring to the role of the chaser, they would be expected to look longer at the chaser when they heard the trained word than when they heard the untrained word.

Method

Participants. Sixteen 14-month-old infants (mean age = 424 days, age range = 396–452 days) participated in this experiment. Three additional infants were tested but excluded from data analysis because they failed to reach our criteria for looking time (see Data Analysis). All of the participants were healthy, full-term infants from Hungarian-speaking families. Parents received information sheets about the experimental procedure and signed informed consent forms after learning the purpose and the procedure of the experiment.

Apparatus. A T60XL eye tracker (Tobii, Danderyd, Sweden) was used to collect the infants' gaze data. The eye tracker was integrated into a 24-in. computer monitor (resolution = 1,920 × 1,200 pixels; refresh rate = 60 Hz). The stimuli were presented on a gray background by a custom-built script written in The Psychophysics Toolbox (Brainerd, 1997) for MATLAB. We implemented the recording of gaze data in MATLAB using the Tobii Analytics Software Development Kit (http://www.tobii.com/en/eye-tracking-research/global/products/software/tobii-analytics-software-development-kit/), which enabled real-time communication with the eye tracker.

Stimuli. The visual stimuli were computer-animated chasing events displayed in a circular area subtending 31.4° in diameter at the center of the monitor on a cyan background. Each event included a pair of objects: a chaser and a target. These objects were selected randomly and without replacement from 18 geometric shapes (each subtending 2° × 2° on the screen) that were rendered with distinctive textures to make their appearance as different as possible (see Fig. S1 in the Supplemental Material).

The trajectories of the chaser and target were generated according to the following rules. The initial distance between the two objects varied uniformly between 4.3° and 12.9°. Both objects started moving at a speed of 13.7° per second, and their direction of motion was updated approximately every 100 ms. The target's direction of motion varied randomly on each update, with a uniform distribution within a 120° angular window centered on its current direction; the chaser's direction was selected randomly on each update, within a 20° angular window centered on an invisible line connecting the chaser to target. The speed of the chaser was kept constant at 13.7° per second. When the distance between the chaser and the target was less than 4.3°, the target accelerated at 0.086° per second^2.

The chaser and the target moved for 5 s (first exposure phase), followed by a 4.2-s pause during which there was no motion: In the training trials, this pause

was the labeling phase, and in the test trials, it constituted the question phase (1.2 s) and the response phase (3 s). Finally, there was an additional 2 s of motion (second exposure phase) during which the chasing action continued. The distance between the objects was always between 4.3° and 12.9°, except during the pause, when the distance was between 7.7° and 12.9°. Fifteen different pairs of trajectories were generated in advance, and each of them was rotated such that the two objects stopped at the same vertical position (i.e., they were horizontally next to each other) during the pause in motion.

In training trials, when the objects paused, an image of a human hand with a downward-pointing index finger appeared above the object that had acted as the chaser during the first exposure phase. The hand moved up and down seven times, repeatedly approaching the object and then withdrawing from it. In test trials, a bull's-eye with rapidly changing colors (1.7° in diameter) appeared between the two objects to attract the infants' attention to the location between the objects. The bull's-eye was visible for the first 1.2 s of the pause (i.e., the question phase).

The auditory stimuli were recordings of a female speaker's ostensive infant-directed utterances, in Hungarian:

- "*Szia baba! Nézd csak!*" ("Hi baby! Look!") during the final 2 s of the first exposure phase in both training and test trials
- "*Itt egy ___. Hű, egy ___!*" ("Here is a ___. Wow, a ___!") during the labeling phase of the training trials
- "*Hol van a ___?*" ("Where is the ___?") during the question phase of the test trials

Four different nonsense words were used as labels (*tacok, bitye, lad*, and *cefó*), and each of them complied with the rules of Hungarian phonotactics.

Procedure. The infants sat on their parents' laps, about 60 cm away from the eye-tracker monitor, in a dimly lit and sound-attenuated room. Parents wore opaque glasses to prevent the eye tracker from catching their gaze and to block their view of the stimuli. First, the eye tracker was calibrated with a five-point calibration procedure using the center and four corners of the screen. The calibration stimuli were displayed successively, and the presentation sequence was randomized across participants. Each stimulus was a bull's-eye with rapidly changing colors. The bull's-eyes started at a size of 2° and then shrank to a size of 1° after 0.5 s. After the five stimuli had all been displayed, the calibration results were computed immediately and reported to the experimenter. If there were fewer than three valid calibrations, the calibration session

was repeated; otherwise, the infants proceeded to the experiment.

The experiment consisted of five training trials and four test trials (Fig. 1; also see Video S1 in the Supplemental Material). At the beginning of each trial, if necessary, the infant's attention was drawn to the display by a rotating spiral and tones. Each trial presented a chasing scene, but the objects' appearance and trajectories differed across trials.

During the training trials, in the last 2 s of the first exposure phase, the infants heard the female voice say, "Hi baby! Look!" This ostensive utterance was intended to inform the infants that they were being addressed. The objects then stopped moving (labeling phase), and a downward-pointing hand moved up and down above the chaser. The infants again heard the voice, which said "Here is a ___. Wow, a ___!" (the blanks represent one of the nonsense words; e.g., *tacok*). The same nonsense word was used in all of the training trials for a particular infant, but the words varied and were counterbalanced across infants.

The test trials were similar to the training trials, except that the labeling phase was replaced with a word-recognition test consisting of a question phase (1.2 s) and a response phase (3 s). During the question phase, after the infants fixated the dynamic bull's-eye displayed between the objects, they were asked to find the object with the label given in the training trials (i.e., the trained-word condition; e.g., "Where is the *tacok*?") or to find the object with a new label (i.e., the untrained-word condition; e.g., "Where is the *bitye*?"). The words were paired (*tacok* always with *bitye, cefó* always with *lad*). Within each pair, one word always served as the trained word and the other as the untrained word for each infant. After the question, the bull's-eye disappeared, and the response phase began, during which we measured the infants' gaze response. There was no motion or sound in the response phase. Each test condition presented a new pair of objects, but within each condition, the two test trials presented the same pair of objects. Trials with the trained and the untrained label were presented in alternation, and whether the first test trial presented the trained or the untrained word was counterbalanced across infants.

Data analysis. To ensure that the infants paid sufficient attention to the stimuli, we applied predefined criteria for inclusion of data for further analyses. Specifically, valid training trials were those in which the infants (a) looked at the screen for at least half of the total time, (b) looked at the screen at least half the time during the first exposure phase, and (c) looked at the screen at least half the time during the labeling phase. Valid test trials were those in which the infants (a) looked at the screen for at least half of the first exposure phase, (b) looked at the
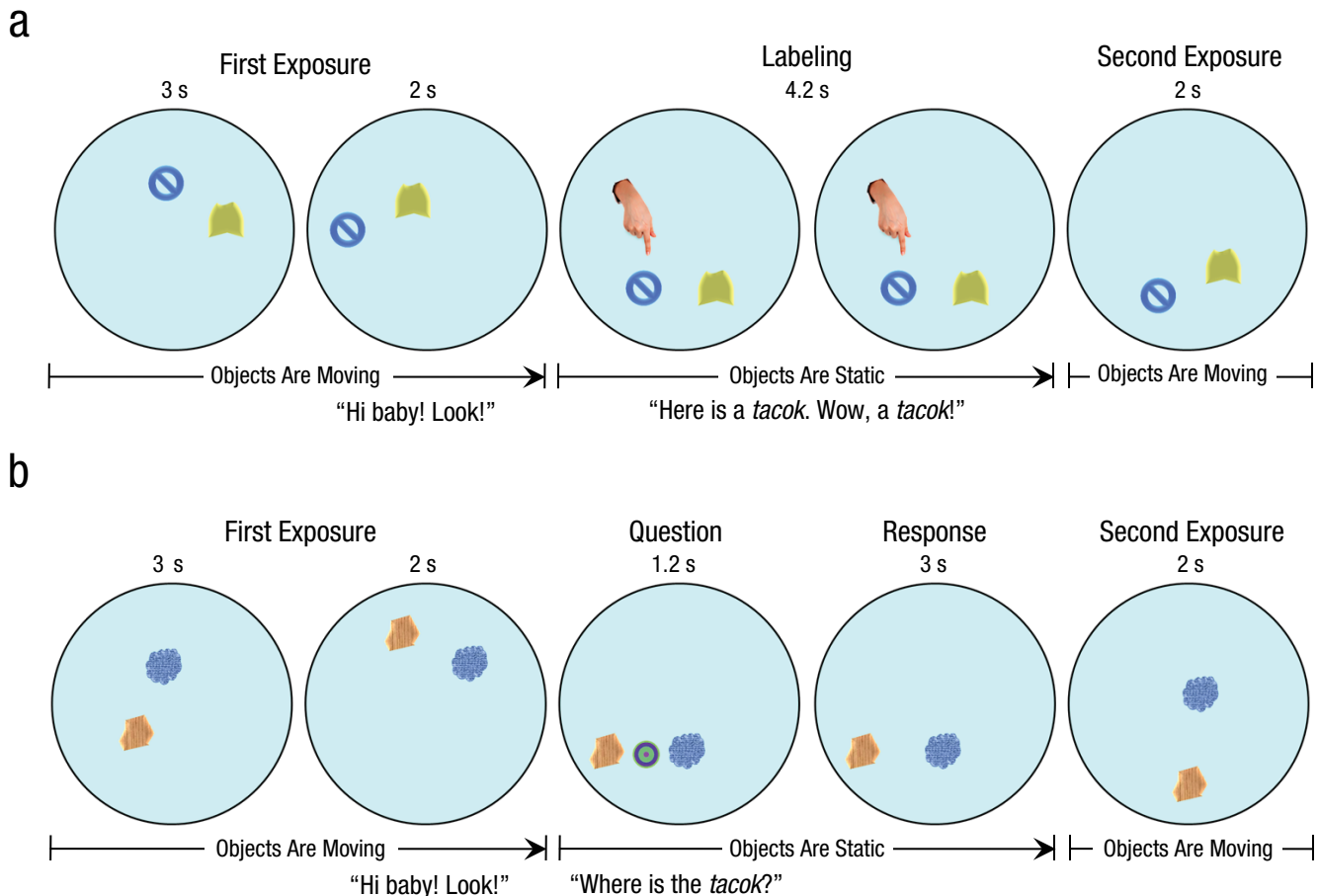
a

First Exposure

3 s   2 s

|——— Objects Are Moving ———→|

"Hi baby! Look!"

Labeling

4.2 s

|——— Objects Are Static ———→|

"Here is a *tacok*. Wow, a *tacok*!"

Second Exposure

2 s

|— Objects Are Moving —|

b

First Exposure

3 s   2 s

|——— Objects Are Moving ———→|

"Hi baby! Look!"

Question

1.2 s

|——— Objects Are Static ———→|

"Where is the *tacok*?"

Response

3 s

Second Exposure

2 s

|— Objects Are Moving —|

Fig. 1. Sequence of events in training and test trials in all experiments. Each training trial (a) had three phases: first exposure, labeling, and second exposure. The objects moved during both exposure phases, but they stopped moving during the labeling phase, during which a moving hand appeared and a labeling utterance was presented twice. Each test trial (b) had four phases: first exposure, question, response, and second exposure. The objects moved during both exposure phases, but they stopped moving during the question and response phases. A bull's-eye with rapidly changing colors appeared in the question phase but then disappeared for the response phase. During the last 2 s of the first exposure phase in both the training and the test trials, infants were addressed by an ostensive utterance.

bull's-eye between the objects at the moment when it disappeared (ensuring that the infants had an equal chance of looking at either object during the response phase), and (c) looked at the screen for at least half the time during the response phase. If an infant did not produce at least three valid training trials and at least one valid test trial in each test condition, he or she was excluded from the analyses.

To examine the infants' looking behavior during the 3-s response phase after hearing the trained or untrained word, we defined two regions of interest: a circle with a diameter of 4° centered on each of the two objects. We determined the cumulative time spent looking at each region of interest (i.e., the sum of all the looks during the response phase of a given trial). Next, we calculated a difference score (Edwards, 2001) by subtracting the cumulative time spent looking at the unlabeled object (i.e., the

target) from the cumulative time spent looking at the labeled object (i.e., the chaser) and then dividing this difference by the combined time spent looking at these objects. Difference scores ranged from +1 (looking only at the chaser) to −1 (looking only at the target). When there was more than one valid test trial within a condition, we calculated the mean of the two difference scores.

For statistical analyses, we used predefined time ranges (0–1 s, 0–2 s, and 0–3 s) to calculate cumulative difference scores during the response phase. We compared these values (a) across conditions (trained word vs. untrained word) using two-tailed paired *t* tests and (b) to zero (i.e., the chance level of no preference for either object) on each condition using one-sample *t* tests. We report Cohen's *d* for effect size. Furthermore, to test whether any effect found was due only to the particular time points we selected, we also performed a permutation-based *t* test
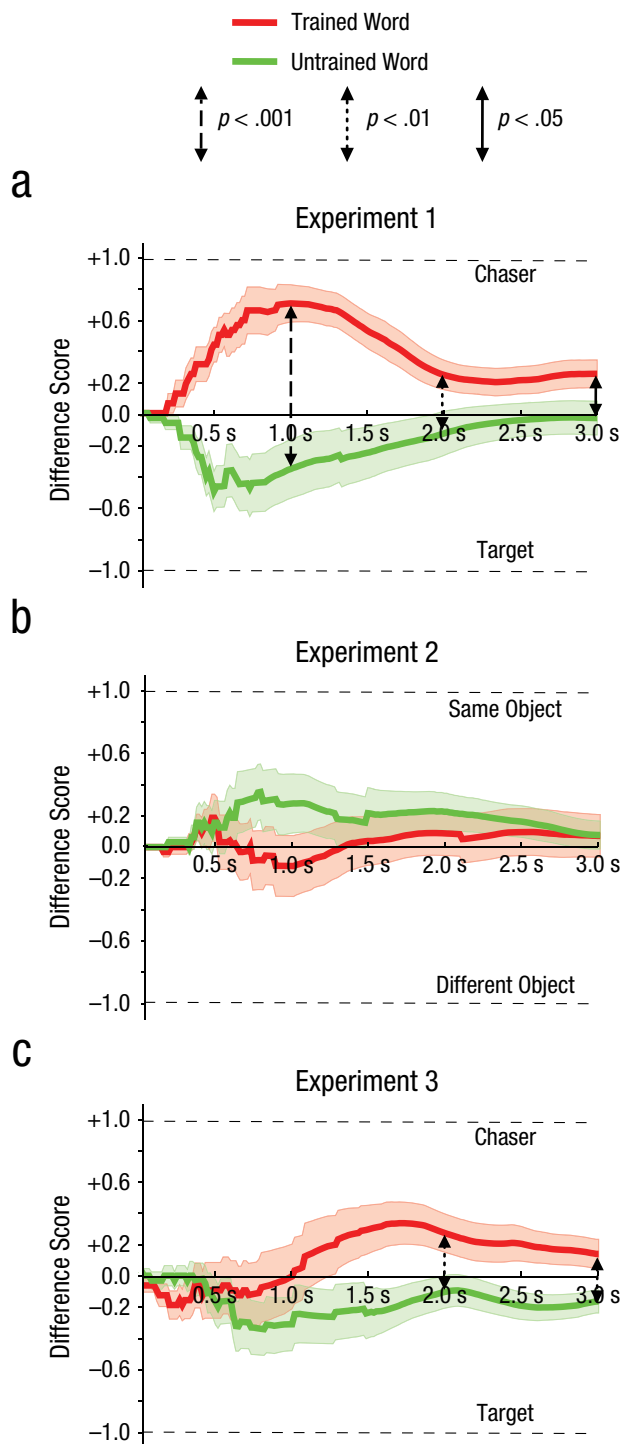
Fig. 2. Relative preference for the objects during the response phase of the test trials. The graphs show difference scores as a function of time in Experiments (a) 1, (b) 2, and (c) 3 for the trained-word condition and the untrained-word condition. The scores were calculated as the difference between cumulative time spent looking at one object (e.g., the chaser in Experiment 1) and cumulative time spent looking at the other object (e.g., the target in Experiment 1) divided by the sum of these values. We statistically compared the difference scores for the trained-word and untrained-word conditions at the predefined time points of 1, 2, and 3 s from the beginning of the response phase; the arrows indicate significant differences between conditions. The dashed lines at the top and bottom of each graph represent difference scores of +1 (looking only at the labeled object) and −1 (looking only at the other object). Shaded areas represent ±1 *SE*.

with 50,000 permutations for each comparison (Blair & Karniski, 1993). This test applies no prior assumptions about the expected time range of the effect and yields the *p* values of statistical differences at all time points with multiple-comparison correction.

Results

The average number of valid training trials was 4.81 (see Table S1 in the Supplemental Material). We analyzed the total looking times in different phases of the test trials. We found no significant difference between the trained-word condition and the untrained-word condition during either the 5-s first exposure phase, $t(15) = 1.72$, $p = .106$, $d = 0.43$, 95% confidence interval (CI) for the looking-time difference = [−0.05, 0.46], or the 3-s response phase, $t(15) = 1.52$, $p = .149$, $d = 0.38$, 95% CI = [−0.04, 0.24].

Figure 2a depicts the time course of changing preference during the response phase. As expected, during the test trials, the infants looked longer at the object that had acted as the chaser when they heard the trained word than when they heard the untrained word. This was true for all three predefined time ranges within the response phase—0–1 s: $t(15) = 5.73$, $p < .001$, $d = 1.43$, 95% CI = [0.66, 1.00]; 0–2 s: $t(15) = 3.23$, $p = .006$, $d = 0.81$, 95% CI = [0.13, 0.62]; 0–3 s: $t(15) = 2.47$, $p = .026$, $d = 0.62$, 95% CI = [0.04, 0.53] (see Fig. 2a). These effects remained significant when looking-time differences during the first exposure phase were included as covariates in the analyses, $p < .001$, $p = .018$, and $p = .037$, respectively, for the three time ranges. Such differences were not due to chance: Using permutation-based *t* tests, we found that the difference scores in the trained-word condition and the untrained-word condition were significantly different from 0.33 to 2.06 s (ps < .050).

In further analyses, we found that the difference scores for all three time ranges were significantly above the chance level in the trained-word condition—0–1 s: $t(15) = 5.88$, $p < .001$, $d = 1.47$, 95% CI = [0.45, 0.96]; 0–2 s: $t(15) = 2.59$, $p = .021$, $d = 0.65$, 95% CI = [0.04, 0.45]; 0–3 s: $t(15) = 2.84$, $p = .012$, $d = 0.71$, 95% CI = [0.06, 0.43]. This finding suggests that the infants identified the chaser as the referent of the trained word. Although the same analysis did not reveal a significant difference from zero in the untrained-word condition, an initial tendency to look longer at the target was observed—0–1 s: $t(15) = 1.91$, $p = .075$, $d = 0.48$, 95% CI = [−0.76, 0.04]; 0–2 s: $t(15) = 0.93$, $p = .368$, $d = 0.23$, 95% CI = [−0.43, 0.17]; 0–3 s: $t(15) = 0.29$, $p = .774$, $d = 0.07$, 95% CI = [−0.26, 0.20]. In additional analyses, we found no effect of order of test trials (trained word or untrained word first) or interaction of this factor with condition (all ps > .250). Thus, we established that infants could learn a word for the behaviorally defined concept "chaser" even when the perceptual appearance of the agent was not diagnostic for its recognition.

Experiment 2

The results of Experiment 1 indicated that the infants detected that it was the action role of the object, and not its appearance, that covaried with the trained label during training trials. It is possible, however, that infants will use any kind of invariance across instances of objects as the basis for mapping the labels onto objects. In that case, they would be expected to do equally well when invariance is defined by visual appearance. Experiment 2 was designed to test this prediction by presenting infants with scenes of two dynamically but independently moving agents, the appearance of which was kept constant across trials.

Method

Sixteen 14-month-old infants (mean age = 430 days, age range = 401–452 days) participated in this experiment. Seven additional infants were tested but excluded from data analysis because they failed to meet the inclusion criteria (5 infants) or we made technical errors (2 infants). The stimuli and the procedure were the same as those of Experiment 1 with the following exceptions (see Video S2 in the Supplemental Material): First, the movement trajectories of the two objects were independent; they followed the rules that generated the target's motion in Experiment 1. Second, the same two objects appeared in each training and test trial for a particular infant. Third, during the training trials, we always labeled the same object. Difference scores for the test trials were calculated in the same way as in Experiment 1 but in this case expressed the strength of preference for the labeled object over the unlabeled object.

Results

The average number of valid training trials was 4.63 (see Table S1). During the test trials, there was no significant difference in the total time the infants spent looking at the screen during the first exposure phases in the trained-word and untrained-word conditions, $t(15) = 0.29$, $p = .776$, $d = 0.07$, 95% CI = [–0.36, 0.47]. Likewise, we found no significant difference in the total time the infants spent looking at the screen during the response phases in the two test conditions, $t(15) = 0.63$, $p = .539$, $d = 0.16$, 95% CI = [–0.27, 0.14].

In the test trials, the infants displayed no preference for either object, regardless of whether they heard the trained or the untrained label (Fig. 2b). Paired t tests revealed no significant effect of test condition on the difference scores in any of the three time ranges—0–1 s: $t(15) = 1.17$, $p = .261$, $d = 0.29$, 95% CI = [–1.00, 0.33]; 0–2 s: $t(15) = 0.65$, $p = .529$, $d = 0.16$, 95% CI = [–0.60,

0.32]; 0–3 s: $t(15) = 0.03$, $p = .978$, $d = 0.01$, 95% CI = [–0.43, 0.42]. In addition, regardless of condition (trained word or untrained word), the infants' preferences within each time range did not statistically differ from chance ($ts < 1.95$, $ps > .07$, $ds < 0.48$). Even when we used permutation-based t tests, we found no significant difference between the difference scores for the trained-word conditions and the untrained-word conditions ($ps > .250$).

We found a significant interaction between the between-subjects factor of order of conditions and the within-subjects factor of test condition in the 0- to 1-s time range, $F(1, 14) = 4.59$, $p = .050$, $\eta_p^2 = .25$. Post hoc comparison revealed that the infants who heard the untrained word first looked longer at the unlabeled object when they heard the trained word than when they heard the untrained word, $t(6) = 4.97$, $p < .001$, $d = 1.88$, but there was no difference when the trained word was presented first, $t(8) = 0.35$, $p = .733$, $d = 0.12$.

Discussion

Under conditions similar to those in which 14-month-old infants had successfully learned a label for the concept of "chaser" (Experiment 1), another group of infants was unable to associate a label with the fixed visual appearance of an agent. This result does not mean that infants at this age are unable to learn labels for objects defined by their appearance (studies show that they are; see, e.g., Werker, Cohen, Lloyd, Casasola, & Stager, 1998). Rather, it suggests that this mapping does not come as rapidly and or as easily as learning words for action roles—at least in dynamic scenes. The successful studies of appearance-based word generalization among infants this age used either hand-held objects or images of static or inertially moving objects on screen. It is possible that the infants found the visual features of moving objects or agents unimportant when they considered the potential meaning of new words, and this prevented them from mapping the label to the object's appearance in this experiment. Nevertheless, the contrast between the results of Experiments 1 and 2 suggests that the kind of invariance that infants take into account when learning a new label may depend on how they construe the referents (e.g., agents or artifacts). This is consistent with an account in which early word learning is driven by the expectation of conceptual content attached to object labels.

Experiment 3

The results of Experiments 1 and 2 suggested that, when possible, infants seek an invariance that defines the concept instantiated by the object, and then apply the novel label to that invariance rather than to an invariance that

is provided by correlated, but possibly accidental, features of the object. If this is the case, infants should favor the chaser as the candidate referent even when the novel label could be equally associated with the fixed appearance of an agent. We investigated this hypothesis in Experiment 3 by fixing the appearance of the chaser and the target during training and swapping their roles during testing.

Method

Sixteen 14-month-old infants (mean age = 424 days, age range = 399–455 days) participated in this experiment. Seven additional infants were tested but were excluded from data analysis because they failed to meet the inclusion criteria (5 infants) or we made technical errors (2 infants).

The stimuli and the procedure were the same as those in Experiment 1 with the following exceptions (see Video S3 in the Supplemental Material): First, the same two objects appeared in each training and test trial for a particular infant. Second, the object that acted as the chaser in the training trials was always the same for a particular infant. Third, the object that was the chaser in the training trials was the target in the test trials, and the object that was the target in the training trials was the chaser in the test trials. Difference scores were calculated as in Experiment 1 but in this case indicated an expression of preference for the object that acted as the chaser during the test trials (i.e., the one that looked like the unlabeled object in the training trials) rather than the object that acted as the target in the test trials (i.e., the object that looked like the labeled object in the training trials).

Results

On average, the infants produced 4.75 valid training trials (see Table S1 in the Supplemental Material). We found no significant difference in total looking times between the two test conditions during either the first exposure phase, $t(15) = 1.63$, $p = .124$, $d = 0.41$, 95% CI = [–0.09, 0.70], or the response phase, $t(15) = 0.28$, $p = .783$, $d = 0.07$, 95% CI = [–0.27, 0.21].

During the response phase of the test trials, the infants looked longer at the agent that acted as the chaser in the test trial after hearing the trained label than after hearing the untrained label (Fig. 2c). Paired *t* tests on the difference scores revealed significant differences in the time ranges of 0 to 2 s and 0 to 3 s—$t(15) = 2.97$, $p = .010$, $d = 0.74$, 95% CI = [0.11, 0.64], and $t(15) = 3.07$, $p = .008$, $d = 0.77$, 95% CI = [0.09, 0.52], respectively. These effects remained significant when looking-time differences during the first exposure phase for these two time ranges were included as covariates in the analyses, $p = .023$ and

$p = .008$, respectively. No such difference was found in the 0- to 1-s time range, $t(15) = 1.17$, $p = .26$, $d = 0.29$, 95% CI = [–0.25, 0.86], which indicates that the infants detected the potential ambiguity of word meaning for the trained label and needed more time to select the agent with the same action role as the correct referent. Permutation-based *t* tests confirmed this result, yielding significant differences between the trained-word and untrained-word conditions from 1.45 s to 1.98 s and from 2.42 s to 3 s, respectively, ps < .050.

In further analyses, we found that the difference score for the 0- to 2-s time range was higher than chance level in the trained-word condition, $t(15) = 2.22$, $p = .042$, $d = 0.55$, 95% CI = [0.01, 0.53]. Within the entire response phase (0–3 s), the difference score was significantly below zero in the untrained-word condition, $t(15) = 2.26$, $p = .039$, $d = 0.57$, 95% CI = [–0.32, –0.01]. This effect indicates that the infants were more likely to link the untrained word to the object that acted like the target but looked like the chaser during the previous training trials than they were to link the untrained word to the current chaser. We found no effect of order of presentation of words (all ps > .250).

Discussion

Our finding indicates that the infants preferred the agent that played the same role (rather than the agent that looked alike) as the referent of the word applied earlier to a chaser. In other words, they mapped the label onto a concept defined by behavior rather than by similarity-defining perceptual features.

In addition, the infants seem to have assumed that the novel word referred to the current target (which looked like the chaser during the training trials). Although this result was not explicitly predicted, it is interesting to note that whenever we found evidence of learning of object labels (Experiments 1 and 3), we also found that the infants tended to look toward the unlabeled object when they heard the novel word. This phenomenon might be the result of applying the logic of mutual exclusivity during mapping novel words to referents (Halberda, 2003; E. M. Markman & Wachtel, 1988). However, a discussion of this effect is beyond the scope of this article.

General Discussion

We found that it was easier for 14-month-old infants to learn a novel label for an agent's action role than for an agent's appearance. The difficulty in linking the label to the agent's visual appearance (or even the failure to do so) cannot be explained by the infants' inability to identify or memorize them, because such information exerted an interference effect when it conflicted with action roles

in Experiment 3, where it delayed the infants' responses. The success of mapping a nonsense word onto the chaser's role must have been based on understanding its goal (i.e., following or catching the other object) rather than its appearance or individual motion patterns. Being a chaser is a rule-governed relational property; as such, it would not be manifest in the static or dynamic perceptual features of the object. The infants were unable to associate a novel word with the target (see Fig. S2 and Experiment 4 in the Supplemental Material). This finding suggests that when searching for word meaning, infants do not consider just any kind of dynamic relational properties; they consider only the ones that indicate action goals. It is also unlikely that the infants linked the word not with the agent but with the action itself, because the objects were static during both labeling and testing. Thus, the infants mapped the novel labels to a concept that cannot be defined purely by perceptual features but is known to be part of their representational repertoire.

Nevertheless, it is possible that the concept that the infants applied in our experiments is more abstract than the one defined by the specific goal of the chaser or the specific interaction in which it was engaged. For instance, instead of representing the object as a "chaser," the infants might have represented it more generally as a "goal-directed agent." In fact, because detecting its goal makes the chaser's behavior predictable, the infants in these experiments might have mapped the label to the (even more) abstract concept of "the predictable one" (note that during labeling, both objects were static; hence this mapping could be achieved only if predictability was attached to the object as a dispositional property). These representations would also have allowed them to transfer the meaning of the word from one situation (training) to another (testing). Further studies should address the question of specificity of representation that infants interpret as the referential content of a word. However, the current study already demonstrated that 14-month-olds learned a label more easily for a behaviorally defined abstract concept (such as "chaser, "goal-directed agent," or "predictable entity") than for an object characterized by certain perceptual features, which suggests that infants expect that words express concepts.

The pragmatic context in which infants are exposed to a novel word may also influence how they interpret it (Akhtar & Tomasello, 2000). The ostensive nature of labeling in our experiments might have induced specific expectations about the link between the referent and the novel label (Csibra & Gergely, 2009; Csibra & Shamsudheen, 2015). However, infants readily acquire words, such as object labels, outside ostensive contexts during the 2nd year of life (Akhtar, Jipson, & Callanan, 2001), and it is possible that many of them are not initially mapped onto preexisting concepts. Nevertheless, our study provides solid evidence that, at least for certain concepts and certain presentation contexts, concept-based word learning is already operating at the earliest stage of lexical acquisition.

Author Contributions

J. Yin and G. Csibra conceived and designed the experiments. J. Yin performed the experiments and analyzed the data. J. Yin and G. Csibra wrote the manuscript.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Supplemental Material

Additional supporting information can be found at http://pss .sagepub.com/content/by/supplemental-data

Open Practices

All data and materials have been made publicly available via Open Science Framework and can be accessed at https://osf .io/aegr4/. The complete Open Practices Disclosure for this article can be found at http://pss.sagepub.com/content/by/ supplemental-data. This article has received badges for Open Data and Open Materials. More information about the Open Practices badges can be found at https://osf.io/tvyxz/ wiki/1.%20View%20the%20Badges/ and http://pss.sagepub .com/content/25/1/3.full.

References

Akhtar, N., Jipson, J., & Callanan, M. A. (2001). Learning words through overhearing. *Child Development*, *72*, 416–430.

Akhtar, N., & Tomasello, M. (2000). The social nature of words and word learning. In R. M. Golinkoff & K. Hirsh-Pasek (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 115–135). New York, NY: Oxford University Press.

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences, USA*, *109*, 3253–3258.

Blair, R. C., & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology*, *30*, 518–524.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433–436.

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science, 27*, 111–133.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*, 148–153.

Csibra, G., & Shamsudheen, R. (2015). Nonverbal generics: Human infants interpret objects as symbols of object kinds. *Annual Review of Psychology, 68*, 689–710.

Edwards, J. R. (2001). Ten difference score myths. *Organizational Research Methods, 4*, 265–287.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*, 165–193.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition, 87*, B23–B34.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development, 3*, 299–321.

Macnamara, J. (1982). *Names for things: A study of human learning*. Cambridge, MA: MIT Press.

Markman, A. B., & Stilwell, C. H. (2001). Role-governed categories. *Journal of Experimental & Theoretical Artificial Intelligence, 13*, 329–358.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology, 20*, 121–157.

Parise, E., & Csibra, G. (2012). Electrophysiological evidence for the understanding of maternal speech by 9-month-old infants. *Psychological Science, 23*, 728–733.

Plunkett, K., Hu, J.-F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition, 106*, 665–681.

Sloutsky, V. M., & Fisher, A. V. (2012). Linguistic labels: Conceptual markers or object features? *Journal of Experimental Child Psychology, 111*, 65–86.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science, 13*, 13–19.

Southgate, V., & Csibra, G. (2009). Inferring the outcome of an ongoing novel action at 13 months. *Developmental Psychology, 45*, 1794–1798.

Swingley, D. (2011). The looking-while-listening procedure. In E. Hoff (Ed.), *Research methods in child language: A practical guide* (pp. 29–42). Hoboken, NJ: Wiley-Blackwell.

Wagner, L., & Carey, S. (2005). 12-month-old infants represent probable ending of motion events. *Infancy, 7*, 75–83.

Waxman, S. R., & Gelman, S. A. (2009). Early word-learning entails reference, not merely associations. *Trends in Cognitive Sciences, 13*, 258–263.

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology, 34*, 1289–1309.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*, 1–34.