BIROn - Birkbeck Institutional Research Online

Title: Concept-based word learning in human infants

Running head: CONCEPT-BASED WORD LEARNING

Authors: Jun Yin[1*] & Gergely Csibra[1,2]

Affiliations:

[1]Cognitive Development Center, Department of Cognitive Science, Central European

University, Nádor u. 9., Budapest 1051, Hungary.

[2]Department of Psychological Sciences, Birkbeck, University of London, UK.

[*] To whom correspondence should be addressed (yinj@ceu.hu).

Abstract

It is debated whether infants initially learn object labels by mapping them onto similarity-defining perceptual features or onto concepts of object kinds. 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 the chaser plays 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 actor's role within the observed interaction rather than as being associated with the actor's appearance. These results demonstrate that infants can learn labels as easily, or even easier, for concepts identified by abstract behavioral characteristics than by perceptual features. Thus, already at early stages of word learning, infants expect that novel words express concepts.

Keywords: word learning, concepts, object kinds, perceptual similarity, chasing action

Introduction

People often label objects when talking to preverbal infants. Although such labeling is usually applied to specific objects, infants are able to 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 a matter of controversies. On the one hand, infants have been proposed to 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). On the other hand, infants, similarly to adults, may be able to conceive 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 employ 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). In order to respond correctly, infants in these tasks are expected to extend the label to other objects whose appearance is similar to the familiar exemplars, 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 belonging to the same kind tend to be similar to each other in visual appearance, these tests of word knowledge and word learning would not allow us to disentangle whether infants' interpretation of object labels is appearance-based or concept-based.

To investigate whether infants appeal to 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 (such as shoes, bananas), where the recognition of kind members does not necessarily depend on grasping concepts that defines these kinds (but could be based on their perceptual features), objects belonging to kinds that are defined in relational or behavioral terms can only be identified if they enter into a relation or are engaged in certain behaviors. Evidence shows that infants can interpret the actions of a human or non-human agent in terms of the goal it pursues (Gergely, Nádasdy, Csibra, & Bíró, 1995; Woodward, 1998). For example, it has been demonstrated that 1-year-old infants can readily understand the goal that the chaser pursues in a chasing scene (Csibra, Bíró, Koós, & Gergely, 2003; Southgate & Csibra, 2009; Wagner & Carey, 2005). Importantly, the concept of a chaser is situationally and behaviorally defined within a relational structure, and the perceptual appearance of the agent is not diagnostic to its recognition as a chaser (it is a role-governed rather than a feature-based category; see Markman & Stilwell, 2010). 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 can map a word onto agents that play the role of a chaser but vary in appearance (Experiment 1), or exhibit fixed appearance without a definite action role (Experiment 2). Experiment 3 offered the direct choice to infants to map a label onto the action role of chasing or onto the appearance of agents. Experiment 4 (reported in the Supplemental Material) tested whether the role of the chasee in a chasing scene enjoys the same conceptual status in the infant mind as that of a chaser.

Experiment 1

A looking-while-listening procedure (Swingley, 2011) was employed to test whether infants could learn a novel word meaning 'chaser.' During training trials, infants were first briefly exposed to an animated agent chasing another one, and then the chaser was labeled with a nonsense word. Crucially, each trial presented a pair of novel-looking actors and a new trajectory of chasing, but the labeled actor was always the chaser. At test, infants were presented with two further novel actors chasing each other, and were asked to find an object as the referent of the word they had heard in the training trials (trained word) or of a different (untrained) word. If infants interpreted the trained word as referring to the role of the chaser, they should look more at the chaser upon hearing the trained word, compared to the untrained word.

Methods

Participants. Sixteen 14-month-old infants (mean age: 424 days; range: 396-452 days) participated in this experiment. Three additional infants were tested but excluded from data analysis due to failing to reach the criteria of 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 understanding the purpose and the procedure of the experiment.

Apparatus. A Tobii T60XL eye tracker (Dandreyd, Sweden) was used to collect infants' gaze data. The eye tracker was integrated into a 24-inch computer monitor with 1920×1200 pixels spatial and 60 Hz temporal resolution. The stimuli were presented on a grey background by a custom-built script written in MATLAB Psychophysics Toolbox. The recording of gaze data was implemented in MATLAB by Tobii Analytics Software Development Kit (Tobii Analytics SDK, 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 of 31.4° at the center of the monitor with cyan color background. Each event included a pair of objects: a chaser and a chasee. These objects were randomly and non-repeatedly sampled from eighteen geometrical shapes (each subtending $2° \times 2°$ on the screen), which were rendered with distinctive textures to make their appearance differ as much as possible (Figure S1). The trajectory of the chaser and chasee was generated according to the following rules. The initial locations of the two objects were chosen randomly with the constraint of having a distance of 4.3° to 12.9° between them. Both objects started moving with the speed of 13.7 °/s, and their moving direction was updated approximately every 100 ms. The chasee's direction of motion randomly varied with uniform distribution within 120° angular window centered on its current direction, while the chaser's direction on each update was selected randomly within a 20° angular window centered on the line connecting the chaser to chasee. The speed of the chaser was kept constant at 13.7 °/s, and the chasee accelerated its speed at 0.086 °/s² when the distance between them became less than 4.3°. The distance between the objects was kept between 4.3° and 12.9° during the whole event, except that their distance was constrained to be at least 7.7° when they stopped at 5 s after the start of the event. After 4.2 s pause, the two objects restored their motion pattern for an additional 2 s as if they had not stopped. Fifteen different trajectories were generated in advance, and each of them was rotated in a way that the two objects stopped at the same vertical position (i.e., they were horizontally next to each other).

During the training trials, when the objects stopped, the image of a human hand with an

extended index finger pointing downwards appeared above the object that had played the role of the chaser previously. This image moved up and down 7 times during the 4.2 s pause, repeatedly approaching the object and then withdrawing from it. During the test trials, when the objects stopped, colored concentric discs (1.7°, 4 to 8 colors) appeared on top of each other in a bull's eye fashion between the two objects. The colors changed continuously every 17 ms during the first 1.2 s of 4.2 s pause, to attract infants' attention the location between the objects before measuring their gaze response.

The auditory stimuli were uttered by a female speaker of Hungarian in an infant-directed manner: "*Szia baba! Nézd csak!*" (Hi baby! Look!) during the ostension phase, "*Itt egy tacok. Hű, egy tacok!*" (Here is a *tacok*. Wow, a *tacok*!") during the labeling phase, and "*Hol van a tacok?*" (Where is the *tacok*?) during the question phase. 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. Infants sat on their parents' lap in a dimly lit and sound attenuated room, about 60 cm away from the eye-tracker monitor. Parents wore opaque glasses to prevent the eye tracker from catching their gaze and to block their sight of the stimuli. First, the eye tracker was calibrated with a five-point calibration procedure using the center and the four corners of screen. The calibration stimuli were displayed successively and the presentation sequence was randomized across participants. Each stimulus started as a 2° colored bull's eye with 4 to 8 colors, which then shrank into 1° size after 0.5 s. After finishing displaying the five stimuli, the calibration results were computed immediately and reported to the experimenter. If the number of valid calibrations was less than three, the experimenter repeated the calibration session, otherwise proceeded to the experiment.

The experiment consisted of five training trials and four test trials (Figure 1 and Video S1). At the beginning of each trial, the experimenter drew the infant's attention to the display by presenting a rotating spiral and tones. Each training trial presented a chasing scene, but the appearance of the objects and their trajectories differed across trials. The two objects moved for 5 s (exposure phase), then stopped for 4.2 s (labeling phase), and finally continued to move for 2 s (more-exposure phase). In the last 2 s of the exposure phase (ostension phase), the infant was addressed by the female voice (Hi baby! Look!). The same word was presented during the labeling phase in each trial (e.g., Here is a *tacok*. Wow, a *tacok*!), but the words varied and were counterbalanced across infants.
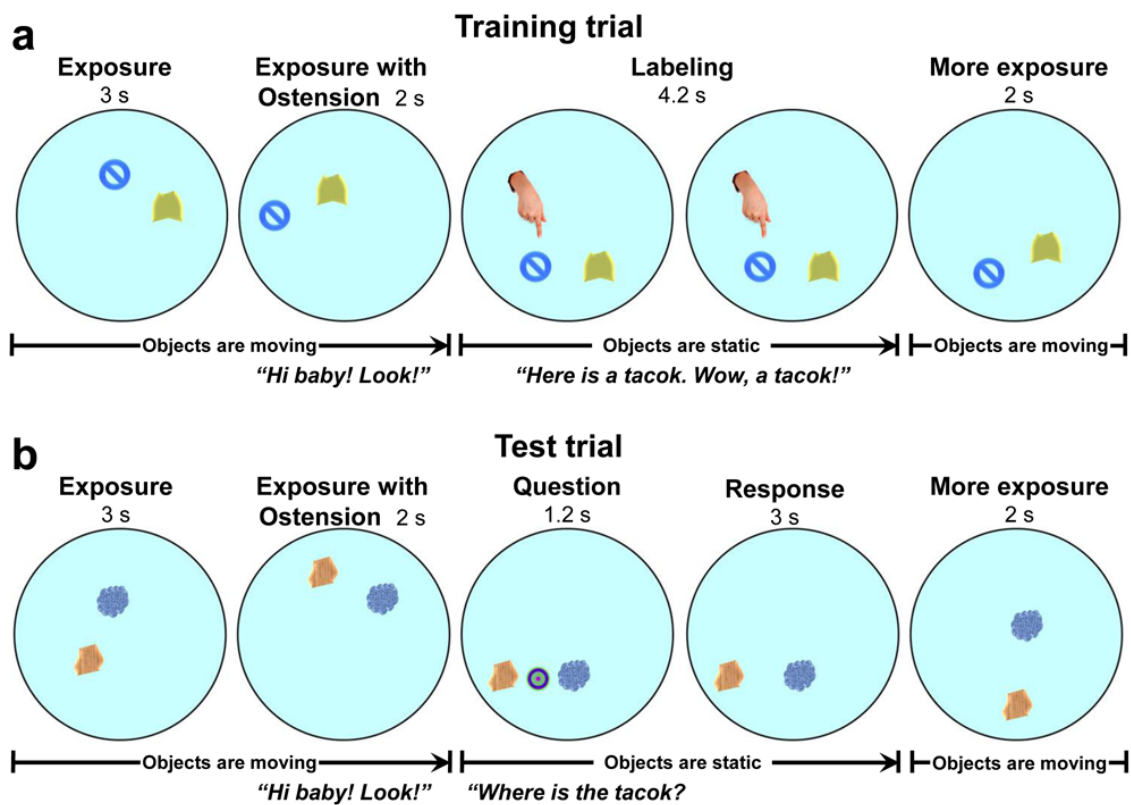

Figure 1. ***Experimental procedure during training (a) and test (b) trials***. In Experiments 1 & 4, each trial presented a new pair of objects. In Experiments 1, 3 & 4, one object chased the other one; in Experiment 2, they moved independently. In Experiment 3, the two objects swapped their roles in chasing during the test trials. In Experiment 1 and 3 the chaser, in Experiment 4 the chasee, and in Experiment 2 an object with constant appearance was labeled during the training trials. Half of the test trials presented the trained word, and the other half presented an untrained word.

The test trials had a similar structure to the training trials, except that the labeling phase was replaced by a word recognition test consisting of a question phase and a response phase. During the question phase, after infants' fixated the dynamic bull's eye displayed between the objects, they were asked to find the object with the label used in the training trials (trained word, e.g., Where is the *tacok*?) or with a new label (untrained word, e.g., Where is the *bitye*?). The words were paired with each other (tacok with bitye, cefó with lad), one serving as trained, the other as untrained label for each infant. After the question (1.2 s), the bull's eye disappeared and the 3-s response phase started, during which no change occurred on the scene and no auditory stimuli were presented. Each test trial presented a new pair of objects. The trials with trained and untrained label alternated, and whether the first test trial presented the trained or untrained word was counterbalanced across participants.

Data analysis. In order to ensure that infants paid sufficient attention to the stimuli, we applied pre-defined criteria for inclusion of data for further analyses. Specifically, for a training trial to be judged valid, infants were required (1) to look at the screen at least 50% of the total time; (2) to look at the screen at least 50% during the exposure phase before two objects stopped; and (3) to look at the screen at least 50% during the labeling phase. For a test trial to be marked valid, infants were required (1) to look at the screen at least 50% during the exposure phase before two objects stopped; (2) to look at the dynamic attention getter between the objects at the point when it disappeared (ensuring that infants had equal amount of chance to look at either object during the response phase); and (3) to look at the screen at least 50% during the response phase. If infants did not produce at least three valid training trials and at least one valid test trial in each test condition according to the aforementioned criteria, they were excluded from the analyses.

To examine infants' looking behavior during the 3-s response phase after hearing the trained or untrained word, regions of interest (ROIs) were defined as circles of 4° diameter centered around the two objects. We determined the cumulative looking time at each ROI (the sum of all the looks during the response phase), and computed a difference score (Edwards, 2001) as a function of time elapsed since the start of the response phase. The difference scores were calculated according to the following formula: $DS = (LT_1 - LT_2) / (LT_1 + LT_2)$, where LT_1 refers to the looking time to the object to which the label was to be attached (i.e., the chaser), and LT_2 refers to the looking time to the other object. DS would range from +1 (looking only at the labeled object) to -1 (looking only at the other object). When there were more than one valid test trials within a condition, we calculated the mean of the two difference scores.

For statistical analyses, we used the pre-defined measures of the values of the cumulative difference scores at 1, 2, and 3 s after the start of the response phase. We compared these values across conditions (trained vs. untrained word) in two-tailed paired t-tests, as well as to the chance level of zero on each condition via one-sample t-tests, and report Cohen's *d* as effect size. Furthermore, to test whether any potential effect found would be due to chance significance at these pre-defined time points, we also performed permutation-based t-tests 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 computes the statistical differences at all time points with multiple-comparison correction.

Results

The average number of valid training trials was 4.81 (Table S1). We analyzed the total looking times in different phases of the test trials. We found no significant difference between trials with trained word and untrained word either during the 5-s-long initial exposure phase

($t(15) = 1.72$, $p = 0.106$, $d = 0.43$, 95%CI = [-0.05 0.46]) or during the 3-s-long response

phase ($t(15) = 1.52$, $p = 0.149$, $d = 0.38$, 95%CI = [-0.04 0.24]).

Figure 2a depicts the dynamically changing preference during the response phase. As

expected, infants looked more at the object that had played the chaser during the test trial

when hearing the trained word than when hearing the untrained word in all three pre-defined

time ranges within the response phase (0-1 s: $t(15) = 5.73$, $p < 0.001$, $d = 1.43$, 95%CI = [0.66

1.00]; 0-2 s: $t(15) = 3.23$, $p = 0.006$, $d = 0.81$, 95%CI = [0.13 0.62]; 0-3 s: $t(15) = 2.47$, $p =$

0.026, $d = 0.62$, 95%CI = [0.04 0.53]; see Figure 2a). These effects remained significant

when looking time differences during the initial exposure phase were included as covariates

in the analyses ($p < 0.001$, $p = 0.018$, $p = 0.037$, respectively, for the three time ranges). Such

differences were not due to chance significance at these pre-defined time points, since using

permutation-based t-tests we also found that the difference scores significantly differed

between the two test conditions from 0.33 s to 2.06 s during the response phase (ps < 0.050).
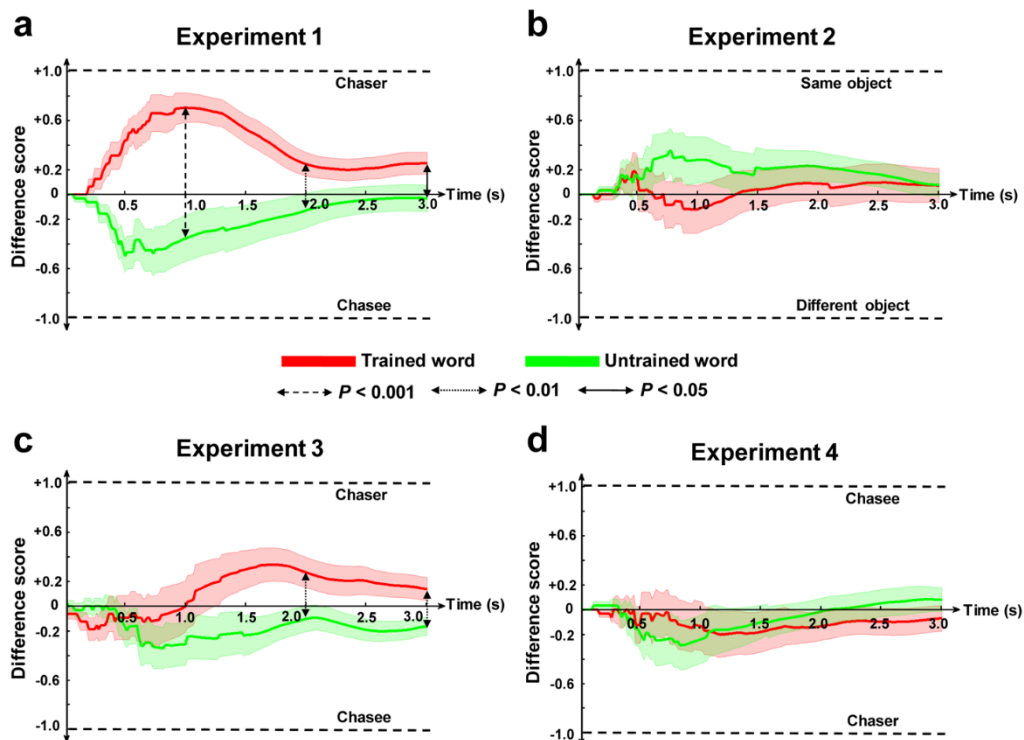
Figure 2. ***Preference to one or the other object during the response phase of the test trials***. Difference scores in Experiments 1 to 4 (a to d) for trained words and untrained words were continuously calculated as the ratio of the difference between cumulative looking time to one (e.g., chaser in Experiment 1) and the other object (e.g., chasee in Experiment 1) divided by the sum of these values. Shaded area represents standard errors (±SE). The difference scores were statistically compared between trials with the trained and untrained words at the pre-defined time points of 1, 2, and 3 s from the beginning of the response phase.

Further analyses revealed that the difference scores in all three time ranges were significantly above the chance level when the trained word was presented (0-1 s: $t(15) = 5.88$, $p < 0.001$, $d = 1.47$, 95%CI = [0.45 0.96]; 0-2 s: $t(15) = 2.59$, $p = 0.021$, $d = 0.65$, 95%CI = [0.04 0.45]; 0-3 s: $t(15) = 2.84$, $p = 0.012$, $d = 0.71$, 95%CI = [0.06 0.43]), suggesting that infants identified the chaser as the referent of the trained word. Although the same analysis did not yield significant difference from zero in the untrained word condition (0-1 s: $t(15) = 1.91$, $p = 0.075$, $d = 0.48$, 95%CI = [-0.76 0.04]; 0-2 s: $t(15) = 0.93$, $p = 0.368$, $d = 0.23$, 95%CI = [-0.43 0.17]; 0-3 s: $t(15) = 0.29$, $p = 0.774$, $d = 0.07$, 95%CI = [-0.26 0.20]), an initial tendency to look longer at the chasee was observed. 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 > 0.250). 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 to its recognition.

Experiment 2

The results of Experiment 1 indicate that infants detected that it was the action role, and not the appearance, of the object that co-varied with the trained label during training trials. It is possible, however, that all that infants cared about was to find any kind of invariance across instances of objects onto which they could map the label. If this is case, they should do equally well when this invariance is defined by visual appearance. Experiment 2 was

designed to test this prediction by presenting participants with scenes of two dynamically but

independently moving (i.e., not chasing) agents. Importantly, the appearance of the two

agents was kept constant across trials, and always the same object was labeled during

training.

Methods

Sixteen 14-month-old infants (mean age: 430 days; range: 401-452 days) participated in

this experiment. Seven additional infants were tested but excluded from data analysis due to

failing to meet the inclusion criteria (5) and technical errors (2).

The stimuli and the procedure were same as those of Experiment 1 with the following

exceptions (Video S2). (1) The movement trajectories of the two objects were independent

and symmetrical. Both of them followed the rules that generated the chasee's motion in

Experiment 1. (2) The same two objects appeared in each training and test trial, and during

the training trials always the same object was labeled. (3) Difference scores were calculated

as expressing preference for the same object against the other (not labeled) object during the

test trials.

Results

The average number of valid training trials was 4.63 (see Table S1). During the test

trials, the total time of looking at the screen during the initial exposure phase showed no

significant difference between the trained and untrained word conditions ($t(15) = 0.29$, $p =$

0.776, $d = 0.07$, 95%CI = [-0.36 0.47]). Likewise, we did not find a significant difference

between two test conditions on the looking times during the response phase ($t(15) = 0.63$, $p =$

0.539, $d = 0.16$, 95%CI = [-0.27 0.14]).

In the test trials, infants displayed no preference for either object, whether they heard the

trained or the untrained label (Figure 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: $t(15) = 1.17$, $p = 0.261$, $d = 0.29$, 95%CI = [-1.00 0.33]; 0-2 s: $t(15) = 0.65$, $p = 0.529$, $d = 0.16$, 95%CI = [-0.60 0.32]; 0-3 s: $t(15) = 0.03$, $p = 0.978$, $d = 0.01$, 95%CI = [-0.43 0.42]). In addition, regardless of the test trial type (trained vs. untrained word), infants' preferences within each time range did not statistically differ from the chance level (ts < 1.95, ps > 0.07, ds < 0.48). Even using permutation-based t-tests, no significant difference was found between the difference scores upon hearing the trained word and an untrained word (ps > 0.250).

When testing for order effects, we found a significant interaction with test condition in the 0-1-s time range ($F(1,14) = 4.59$, $p = 0.050$, $\eta_p^2 = 0.25$). Post-hoc comparison revealed that infants who heard the untrained word first looked more at the unlabeled object upon hearing the trained word than the untrained word ($t(6) = 4.97$, $p < 0.001$, $d = 1.88$), but there was no difference when the trained word was presented first ($t(8) = 0.35$, $p = 0.733$, $d = 0.12$).

Discussion

Under similar conditions in which they successfully learned a label for the concept of 'chaser' (Experiment 1), 14-month-old infants were unable to associate a label with the constant visual appearance of an agent. This result does not entail 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), but it suggests that this mapping does not come as rapidly and easily as learning words for action roles — at least in dynamic scenes. The successful studies of appearance-based word generalization at this age used either hand-held objects or static or inertly moving images of objects on screen. It is possible that infants find the visual features of moving objects/agents less relevant when they consider the potential meaning of new words, and this prevented them from mapping the label to

object 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 the account that proposes that early word learning is driven by the expectation of conceptual content attached to object labels.

Experiment 3

The results of Experiments 1 & 2 suggest that, whenever it is possible, infants seek to find an invariance that defines the concept instantiated by the object onto which the novel label is applied, rather than 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 equally be associated with constant appearance of an agent. We investigated this hypothesis in Experiment 3 by keeping the appearance of the chaser and the chasee constant during training, and swapping their roles during test.

Methods

Sixteen 14-month-old infants (mean age: 424 days; range: 399-455 days) participated in this experiment. Seven additional infants were tested but were excluded from data analysis due to failing to meet the inclusion criteria (5) and technical errors (2).

The stimuli and the procedure were same as those of Experiment 1 with the following exceptions (Video S3): (1) The same two objects appeared in each training and test trial. (2) During the training trials, the same object played the role of the chaser. (3) During the test trials, the same two objects performed the chasing action, but their appearance (or role in the action) was swapped. (4) Difference scores were calculated as expressing preference for the

object that played the role of the chaser during the test trials against the object that looked the

same as the labeled object (i.e., the chasee in the test) during the training trials.

Results

On average, infants produced 4.75 valid training trials (see Table S1). We did not find

significant differences of total looking times between the two test conditions during either the

initial exposure phase (t(15)= 1.63, p = 0.124, d = 0.41, 95%CI = [-0.09 0.70]) or the

response phase (t(15) = 0.28, p = 0.783, d = 0.07, 95%CI = [-0.27 0.21]).

During the response phase, infants looked longer at the actor who played the role of the

chaser in the test trial after having heard the trained label than after having heard the

untrained label (Figure 2c). Paired t-tests on the difference scores revealed that in the time

ranges of 0-2 s (t(15) = 2.97, p = 0.010, d = 0.74, 95%CI = [0.11 0.64]) and 0-3 s (t(15) =

3.07, p = 0.008, d = 0.77, 95%CI = [0.09 0.52]) infants looked significantly longer at the

object acting as the chaser in the test when hearing the trained word than when hearing an

untrained word. These effects remained significant when looking-time differences during the

initial exposure phase were included as covariates in the analyses (p = 0.023, p = 0.008,

respectively, for the two time ranges). No such difference was found in the time range of 0-1

s (t(15) = 1.17, p = 0.26, d = 0.29, 95%CI = [-0.25 0.86]), indicating that 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 (ps < 0.050).

Further analyses found that the difference score computed from 0 to 2 s was higher than

chancel level upon hearing the trained word (t(15) = 2.22, p = 0.042, d = 0.55, 95%CI = [0.01

0.53]), and within the entire response phase (0-3 s) the difference score was significantly

lower than baseline when presenting the untrained word ($t(15) = 2.26$, $p = 0.039$, $d = 0.57$, 95%CI = [-0.32 -0.01]). The latter effect indicates that infants were more likely to link the untrained word to the object that acted as the chasee but looked like the chaser during the previous training trials than to the current chaser. We found no effect of order of presentation of words (all ps > 0.250).

Discussion

Our finding indicates that infants prefer to choose the actor who played the same role, rather than the one who had the same visual appearance, as the referent of the word applied earlier to a chaser. In other words, they rather map the label onto a behaviorally defined concept than onto a featurally defined object.

In addition, the participants seem to have assumed that the novel word referred to the current chasee (who looked like the chaser during the training trials). While this result was not explicitly predicted, it is interesting to note that whenever we found evidence of learning object labels (Experiments 1 and 3), we also found that infants tended to look towards the label-less object upon hearing the novel word. This phenomenon might be the result of applying the logic of mutual exclusivity during mapping novel words to referents (Halberda, 2003; Markman & Wachtel, 1988). However, discussing this effect is beyond the scope of this paper.

General Discussion

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

action roles in Experiment 3, where it delayed infants' responses. The success of mapping a nonsense word onto the chaser's role must have been based on understanding the goal it pursued (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, which would not be manifest in the static or dynamic perceptual features of the object. Moreover, the fact that infants were unable to associate a novel word with the chasee (Experiment 4 in the Supplemental Material) suggests that they consider not any kind of dynamic relational properties but only the ones that indicate action goals when searching for word meaning. It is also unlikely that infants linked the word not with the agent but with the action itself, because the objects were static during both labeling and testing. Thus, 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 infants appealed to in our experiments is more abstract than what was defined by the specific goal that the chaser pursued or the specific interaction in which it was engaged. For instance, instead of representing the object as a 'chaser,' infants might have represented it more generally as a 'goal-directed agent.' In fact, since detecting its goal makes the chaser's behavior predictable, 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 only be achieved if predictability is 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 (test). Further studies should address the question of specificity of representation that enables infants to map a word onto. However, the current study already

demonstrates that 14-month-olds learn 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.

ᴛhe pragmatic context in which infants are exposed to a novel word may also influence how they will 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 second year of life (Akhtar, Jipson, & Callanan, 2001), and it is possible that many of them are not initially mapped onto pre-existing 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.Y. and G.C. conceived and designed the experiments. J.Y. performed the experiments and analyzed the data. J.Y. and G.C. wrote the manuscript.

References

Akhtar, N., Jipson, J., & Callanan, M. A. (2001). Learning words through overhearing. *Child Development*, *72*(2), 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 of the United States of America*, *109*(9), 3253–3258.

Blair, R. C., & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology*, *30*(5), 518–24.

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, *27*(1), 111–133.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 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*(3), 265–287.

Gergely G., Nádasdy Z., Csibra G., & Biró S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*(2), 165-193.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, *87*(1), B23–B34.

Macnamara, J. (1982). *Names for Things: A Study of Human Learning.* Cambridge, MA: MIT Press.

Markman, A. B., & Stilwell, C. H. (2010). Role-governed categories. *Journal of Experimental & Theoretical Artificial Intelligence*, *13*(4), 329–358.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121–157.

Parise, E., & Csibra, G. (2012). Electrophysiological evidence for the understanding of maternal speech by 9-month-old infants. *Psychological Science*, *23*(7), 728–33.

Plunkett, K., Hu, J.-F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition*, *106*(2), 665–681.

Sloutsky, V. M., & Fisher, A. V. (2012). Linguistic labels: Conceptual markers or object features? *Journal of Experimental Child Psychology*, *111*(1), 65–86.

Southgate, V., & Csibra, G. (2009). Inferring the outcome of an ongoing novel action at 13 months, *45*(6), 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*(1): 75-83.

Waxman, S. R., & Gelman, S. A. (2009). Early word-learning entails reference, not merely

associations. *Trends in Cognitive Sciences*, *13*(6), 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*(6),

   1289–1309.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach.

   Cognition, *69*(1), 1–34.

Supplementary Material

Experiment 4

Experiments 1 to 3 found that it was easier for 14-month-old infants to learn a novel label for an abstract concept (i.e., the action role that the chaser plays in the chasing event) than for certain combinations of perceptual features (i.e., visual appearance). Experiment 4 attempted to answer the question of what makes 'chaser' a good candidate concept for mapping a novel label on. One possibility is that it is the well-defined goal of the chaser's action that invites the binding of a new label. In this case, the chasee may not be interpreted as the potential referent of a novel word, because this actor, though its presence is a necessary enabling condition for a chasing action, does not pursue a goal (for example, it was not always showing evidence of 'fleeing' from the chaser in our animations). Another possibility is that what makes 'chaser' an easily nameable concept is that it is a social role in a well-defined interaction. In this case, its complementary concept ('chasee') would also provide a case for easy label learning.

Methods

Sixteen 14-month-old infants (mean age: 432 days; range: 405-453 days) participated in this experiment. Five additional infants were tested but excluded from data analysis due to failing to meet the inclusion criteria (4) or technical errors (1).

Stimuli, procedure and data analysis were almost the same as in Experiment 1, except that, in the training trials, the agent acting as the chasee was labeled, and difference scores were calculated as expressing preference for the chasee against the chaser.

Results

The number of valid training trials was 4.63 on average (Table S1). During the test trials,

there was no significant difference between looking times to the test conditions during either the initial exposure phase ($t(15) = 0.45$, $p = 0.656$, $d = 0.11$, 95%CI = [-0.33 0.21]) or the response phase ($t(15) = 0.61$, $p = 0.549$, $d = 0.15$, 95%CI = [-0.10 0.17]).

Infants did not produce significantly different difference scores for the trained word and the untrained word (0-1 s: $t(15) = 0.32$, $p = 0.753$, $d = 0.08$, 95%CI = [-0.62 0.83]; 0-2 s: $t(15) = 0.67$, $p = 0.513$, $d = 0.17$, 95%CI = [-0.50 0.26]; 0-3 s: $t(15) = 1.07$, $p = 0.302$, $d = 0.27$, 95%CI = [-0.46 0.15]; Figure 2d), or in comparisons with chance level (ts < 1.30, ps > 0.230, ds < 0.32) . Permutation-based t-tests did not reveal significant differences either (ps > 0.250). Thus, 14-month-old infants do not readily recruit the concept of chasee when interpreting a chasing scene (at least in the implementation of chasing that we adopted). We found no effect of order of presentation of words (ps > 0.250).

Discussion

The current finding suggests that it is probably the action to a well-defined goal that makes a chaser a good candidate for attaching a label to. Note that this result does not entail that infants do not have a concept of chasee. Perhaps a different chasing scene, in which the chasee adjusts its behavior to that of the chaser, making the event a real inter-action, would successfully facilitate infants' linking a new word to the available concept of chasee.