Twelve-month-olds disambiguate new words using mutual-exclusivity inferences

Barbara Pomiechowska [1], Gábor Bródy [1,2], Gergely Csibra [1,3], Teodora Gliga [4]

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

[2] Department of Cognitive, Linguistic, and Psychological Sciences, Brown University

[3] Birkbeck College, University of London

[4] School of Psychology, University of East Anglia

Corresponding author:
Barbara Pomiechowska
PomiechowskaB@ceu.edu
Department of Cognitive Science
Central European University
Nador utca 9
1051 Budapest
Hungary

Abstract

Representing objects in terms of their kinds enables inferences based on the long-term knowledge made available through kind concepts. For example, children readily use lexical knowledge linked to familiar kind concepts to disambiguate new words (e.g., “find the toma”): they exclude members of familiar kinds falling under familiar kind labels (e.g., a ball) as potential referents and link new labels to available unfamiliar objects (e.g., a funnel), a phenomenon dubbed as ‘mutual exclusivity’. Younger infants’ failure in mutual exclusivity tasks has been commonly interpreted as a limitation of early word-learning or inferential abilities. Here, we investigated an alternative explanation, according to which infants do not spontaneously represent familiar objects under kind concepts, hence lacking access to the information necessary for rejecting them as referents of novel labels. Building on findings about conceptual development and communication, we hypothesized that nonverbal communication could prompt infants to set up kind-based representations which, in turn, would promote mutual exclusivity inferences. This hypothesis was tested in a looking-while-listening task involving novel word disambiguation. Twelve-month-olds saw pairs of objects, one familiar and one unfamiliar, and heard familiar kind labels or novel words. Across two experiments providing a cross-lab replication in two different languages, infants successfully disambiguated novel words when the familiar object had been pointed at before labeling, but not when it had been highlighted in a non-communicative manner (Experiment 1) or not highlighted at all (Experiment 2). Nonverbal communication induced infants to recruit kind-based representations of familiar objects that they failed to recruit in its absence and that, once activated, supported mutual-exclusivity inferences. Developmental changes in children’s appreciation of communicative contexts may modulate the expression of early inferential and word learning competences.

Keywords

infancy, conceptual development, word learning, mutual exclusivity, pointing, inference, object representation, eye tracking

1. Introduction

Although there are multiple ways to represent objects, conceptual representation in terms of object kinds (e.g., whether an object is a cherry or a pebble) is a matter of routine for the human mind and carries important inferential benefits. Providing access to long-term knowledge (e.g., a cherry, unlike a pebble, is edible and tasty; in English, it is called "cherry" while in French it is called "cerise"), conceptual representation is a source of premises for inferences guiding behaviour and communication. Here, we investigated the relationship between communication, kind-based object representation and inferential processes in early development. Building on evidence from studies on conceptual and lexical development, we hypothesized that young infants do not spontaneously deploy their kind concepts to represent perceived objects, which, in turn, prevents them from easily accessing the associated kind labels and limits the range of inferences they can make to interpret verbal communication. We tested this proposal using an adaptation of a verbal reference disambiguation task, which produces the so-called mutual-exclusivity effect in older children and adults.

1.1. The Role of Communication in Object Representation

While infants as young as 6 to 9 months of age readily encode their visual experience through core-cognition concepts (Carey, 2009), such as human/agent versus object (Kibbe & Leslie, 2019; Bonatti et al., 2002; Surian & Caldi, 2010), basic-level kind concepts (or sortals, Xu, 2007), such as ball or car, are not adopted for object representation before about one year of age. For example, object individuation experiments found that upon seeing a ball and a car emerging in an alternating manner from behind an opaque screen, a situation in which kind-based representation enables inferring the presence of two hidden objects, 10-month-olds were unable to decide how many objects were there, while 12-month-olds succeeded (e.g., Xu & Carey, 1996; Van de Walle et al., 2000). Even more strikingly, younger infants also failed to draw on kind concepts to make sense of a visual scene when provided with uninterrupted perceptual access to the target objects (Xu et al., 1999). One interpretation of these difficulties in kind-based individuation is that basic-level kind concepts and object categories they pick out emerge late during the first year of life, and their acquisition might depend on linguistic input (Xu, 2002; Carey, 2009; see also Perszyk & Waxman, 2018). However, evidence from studies using labelling requires considering an alternative interpretation: namely, infants might be struggling not with the lack of kind concepts, but the recruitment of these concepts for object representation.

By 9 months of age, infants readily set up object representations that include kind information in response to familiar kind labels. Simply naming the objects emerging in turns from behind an occluder in a typical object-individuation scenario (e.g., "Look, [baby's name], a **duck**" upon bringing out object 1, and "Look, [baby's name], a **ball**" upon bringing out object 2) enables 9-month-olds to use the concepts indexed by the familiar labels (DUCK, BALL) to successfully establish the presence of two occluded objects (Xu, 2002, for a review, Xu, 2007). Furthermore, upon hearing a familiar label that does not match the subsequently presented object (e.g., when someone points to a screen announcing that it is hiding a shoe, while a duck is revealed as the screen goes down), infants at this age, like adults, produce a neural response considered to index semantic violation (i.e., N400 even-related potential, Parise & Csibra, 2012). Critically, when the order of presentation is swapped such that objects, instead of words, are used as primes, the semantic violation effect is not present before 14 months of age (Friedrich & Friederici, 2005a, b; cf. Friedrich et al., 2017). Together, these findings suggest that merely seeing familiar objects does not automatically activate kind representations and the lexical information associated with them even though they are part of infants' conceptual repertoire and can be elicited through labelling. This begs the question of why young infants do not spontaneously make use of kind knowledge in familiar object representation.

At least two, perhaps complementary, answers to this question emerge. First, activating a kind representation based on visual evidence might be more difficult than based on a label. Conceptual representations, including kind representations, are assumed to be information structures (Barrett, 2015), indexed by internal mental symbols (Sperber & Wilson, 1986), encompassing lexical information (i.e., labels used to express concepts in natural language) and non-linguistic information (e.g., perceptual features and feature-combinations characterizing the entities a given concept denotes). As each kind label unambiguously points to the kind symbol it is connected to, naming an object directly activates a kind-based representation. In contrast, the visual data generated by looking at an object only partly overlap with the features corresponding to a particular object kind stored in the long-term memory (Murphy, 2004) because every exemplar is different; this partial overlap may be insufficient for young infants to activate the internal kind symbols and, hence, represent a viewed object as an instance of a familiar kind. Second, thinking of objects in terms of their kinds is necessary and beneficial only if knowledge associated with kind concepts is useful for the cognitive tasks at hand. Therefore, kind-based representation of a perceived entity, although achievable, might not be used as the default. Instead, it might be conditional on the context and one’s current goals (even in adults, e.g., accessing any conceptual representation beyond the concept of object may be irrelevant if all we want is to avoid collision).

One task that cannot be accomplished without conceptual object representation is the interpretation of communication. Human referential communication works by invoking meaning in the form of conceptual representations (Pinker, 2007; Macnamara, 1982; see also Bloom, 2000). In verbal communication, referents are typically introduced under a linguistic description marking a specific conceptual perspective determined by the purpose of the communication (Clark, 1997; Geeraerts et al., 2012). For example, referring to an animal as “the dog” underlines the basic-level kind to which it belongs, while referring to it as “Sebastian, my new puppy” emphasizes its individuality and provides individual-specific information. As discussed above, the ability to adopt particular conceptual perspectives while interpreting verbal descriptions is manifest by 9 months of age, as infant respond to familiar kind labels by setting up kind-based representations of referred objects. In addition, they take novel words to refer to object kinds and categories as indicated in a variety of experimental tasks (e.g., Dewar & Xu, 2009; Balaban & Waxman 1997; Ferry et al. 2010, 2013; Yin & Csibra, 2015; Pomiechowska & Gliga, 2019) and, around their first birthday, they recognize that novel words might signal other conceptual perspectives besides object kinds (e.g., individuals, LaTourrette & Waxman, 2020; Pickron et al., 2018; action roles, Yin & Csibra, 2015; properties, Waxman & Booth, 2003).

Although language is arguably the most convenient tool to express different concepts under which a referent should be considered, conceptual perspective is not always made explicit verbally. In particular, in nonverbal communication, when no linguistic description is available, in order to interpret the communicated message, the addressee has to supply a description of the referent on their own. The right level of representation can be worked out by recruiting the concept that is most accessible or most relevant given the context (Sperber & Wilson, 1986; see also Tomasello, 2008). A wealth of experimental evidence indicates that, for familiar objects, adults typically set up kind-based conceptual representations, that by definition, correspond to basic-level object categories (e.g., cherry, dog, computer; Rosch et al., 1976; Rosch, 1978; Tanaka & Taylor, 1991; Lin et al., 1997).

Although to date there has been no systematic investigation of whether nonverbal communication or, more widely, communication that does not involve labelling triggers some level of conceptual encoding of referred objects also in early development, initial evidence suggests that it induces infants to set up kind-based representations of familiar objects. For example, introducing an object in a communicative manner but without labelling it results in a representation rich in information about its category features (Xu, Carey, & Quint, 2004; Shamsudheen, 2020; Ferry et al., 2013; Ferguson & Waxman, 2016; Thiele et al., under review) and/or functions (Futó et al., 2010), which would not necessarily be encoded in a non-communicative context. Furthermore, a closer look at the methods of studies reporting successful kind-based object individuation in one-year-olds reveals that, while refraining from labelling objects, paradigms systematically involved ostensive communication and object-directed referential actions. For example, in the tasks described above (Xu et al., 1999; Xu, Carey, & Quint, 2004), infants were ostensively addressed (e.g., "Look at this, *INFANT'S NAME*. Now.") and the objects were highlighted via tapping them against the table, an action that 12-month-olds have likely interpreted as a deictic referential signal (Pomiechowska & Csibra, under review). It is thus possible that infants' success in using kind-based concepts for object representation documented in the literature has been facilitated by communication about the presented items, which triggered the retrieval of the available conceptual knowledge.

Building on these theoretical and methodological observations, we hypothesized that infants' use of conceptual representation might not be spontaneous but contingent on the cognitive tasks they engage in, such as interpretation of communicative acts directed at them. To examine this hypothesis in the present study, we investigated whether the nature of representations that infants recruit for objects depends on whether these objects are introduced in a communicative manner or not. More specifically, we compared object representations that 12-month-olds set up for familiar objects in the following situations: (1) when the target object is highlighted by a nonverbal communicative action in the form of pointing (Experiments 1-2), (2) when the target object is highlighted non-communicatively by moving vertically on the screen (Experiment 1), and, finally, (3) when the target object is not highlighted at all (Experiment 2). If infants are compelled to retrieve a conceptual description of a familiar object targeted by a nonverbal communicative action, the information that comes with a conceptual representation (e.g., label, function, properties) should become available to them and could support inferential processes that rely on this information. Here, we focused on inferences involved in disambiguating novel words: we reasoned that young infants might fail to apply mutual exclusivity to interpret these words simply because the representational formats they spontaneously use for familiar objects do not provide any links to lexical information.

Arguably, our idea rests on the assumption that, when prompted by nonverbal referential communication, infants would categorize familiar objects in terms of their kinds. Although any concept could be recruited to make sense of nonverbal communication (e.g., kind, individual, property, action), those that are most available and easiest to access should be prioritized (Bloom, 2000; Horton & Markman, 1980). The experimental literature indicates that infants prioritize kind concepts over other representations when interpreting verbal communication about familiar objects (Dewar & Xu, 2009; Pomiechowska & Gliga, 2019; Xu et al., 1999; Xu, Carey, & Quint, 2004; Futo et al., 2010; see also Yin & Csibra, 2015). Importantly, this interpretation might stem from the structure of the speech input infants receive. Parents commonly talk about object kinds and employ kind labels (e.g., “dog”) more frequently than labels describing other levels of categorization (e.g., “mammal” or “Golden retriever”, Brown, 1958; Callanan, 1985). We thus reasoned that infants should favour kind-based representations also when interpreting nonverbal communication.

1.2. Conceptual Representation at the Service of Word Disambiguation

When presented with two objects, one belonging to a familiar kind (e.g., dog) and the other one to an unfamiliar kind (e.g., phototube), and requested to "find the moxi", monolingual adults and young children interpret this expression as referring to the unfamiliar object (Markman & Watchel, 1988; Merriman et al., 1989; Mervis & Bertrand, 1994; Bion et al., 2013; Halberda, 2006; Diesendruck & Markson, 2001; Houston-Price et al., 2010; Golinkoff et al., 1992; 1996; Spiegel & Halberda, 2011; Markman, Wasow, & Hanson, 2003; for bilinguals see, e.g., Byers-Heinlein & Werker, 2009, 2013; Byers-Heinlein, Chen & Xu, 2014; Houston-Price et al., 2010; Au & Glusman, 1990; Davidson & Tell, 2005; for a meta-analysis, see Lewis et al., 2020). This interpretation rests on recognizing the familiar object as a member of the kind dog and using the label associated with this concept ("dog") to discard this object as a potential referent of the novel word ("moxi"). Importantly, infants below 16-18 months of age, even when they know the name of the familiar object, seem unable to decide what the novel word refers to, thus failing in mutual exclusivity tasks. This apparent developmental shift in performance was consistently reported in a number of studies (Halberda, 2003; Bergelson & Aslin, 2017; Garrison et al., 2020; cf. Mather & Plunkett, 2010a).

Various explanations for disambiguation through mutual exclusivity have been advanced. Some proposed that children are equipped with specific lexical principles that guide their interpretation of new words (e.g., mutual exclusivity principle, Markman & Watchel, 1989; Markman, 1990; novel-name nameless-category, Golinkoff, Mervis, & Hirsh-Pasek, 1994; Mervis & Bertrand, 1994). Others suggested that their behaviour is underpinned by pragmatic factors (Clark, 1988, 1990; Diesendruk & Markson, 2001; Bloom, 2000) or development of word-learning heuristics based on experience with lexical regularities in one's linguistic environment (Byers-Heinlein & Werker, 2009, 2013; Lewis et al., 2020; see also, Lewis & Frank, 2013; Frank, Goodman, & Tenenbaum, 2009). Regardless of what motivates children to avoid lexical and/or referential overlap, passing mutual exclusivity tasks requires applying a sequence of inferential rules (e.g., disjunctive syllogism) involving the following inferential steps (Halberda, 2003, 2006): (1) “moxi” describes object A or “moxi” describes object B (A *OR* B); (2) “find the moxi” does not refer to object A because this object falls under the label “dog” (*NOT* A); (3) “find the moxi” must refer to the object B (*THERFORE* B). Hence, it was previously assumed that young infants fail to show the mutual exclusivity effects because they lack some part of the necessary inferential apparatus, do not have access particular principles regarding word meaning, or have not had time to develop necessary word-learning heuristics.

However, disambiguation through mutual exclusivity may also fail if infants do not automatically represent familiar objects under familiar kind concepts and associated kind labels (a prerequisite for step 2, in the disjunctive syllogism described above). Without setting up object representations that mediate lexical knowledge, there is no way to contrast the newly encountered novel word with the familiar object’s name stored in one’s long-term knowledge. In other words, infants might be equipped with the computational tools and knowledge to disambiguate novel words, but simply stand no chance of applying them because they do not represent the familiar objects in the appropriate format that would support access to lexical information. If it is then the lack of spontaneous categorization under kind concepts that prevents them from applying mutual exclusivity, prompting infants to deploy conceptual representation should enable them to disambiguate novel words.

1.3. The Present Experiments

In the present experiments we investigated whether nonverbal communication makes infants deploy kind concepts to represent familiar objects. More specifically, we tested the hypothesis that in 12-month-olds kind-based object representation can be triggered through nonverbal communication about familiar objects, and this should be reflected in successful novel word disambiguation. We used an adaptation of a standard mutual-exclusivity looking-while-listening task modelled on that used by Halberda (2003, 2006), but we additionally introduced a manipulation aimed at triggering conceptual object representation before labelling. Like in the original task, on each trial, infants were first presented with two objects, an exemplar of a familiar kind whose label they should know (e.g., a ball) and an unfamiliar item (e.g., a paper puncher). Then, unlike in the original task, on some trials the familiar object was targeted by a brief nonverbal communicative action in the form of pointing aimed at triggering categorization / conceptual representation (experimental condition). On other trials (control conditions) the familiar object moved repeatedly along vertical axis (Experiment 1) or was not highlighted at all, as in original mutual exclusivity tasks (Experiment 2). Following this, infants were either prompted with a phrase containing a familiar word corresponding to the familiar object (e.g., “Where is the ball?”) or an unfamiliar word (e.g., “Where is the moxi?”). Their looking behaviour was recorded to assess referent selection. Both manipulations (highlighting before labeling and word choice during labeling) were applied within subjects (N = 20 per experiment). If (i) conceptual representation of the familiar object is necessary to disambiguate novel words, and (ii) infants do not spontaneously encode familiar objects in terms of their kinds but (iii) nonverbal communication triggers such encoding, infants should display successful disambiguation through mutual exclusivity in the experimental conditions only.

We chose to test 12-month-olds for three main reasons. First, they were previously shown to appreciate nonverbal communication (e.g., Senju & Csibra, 2008; Hernik & Broesch, 2019; Tauzin & Gergely, 2018, 2019). Second, their conceptual repertoire contains a handful of lexicalized basic-level object kind concepts (Bergelson & Aslin, 2017;

Bergelson & Swingley, 2012; Parise & Csibra, 2012). Last, there is initial evidence that 12-month-old infants can apply deductive inference involving exclusion in nonverbal contexts (Cesana-Arlotti et al., 2018, 2020; Halberda, 2018; cf. Leahy & Carey, 2020).

Nonverbal communication was operationalized as pointing based on the evidence from interpretation and production studies indicating that by 12 months of age infants both recognize as well as use pointing as a communicative action expressing deictic reference. While several studies confirm that infants expect pointing gestures produced by others to be directed at specific objects or locations, whether points are accompanied by gaze and postural information (i.e., the actor/experimenter turns his head in the direction of the point, Gliga & Csibra, 2009; Behne et al., 2012) or not (i.e., only the pointing hand is visible, Daum et al., 2013; Pomiechowska & Csibra, under review), the strongest evidence that pointing is understood as a referential act comes from studies demonstrating that 12-month-olds expect points to co-refer with concurrently produced labels. Gliga and Csibra (2009) showed infants videos of an actress pointing to the location of a hidden familiar object while naming it. Infants looked longer when the named object was revealed not at the location indicated by the pointing gestures, but on the opposite side of the display. Further experimental evidence demonstrated that when one of two unfamiliar objects is pointed at during labeling, infants take it to be the referent of the novel label, an interpretation they do not deploy when the object is targeted by an instrumental action such as grasping (Pomiechowska & Csibra, under review). In production, 12-month-olds not only deploy pointing to selectively indicate objects they want to learn about (Kovacs et al., 2014) or inform about (Liszkowski et al., 2004, 2007), but also to refer to kinds of absent objects (e.g., infants request more of the items these containers previously carried by pointing to empty containers, Bohn et al., 2015, 2018).

2. Experiment 1

We sought to find out whether infants' referent selection in response to familiar and novel words delivered individually in the presence of two objects (one familiar, one unfamiliar) would be affected by the way they represent the familiar object prior to labelling. To this aim, we highlighted the familiar object prior to labelling in one of two ways. In the experimental condition, highlighting consisted in pointing towards the familiar object, a nonverbal communicative action that was expected to trigger conceptual representation of the target. In the control condition, highlighting consisted in the object moving up and down, which was expected to attract infants' attention to the object the same way as pointing but without eliciting conceptual encoding.

If infants at this age spontaneously draw on kind concepts to represent familiar objects and their difficulties at disambiguating new words are due to the immaturity of their inferential apparatus or lack of word-learning skills, their performance should be comparable across highlighting conditions: they should succeed in recognizing the familiar words and fail to disambiguate the novel ones. If, on the other hand, early novel word disambiguation performance is limited by the absence of spontaneous conceptual encoding of familiar objects, which, however, can be triggered by nonverbal communication about an object (but not by its movement), infants should be able to disambiguate novel words only in the experimental condition.

2.1 Method

2.1.1. Participants

Twenty 12-month-olds (M = 12 months 3 days, range: 11 months 15 days to 12 months 27 days) from monolingual English-speaking families were included in the analysis.

All infants were born full-term and healthy. An additional 10 infants were tested but had to be excluded from the analysis, having failed to complete the experiment (n = 6) or provide enough data (n = 4, for inclusion criteria see below). Infants were recruited through advertising in the local press. All caregivers provided written informed consent. Participation was rewarded with a small gift. Our sample size was selected a priori based on the review of word-recognition literature in one-year-olds (e.g., Bergelson & Aslin, 2017; Pomiechowska & Gliga, 2019). The study was approved by the local ethical committee.

2.1.2. Apparatus

We collected infants' binocular gaze data using a TOBII TX300 eye tracker with an integrated 23-inch monitor (resolution 1920 x 1080 px, refresh rate: 60 Hz, Tobii Technology, Danderyd, Sweden: *https://www.tobii.com/).* The data were sampled at 120 Hz. External speakers delivered the sound. The stimuli presentation and data collection were administered through MATLAB, with Psychophysics Toolbox for stimuli presentation (Brainard, 1997) and the Tobii Pro Analytics for data collection (https://www.tobiipro.com/productlisting/tobii-pro-sdk/).

2.1.3. Stimuli

Object stimuli were colour photographs representing 16 kinds of objects. Eight of them were selected to be familiar to the infants (ball, banana, car, cup, shoe, spoon, teddy, telephone) and the other 8 to be unfamiliar (cheese grater, flashlight, garlic press, hourglass, massage tool, padlock, paper puncher, shell) based on the previous experimental research (Parise & Csibra, 2012; Bergelson & Swingley, 2012) and parental reports (Hamilton, Plunkett, & Schafer, 2000). Two different photographs per object kind were used. Their display size was 375 x 375 px, subtending approximately 9° of visual angle during presentation. The depicted objects were matched in height, and whenever possible also in width.

Word stimuli were eight familiar words in the form of count nouns corresponding to the familiar object kinds (i.e., “ball”, “banana”, “car”, “cup”, “shoe”, “spoon”, “teddy”, “telephone”) and four pseudo-words conform to English phonotactics (“dax”, “moxi”, “peko”, “wug”). The pseudo-words were selected not to overlap in their onset phonemes with any the familiar words. All word stimuli were embedded in the following carrier phrase: *Hi baby! Look! Where is the [LABEL]? [LABEL]! [LABEL]!* The speech was recorded by a female native speaker of British English using infant-directed prosody. There was a single token of each phrase, ensuring that infants looking responses would not be affected by auditory differences between labelling phrases.

All stimuli and sample animations are available at https://osf.io/rn6zv/.

2.1.4. Design

We used a two-by-two within-subject factorial design. First, we manipulated the way in which the familiar object was highlighted before labelling. In the experimental condition, infants’ attention to the familiar object was called by pointing, while in the control condition by the object’s small vertical movement. Second, we varied whether the word uttered after highlighting was the basic-level name of the familiar object or a novel label. The experiment consisted of 16 trials (4 per each crossing of factors) delivered in two blocks of 8 trials blocked by highlighting condition, with half of the infants watching the experimental condition first, and the other half watching the control condition first. The blocks were presented with a short break in between (30 s), during which infants were shown an animation of geometric shapes moving to the sound of soft music. This break was introduced to help maintain infants’ interest in the task.

2.1.4.1. Experimental Condition. On each trial, infants saw two objects placed on the opposite sides of a white table (Figure 1): one belonged to a familiar kind and the other was an unfamiliar object. At the beginning of the trial, the objects appeared simultaneously and remained still in silence for 2 s. This period of time was used as a within-trial baseline phase to measure infants' spontaneous image preference. Then, a downward pointing hand appeared above the familiar object. It moved down towards the object and up away from it three times, without entering in contact with the object. This was accompanied by a jingle

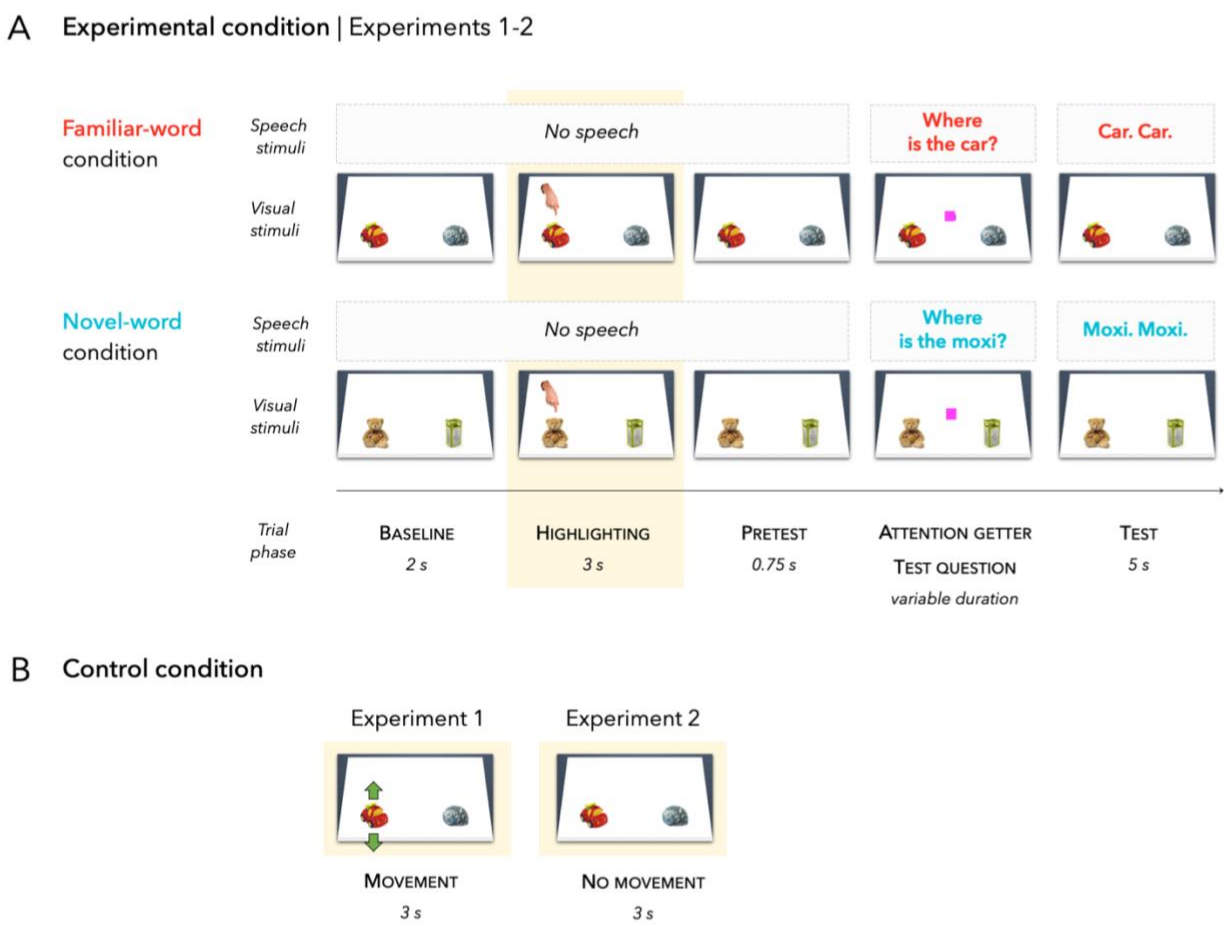

Figure 1. Experimental design across Experiments 1-2. The trial structure and timing were kept constant throughout the study. We used a two-by-two design, with two highlighting conditions (experimental v. control) and two labeling conditions (familiar v. novel word) in each experiment. (**A**) *Experimental conditions* involved pointing to the familiar object and were identical across experiments (except for language: English in Experiment 1; Hungarian in Experiment 2). The schematic represents trial structure across familiar- and novel-word trials. (**B**) *Control condition*s differed between experiments: in the control condition of Experiment 1, the familiar object moved vertically; in the control condition of Experiment 2, both objects remained still.

sound repeated two times (for a total duration of 0.9 s). This highlighting phase lasted for a total duration of 3 s. After the pointing hand disappeared, the objects were presented in silence for 0.75 s before infants' attention was brought to the middle of the display by a dynamic attention getter (a pulsating square expanding and shrinking between 50 x 50 px and

150 x 150 px, i.e., between 1.00° and 2.35° visual angle). This was done to ensure that infants were at an equal distance from the two objects at the onset of the labelling phase. The attention getter was gaze contingent, such that looking at it continuously for 0.5 s would initiate the onset of the labelling phase. The disappearance of the attention getter was synchronized to the offset of the label and marked the beginning of the test phase. The test

Trial	Block	Highlighting condition	Word condition	Word stimulus	Side of the target object	Object stimuli
1	1	Experimental	Familiar	*banana*	left	
2	1	Experimental	Novel	*moxi*	left	
3	1	Experimental	Novel	*dax*	right	
4	1	Experimental	Familiar	*shoe*	right	
5	1	Experimental	Novel	*dax*	left	
6	1	Experimental	Familiar	*shoe*	left	
7	1	Experimental	Familiar	*banana*	right	
8	1	Experimental	Novel	*moxi*	right	
Attention-getter video (30 s)						
9	2	Control	Novel	*peko*	left	
10	2	Control	Familiar	*telephone*	left	
11	2	Control	Familiar	*teddy*	right	
12	2	Control	Novel	*wug*	right	
13	2	Control	Familiar	*teddy*	left	
14	2	Control	Novel	*wug*	left	
15	2	Control	Novel	*peko*	right	
16	2	Control	Familiar	*telephone*	right	

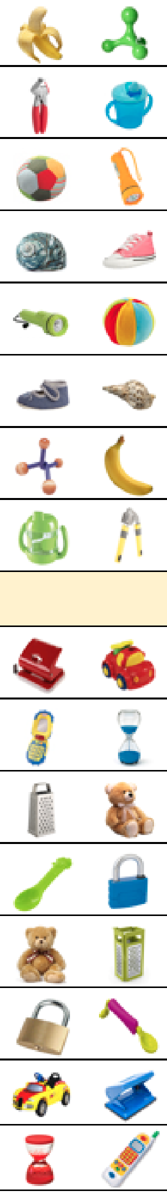

Figure 2. Sample experimental list (see section 2.1.4.3 *Counterbalancing).*

phase lasted for 5 s, during which the isolated label was repeated two more times (starting at 1.5 s and 3.5 s relative to the onset of the test phase).

2.1.4.2. Control Condition. The trial structure and timing were identical to the experimental condition. The only difference between conditions was the way in which the familiar object was highlighted before labelling: instead of being pointed at, it moved up and down three times.

2.1.4.3 Counterbalancing. For each participant, the novel and familiar objects were randomly paired into 8 pairs (Figure 2). Four pairs were then assigned to the experimental condition and the other four to the control condition. Within each highlighting condition (experimental v. control), each pair of objects was assigned a unique target word (either familiar or novel) and this pair was presented twice, with different exemplars of the corresponding categories in each presentation. For example, the labelling phrase "where is the banana?" was used in two trials within the assigned highlighting condition, and each time it was accompanied by different images of a banana and the corresponding novel object[1]. This resulted in 8 trials per highlighting condition: 4 using familiar words and 4 using novel words. The object locations were swapped from one presentation to another.

The order of trial presentation within the block was randomized with the following constraints: (1) there should not be more than two trials from the same word condition (familiar v. novel) in a row, (2) the target object should not be displayed more than two times

[1] The target word/object pair assignments were counterbalanced across but not within participants to avoid the following confound: had we used within participants counterbalancing, for some object pairs the familiar-label trial (e.g., "Where is the banana"?) would come before the novel-label trial (e.g., "Where is the moxi"?). It would be then impossible to determine the origins of infants' success at applying mutual exclusivity because we would not be able to tell apart whether (1) they retrieved the familiar label on their own after having recruited kind-based representation of the familiar object or (2) simply memorized it from the previous familiar-label trial.

on the same side, and (3) the same target word should not be presented on two consecutive trials.

2.1.5. Procedure

Infants were tested individually in a soundproof and dimly lit eye-tracking lab, sitting in a car seat approximately 60 cm from the eye-tracking camera and the monitor. Caregivers sat behind them and were instructed to remain silent and refrain from interacting with the child. Each session started by a 5-point calibration routine, performed until 4 out 5 points were successfully calibrated. The task started immediately after the calibration. The testing, including the calibration routine, lasted approximately 7 minutes.

2.1.6. Data Processing and Analysis

The analysis was performed only on valid trials, defined as trials during which infants provided at least 50% of on-screen data during baseline, highlighting, labelling, and test phases. On average, infants contributed 12.6 trials (*SD* = 2.59, range: 8 to 16; experimental condition with familiar words: *M* = 3.25 trials, *SD* = 1.02; experimental condition with novel words: *M* = 3.20 trials, *SD* = 0.89; control condition with familiar words: *M* = 2.95 trials, *SD* = 0.99; control condition with novel words: *M* = 3.20 trials, *SD* = 0.95).

To quantify infants' looking behaviour, we defined two square-shaped areas of interest (AOIs), one centred around the familiar object and the other one around the unfamiliar object. The AOIs' surface was slightly larger (505 x 505 px) than the surface of the displayed pictures (375 x 375 px) to accommodate sparse gaze points. These AOIs included the surface covered by the moving object (in the control condition), but did not include the surface covered by the hand movement (in the experimental condition).

2.1.6.1. Main Analysis. The aim of our main analysis was to assess whether the infants' looking patterns at test were affected by their interpretation of labels. Our main outcome measure was a baseline-corrected proportion of looking at the familiar object (*corrPLF*), which reflects how infants' looking behaviour at test was affected by naming relative to a silent baseline phase within each trial. This measure adjusts for preferences that infants might have for individual picture stimuli. For each trial, we first calculated the proportion of looking at the familiar object during the baseline and test phases. This was done by dividing the sum of AOI hits recorded in the familiar object AOI by the total AOI hits recorded in the familiar or unfamiliar objects' AOIs: $PLF = AOI\_HITS_{FAMILIAR\_OBJECT} / (AOI\_HITS_{FAMILIAR\_OBJECT} + AOI\_HITS_{UNFAMILIAR\_OBJECT})$. We then subtracted the PLF at baseline from the PLF at test: $corrPLF = PLF_{TEST} – PLF_{BASELINE}$. Such derived *corrPLF* values range from -1 to 1. Positive *corrPLF* values indicate an increase in looking to the familiar object relative to baseline, which corresponds to the decrease in looking at the unfamiliar object. Negative *corrPLF* values indicate a decrease in looking to the familiar object, which corresponds to the increase in looking to the unfamiliar object. Therefore, positive *corrPLF* scores suggest that infants recognize the familiar words, and negative *corrPLF* scores suggest that they link novel words to the unfamiliar objects. The data from the valid trials were averaged across trials by highlighting (experimental v. control) and word (familiar v. novel) conditions.

While our main analysis was performed on the fixation data summed over the duration of the test phase (Figure 3), to illustrate the time course of infants' responses we plotted also the evolution of the *corrPLF* across time (Figure 5). The time course data were derived within-trial by subtracting the average $PLF_{BASELINE}$ from PLF_{TEST} at each sample (i.e., time point) of the test phase. Then, we computed individual averages by averaging over trials within each condition separately for each participant, and finally we computed grand averages for each condition by averaging over participants.

2.1.6.2. Additional Analyses. An additional analysis was performed to investigate how infants' attention was distributed during the highlighting phase and, more specifically, whether the familiar objects were attended equally when being pointed at and when moving vertically on their own. To this aim we computed and compared the proportion of looking at the familiar object during the highlighting phase ($PLF_{HIGHLIGHTING}$) across conditions (experimental v. control).

2.2. Results

A preliminary mixed-model ANOVA on average *corrPLF* data with block order (experimental condition in block 1 v. in block 2) as a between-subject factor, as well as condition (experimental v. control) and word (familiar v. novel) as within-subject factors yielded no significant effect of block order nor significant interactions with this factor, all *p*s > 0.35. Thus, block order was removed from further analyses.

A repeated-measures ANOVA with condition (experimental v. control) and word (familiar v. novel) as within-subject factors revealed a significant interaction between word and condition, $F(1,19) = 7.098$, $p = .015$, $\eta_p^2 = .27$ (Figure 3). To resolve this interaction, we compared the responses to familiar and novel words separately within each condition using paired-samples *t* tests (two-tailed). In the experimental condition, familiar words ($M = .14$, $SD = .26$) elicited a significantly different looking pattern than novel words ($M = -.11$, $SD = .21$), $t(19) = 2.841$, $p = .010$, $d = 0.64$, 95% CI = [.07, .43], while in the control condition they did not, $t(19) = 0.284$, $p = .780$, $d = 0.06$, 95% CI = [-.14, .10], (familiar words: $M = -.01$, $SD = .15$; novel words: $M = .01$, $SD = .21$). Furthermore, only in the experimental condition the responses differed significantly from baseline (assessed by two-tailed one-sample *t* tests against 0). This was true whether infants heard familiar words, $t(19) = 2.382$, $p = .028$, $d = 0.53$, 95% CI = [0.02, 0.26], in which case they tended to increase their looking towards the familiar object, or novel words, $t(19) = 2.294$, $p = .033$, $d = 0.51$, 95% CI = [-.01, -.21], which made them increase their looking to the novel object. In the control condition, infants' looking remained at the baseline level for both types of words, *p*s > .70. Figure 5 shows the time course of infants' *corrPLF* responses spilt by condition. It indicates that in the

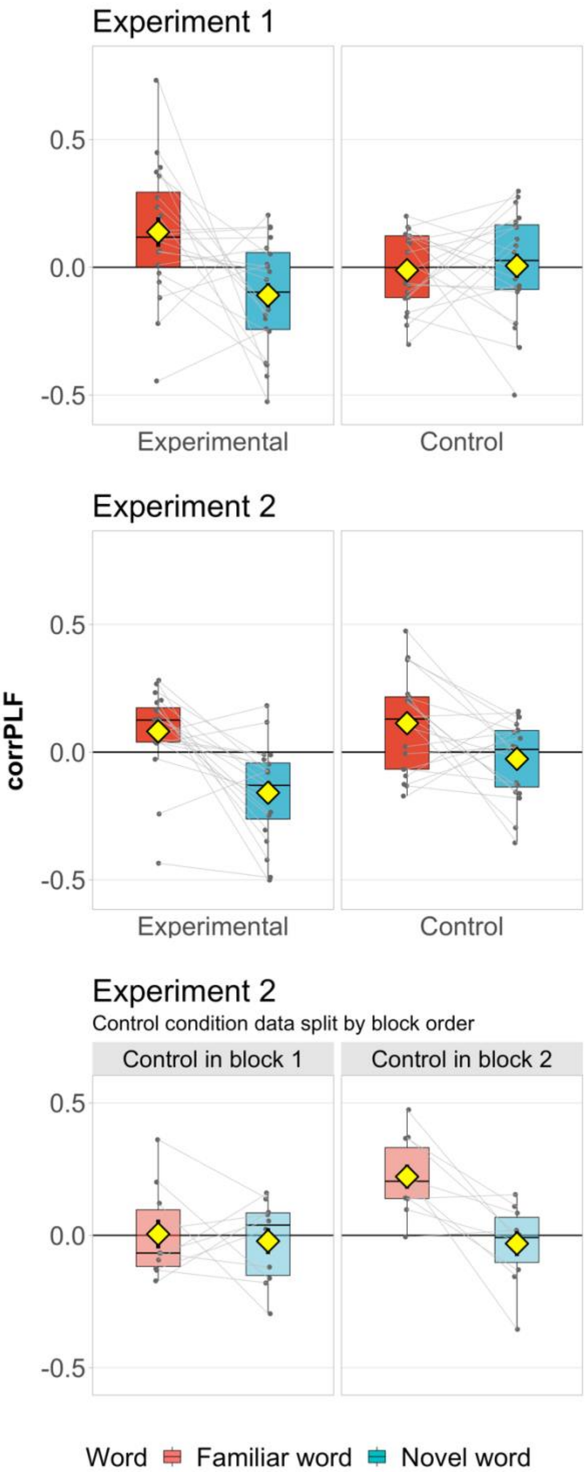

Figure 3. Results of Experiments 1 and 2: Baseline-corrected proportion of looking at the familiar object (*corrPLF*). The *corrPLF* indicates changes in looking to the familiar object during the test phase relative to the silent baseline phase. Positive values indicate an increase and negative values indicate a decrease in looking at the familiar object, which are equivalent, respectively, with a decrease and an increase in looking at the unfamiliar object. Yellow diamonds represent group means. Black horizontal lines indicate medians. The bottom and the top of the boxes represent the first and the third quartiles. Whiskers extend from the middle quartiles to the smallest and largest values within 1.5 times the interquartile range. Points connected across boxes represent individual means.

experimental condition their responses to familiar and novel words differed from each other and from baseline in a sustained manner through most of the test phase. No clear differentiation between words or either word and baseline can be observed in the control condition.

Additional analyses revealed that during the highlighting phase infants oriented preferentially to the familiar objects (Figure 4). Importantly, the pointing (*M* = .74, *SD* = .13), in the experimental condition, and the movement (*M* = .76, *SD* = .12), in the control

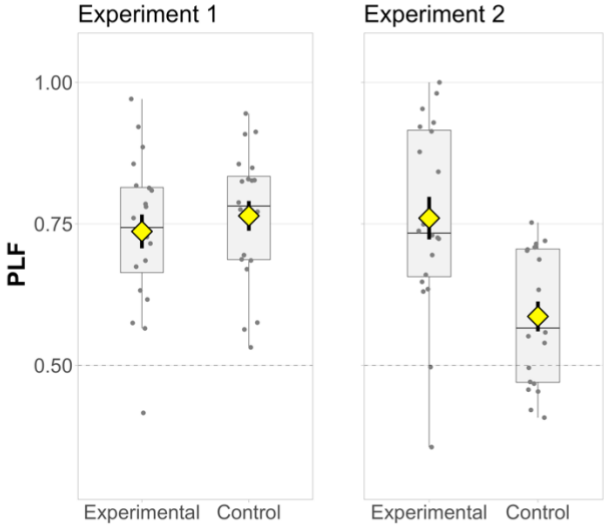

Figure 4. Proportion of looking at the familiar object (*PLF*) during the highlighting phase in Experiments 1 and 2**.** Yellow diamonds represent mean *PLF*. Bars represent the standard error of the mean. The bottom and the top of the boxes represent the first and the third quartiles. Whiskers extend from the middle quartiles to the smallest and largest values within 1.5 times the interquartile range. Dots represent individual averages.

condition, induced similar amount of looking at the highlighted object, $t(19)$ =.668, p = .512, d = 0.15, 95% CI = [-.06, .11]. In both conditions infants looked at the familiar object longer than expected by chance of (.5): $t(19)$ = 7.925, p < .001, d = 1.77, 95% CI = [.67, .80] in the experimental condition; and $t(19)$ = 10.035, p < .001, d = 2.24, 95% CI = [.71, .82] in the control condition. Their level of looking was not affected by block order, as revealed by a mixed-model ANOVA with block order as a between-subject factor and condition as a within-subject factor, all *p*s >.25.

2.3. Discussion

Conforming to our predictions, the current results indicate that the ways in which familiar objects were highlighted before labelling distinctly influenced infants' interpretation of speech. Only in the experimental condition, when labelling followed nonverbal communication about familiar objects, infants reacted differently to familiar and novel words. Relative to a silent baseline recorded before highlighting and labelling, at test, they increased their looking to familiar objects after hearing familiar words and, conversely, they increased their looking to novel objects after hearing novel words. This pattern of looking behaviour indicates that infants were able to interpret both familiar and novel words. More specifically, they identified familiar objects as the referents of expressions including familiar labels and rejected familiar objects as referents of expressions including novel labels. Therefore, these results provide evidence of word disambiguation via mutual exclusivity at 12 months of age, which is younger than had been assumed possible in the literature (Halberda, 2003; Bergelson & Aslin, 2017), and support the hypothesis that nonverbal communication promotes conceptual representation of the referent, making its content available to inferential processes. For example, once infants represent a familiar object using a kind concept, they may access the associated familiar label and then use it to disambiguate the novel words.

It is important to note that the pattern of results observed in the experimental condition cannot be explained by differences in the distribution of visual attention between conditions during highlighting: the amount of time infants spent fixating familiar objects in response to pointing in the experimental condition and to movement in the control condition was comparable. Explanations that would attribute infants' responses to novel words to mechanisms that match auditory to visual novelty (e.g., Horst et al., 2011; Mather & Plunkett, 2012; Mather, 2013, cf. Mather & Plunkett, 2010a) can also be ruled out. Heuristics to match novel words and unfamiliar objects would lead to increased orienting to unfamiliar objects (relative to the infants' baseline looking behaviour) irrespective of what happened before labelling. While such heuristics might contribute to disambiguation via mutual exclusivity in some experimental contexts (and mostly in older children, Horst et al., 2011; Mather & Plunkett, 2012), they cannot explain the current results. Since the age group tested here was younger than the children who were shown to use novelty-matching strategies, it is possible that such strategies require more language experience to develop (see also Lewis et al., 2020).

Interestingly, in the control condition, in which highlighting was operationalized through vertical movement of the familiar objects, infants failed to display evidence of either familiar word recognition or novel word disambiguation. We predicted that they would have difficulties interpreting novel words because non-communicative highlighting would not prompt them to represent familiar objects as members of particular kinds. However, we did not expect that they would fail to find the referents of familiar words. In line with the previous word-recognition studies, we reasoned that 12-month-olds should be able to match familiar objects with familiar kind labels, yet studies testing infants of this age used static objects. Although in older infants small movements of the labelled objects do not seem to affect familiar word recognition (Swingley & Aslin, 2002), it is possible that at the age tested here movement induces a conceptual description of the moving object that does not match the kind label used to test their word knowledge. One possibility is that infants identified the self-moving object as an agent, a concept that would not match the familiar kind label provided at test. In Experiment 2 we aimed to replicate our findings from the experimental condition of Experiment 1, while also asking whether it was the movement of the object that conflicted with the interpretation of familiar labels.

3. Experiment 2

In a new sample of participants, word-recognition and referent disambiguation performance were contrasted across two conditions: an experimental condition, in which familiar objects were targeted by nonverbal communication in the form of pointing that took place before labelling (as in Experiment 1), and a control condition in which there was no highlighting before labelling. If it was movement that disrupted familiar-word recognition in Experiment 1, infants in the present control condition should display evidence of recognizing these words (i.e., looking longer to familiar objects upon hearing familiar words) without succeeding in disambiguating the novel words (i.e., showing no preference for either the familiar or the unfamiliar objects upon hearing novel words). If nonverbal reference to a familiar object is a reliable trigger for kind-based representation, we should observe the same pattern of results in the present experimental condition as we found in the experimental condition of Experiment 1.

Experiment 2 was carried out in a different laboratory and in a different language: Hungarian.

3.1. Method

3.1.1. Participants

As in Experiment 1, our final sample included 20 12-month-olds (M = 11 months 28 days , range: 11 months 17 days to 12 months 11 days). Twelve infants were excluded from the analysis (n = 7 failed to provide enough data, n = 5 failed to complete the experiment). All infants were monolingual, living in Hungarian-speaking families who volunteered to participate in our research. Caregivers provided written informed consent. Infants' participation was rewarded with a small gift. The local ethical committee approved the research. The sample size was defined to match the one of Experiment 1.

3.1.2. Apparatus

The testing took place in a different lab, using a TOBII T60XL eye tracker (Tobii Technology, Danderyd, Sweden: https://www.tobii.com/) with 60 Hz sampling rate equipped with a 24-inch integrated monitor (resolution: 1920 x 1200 px, refresh rate: 60 Hz). The same methods of stimulus presentation and data collection were used as in Experiment 1.

3.1.3. Stimuli

We used the same visual stimuli as in Experiment 1, while our speech stimuli were adapted into Hungarian. Familiar kind labels were translated into Hungarian: “labda” (ball), “banán” (banana), “autó” (car), “pohár” (cup), “cipő” (shoe), “kanál” (spoon), “maci” (teddy), “telefon” (telephone). Novel labels were four pseudo-words selected to conform to Hungarian phonotactics (“dupi”, “gete”, “lim”, “rap”) and matched in structure with the pseudo-words from Experiment 1. As before, the words were embedded into a carrier phrase: “Szia baba, nézd csak! Hol van a [*LABEL*]? [*LABEL*]! [*LABEL*]!” (“Hi baby! Look! Where is the [*LABEL*]? [*LABEL*]! [*LABEL*]!”). The speech recording was performed by a female native speaker of Hungarian using infant-directed speech.

3.1.4. Design

The experimental condition was identical as in Experiment 1: familiar objects were targeted by nonverbal communication in the form of pointing before labeling. In the control condition, the familiar object was not highlighted in any way. Instead, after the baseline phase (2 s), both objects remained still for an additional 3 s to match the timing and structure between the control and experimental trials. The same jingle sound as used in the experimental condition (repeated twice for a total duration of 0.9 sec) was played (beginning 2 s counted from the trial onset).

3.1.5. Procedure and Data Analysis

The procedure and data analysis protocols were identical to the ones applied in Experiment 1. Overall, infants provided 14.10 trials (*SD* = 1.92, range: 8 to 16; experimental condition with familiar words: *M* = 3.40 trials; experimental condition with novel words: *M* = 3.35 trials; control condition with familiar words: *M* = 3.85 trials; control condition with novel words: *M* = 3.50 trials).

3.2. Results

Figure 3 shows the *corrPLF* data. These data were entered into a mixed-model ANOVA with block order (experimental condition in block 1 v. in block 2) as a between-subject factor, as well as condition (experimental v. control) and word (familiar v. novel) as within-subject factors. Unlike in Experiment 1, this analysis yielded significant main effects of block order, $F(1,18) = 12.654$, $p = .002$, $\eta_p^2 = .41$, and word, $F(1,18) = 48.172$, $p < .001$, $\eta_p^2 = .73$, and a significant three-way interaction, $F(1,18) = 4.568$, $p = .047$, $\eta_p^2 = .20$, indicating that infants' performance was affected by the order in which the highlighting conditions were delivered. To resolve this interaction, we performed separate ANOVAs within each highlighting condition. These analyses used word (familiar v. novel) as a within-subject factor and block order (experimental condition in block 1 v. in block 2) as a between-subject factor.

In the experimental condition, we observed only a significant main effect of word, $F(1,18) = 20.339$, $p < .001$, $\eta_p^2 = .53$, indicating that infants' looking behaviour reliably varied between familiar (*M* = .08, *SD* = .17) and novel words (*M* = -.16, *SD* = .19) and was not affected by block order. Comparisons to baseline on the data collapsed across block orders indicated above-baseline looking to familiar objects for familiar words, $t(19) = 2.152$, $p = .045$, $d = 0.48$, 95% CI = [0.01, 0.16], and below-baseline looking to familiar objects for novel words, $t(19) = 3.825$, $p = .001$, $d = 0.86$, 95% CI = [-.25, -.07]. These results replicate the one observed in Experiment 1, showing that, on one hand, infants could identify familiar objects as the referents of familiar words and, on the other hand, they could make sense of new words by relating them to novel objects.

In contrast, in the control condition there was a significant interaction between word and block order, $F(1,18) = 5.122$, $p = .036$, $\eta_p^2 = .22$, as well as significant main effects of word, $F(1,18) = 7.829$, $p = .012$, $\eta_p^2 = .30$, and block order, $F(1,18) = 4.562$, $p = .047$, $\eta_p^2 = .20$. The interaction was due to the fact that the infants' looking patterns differed depending on which block the control condition was administered in (first v. second). Namely, when the control condition was delivered in the first block, infants displayed comparable looking behaviour following familiar ($M = .01$, $SD = .17$) and novel ($M = -.03$, $SD = .15$) words, $t(9) = .350$, $p = .735$, $d = 0.11$, 95% CI = [-.14, .19], and neither differed from baseline, ps >.60. In contrast, when the control condition was delivered in the second block, looking behaviour was affected differently by the familiar ($M = .22$, $SD = .15$) and novel ($M = -.02$, $SD = .15$) words, $t(9) = 3.928$, $p = .004$, $d = 1.24$, 95% CI = [.11, .40]. In addition, looking to the familiar objects in response to the familiar words rose above baseline, $t(9) = 4.844$, $p < .001$, $d = 1.53$, 95% CI = [0.11, 0.32], while looking in response to novel words did not differ from baseline, $t(9) = .660$, $p = .526$, $d = 0.21$, 95% CI = [-.14, .08]. These order effects indicate that infants were successful at identifying referents of familiar words when these were not highlighted by movement, but only when the control block was preceded by an experimental block.

The time course of the average *corrPLF* is shown in Figure 5. In the experimental condition, infants' responses to familiar and novel words appear to differ from each other and from baseline in a sustained manner until the end of the test phase. In the control condition,

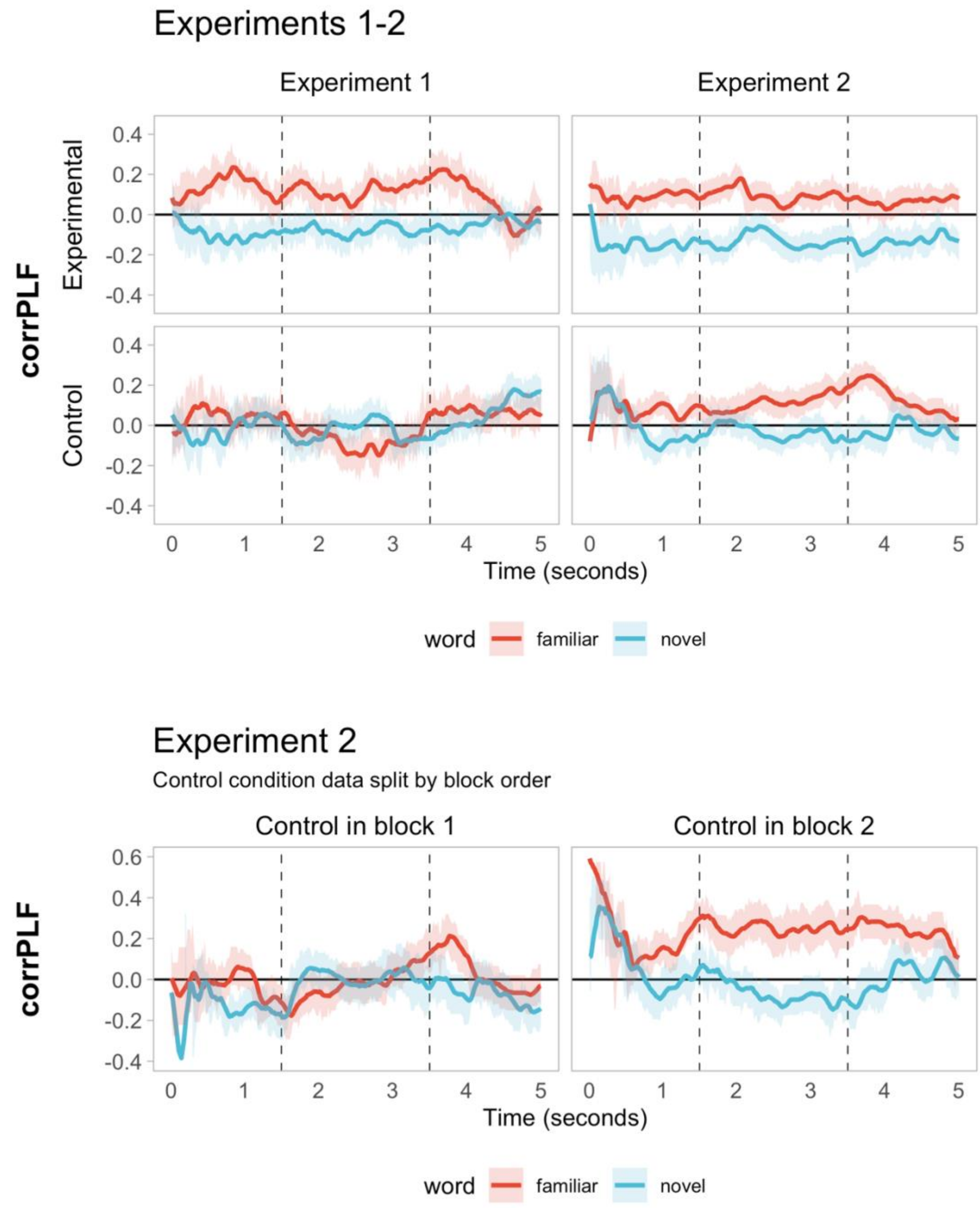

Figure 5. Time course of the average *corrPLF* responses. The ordinate shows the average *corrPLF* computed by averaging over subjects at each time point (see *Data Processing and Analysis)*. Error bars indicate SEMs. Zero corresponds to the offset of the labelling phrase (e.g., "*Where is the shoe?")*. Dotted lines indicate the onset of the isolated target label that was repeated two more times during the test phase (e.g., "*Shoe. Shoe.")*.

on the one hand, the positive response to familiar words seems to strengthen as the test unfolds. On the other hand, the looking patterns elicited by novel words do not reliably diverge from the looking patterns displayed at baseline.

As in Experiment 1, we assessed infants' attention to the familiar object during pointing in the experimental condition and during the corresponding time window in the control condition (Figure 4). Infants spent more time fixating the familiar object when it was targeted by pointing ($M = .76$, $SD = .17$) than when it was not highlighted ($M = .59$, $SD = .12$), $Z = 2.688$, $p = .007$, as assessed by a signed-ranks Wilcoxon test (applied because the data in the control condition were not normally distributed). However, in both conditions looking at the familiar object was higher than expected by chance (of .5), experimental condition: $t(19) = 6.91$, $p < .001$, 95% CI = [.68, .83]; control condition: Wilcoxon, $Z = 2.501$, $p = .012$.

3.3. Discussion

In line with our predictions, the results of the experimental condition replicated the results observed in the experimental condition of Experiment 1. When pointing to the familiar object preceded labelling, infants recognized the familiar words, as evidenced by an increase in looking at the familiar objects upon hearing familiar labels (relative to a silent baseline), and successfully interpreted the novel words, as evidenced by an increase in looking to unfamiliar objects when presented with novel labels. This pattern of responses to novel words was only observed in the experimental condition. In the control condition, in which the familiar objects were not highlighted before labelling, hearing novel words did not affect infants' baseline looking preferences. Jointly, these results indicate that nonverbal communication in the form of pointing was critical to eliciting a kind-based representation of familiar objects.

Besides, in the control condition, infants provided evidence for understanding the familiar words only when this condition was presented in the second block. This suggests that a certain amount of exposure to pointing in the experimental condition administered in the first block was necessary for them to make sense of labelling delivered without extra referential signals. We speculate that, when the control condition was presented in the first block, it did not occur to the infants that the disembodied speech they were hearing was supposed to be connected to the objects on the screen. Earlier studies in which infants at this age succeeded in linking familiar words to familiar objects used pointing gestures (e.g., Gliga & Csibra, 2008), live speech produced by the parents (e.g., Bergelson & Swingley, 2012), or both (Parise & Csibra, 2012), but neither of these were included in our control condition. However, infants successfully recognized the familiar words when the control condition was presented in the second block, presumably because the preceding experience with familiar objects explicitly designated as referents by pointing gestures led them to develop an expectation for referential relations between words they heard and objects. In other words, when the experimental condition preceded the control condition, infants might have understood the situation as a naming game during the first block and carried this interpretation on to the second block. If this explanation is correct, it highlights that not only the conceptualization of objects and, in particular, objects seen on a screen, is not automatic at this age, but neither is linking disembodied words to objects that happen to be in front of the infants.

Although familiar word-recognition was affected by the order in which the experimental and control conditions were presented, our findings provide preliminary confirmation that the movement of familiar objects disrupted the interpretation of familiar words in Experiment 1 (control condition). It could have been argued that infants' failure to disambiguate the novel words in the control condition of Experiment 1 was also due to this disruption. However, the fact that in Experiment 2 infants managed to find object referents upon hearing familiar but still could not find the referents of novel words excludes this possibility. Altogether, the patterns of results in the control conditions of Experiments 1 and 2 are consistent with the proposal that kind-based representation of familiar objects is not automatic but is reliably deployed in response to nonverbal communication.

4. General Discussion

Across two experiments, conducted in different labs and using different languages, we established that 12-month-olds were able to disambiguate novel words in a mutual-exclusivity context. More specifically, infants discarded the familiar-kind objects whose labels they knew as referents of novel words. Critically, this interpretation was only deployed when the familiar object was highlighted by a nonverbal communicative action (i.e., via pointing) before labelling. When it was highlighted in a non-communicative manner (i.e., via vertical movement in the control condition of Experiment 1) or when it was not highlighted at all (in the control condition of Experiment 2), infants did not display evidence of referent selection in response to novel words. These results point to two theoretically important conclusions. First, nonverbal communication about a familiar object makes infants represent this object conceptually as an exemplar of an object kind. Second, this representational format enables the operation of mutual-exclusivity inferences responsible for disambiguation of new words. We discuss these points in turn.

4.1. Nonverbal Communication Gives Rise to Conceptual Representation

Human communication carries meaning in the form of concepts (e.g., Macnamara, 1982). Interpreting communication requires setting up specific conceptualizations of objects and events of interest. In verbal communication, the choice of words indicates which among the available concepts should be recruited (Clark, 1997; Lakoff & Johnson, 1980), for example, whether an object should be thought of as a cat, Garfield, a homeless animal, or a family pet. In nonverbal communication, on the other hand, the addressee has to posit a conceptual perspective based on the current context by using her long-term knowledge and experience (Sperber & Wilson, 1986; Tomasello, 2008). Our results provide evidence that 12-month-olds respond to nonverbal communication by setting up kind-based representations of familiar objects referred to through pointing.

Why do infants resort to kind concepts when interpreting nonverbal communication to familiar objects, even though there are many other possible conceptual perspectives? We believe that infants default on kind level of conceptual encoding because it is most easily accessible to them, most relevant in the absence of contextual information indicating otherwise, or both. Adults, for instance, privilege basic-level kind concepts because of their high informativeness (Murphy & Brownell, 1985; Mervis & Crisafi, 1982; Rosh et al., 1976; Rogers & Patterson, 2007) and the resulting frequent use of labels associated with them in communication (Murphy, 2004). Basic-level kind labels are abundant in parental speech and used more often than labels pertaining to other levels of categorization (i.e., parents are more likely to say “cat” rather than “vertebrate” or “European shorthair”, Brown, 1958; Callanan, 1985). This kind of input statistics might lead children to develop expectations about which concepts are most recurrent and, consequently, particularly pertinent in communication. Supporting this idea previous studies show that around one year of age, infants expect verbal labels to be linked to object kinds (Dewar & Xu, 2009; Pomiechowska & Gliga, 2019). The prevalence of kind-based representations continues beyond infancy. Even 3- to 4-years-old children tend to assume that in phrases such as “They like fruit”, “they” stands for a kind (*LEMURS*) rather than a particular set (*these two lemurs*) (Meyer & Baldwin, 2013).

Although privileged by infants in the course of reference assignment, kind-based concepts are not recruited by default when a familiar object becomes available for visual inspection. This might be because infants do not feel compelled to engage in conceptual processing when not justified by the task at hand. Alternatively, they might consider different types of available conceptual representations (e.g., pertaining to kinds, individuals, properties, etc.) and randomly select one, being unable to determine from the contextual cues which one is the most relevant. Beyond communication, tasks that may warrant kind-based encoding involve interpretation of selective goal-directed actions and/preference attribution (Spaepen & Spelke, 2012) and tracking objects through occlusion using working memory (Pomiechowska & Gliga, 2021; see also Kibbe & Leslie, 2019). Finally, it is also important to note that even beyond infancy, how an object is represented depends on the task that one is engaged in. For example, when one has to act on objects, their affordance properties may be privileged over conceptual object representations (e.g., Hodges et al., 1999; Norman, 2002).

We propose that, in the absence of communication, one-year-old infants may not have a reason to use kind-based representations when thinking about the objects they experience and, in particular, about images of objects displayed to them in computerized tasks. Later on, as a result of experience, including communication through various media, adults and older children come to regard computers and screens as representational devices that are used for communication. Therefore, they might spontaneously think of items shown them on a computer screen as ‘presented to them,’ i.e., as referents. In line with this idea, one explanation of why children older than 17-18-month-olds succeed in mutual exclusivity tasks is that they automatically take object images on the screen as communicated to them, interpreting these images as iconic representations of objects of certain kinds (see also, Revencu & Csibra, 2020a,b). Indeed, by this age, infants begin to spontaneously categorize screen-based images in terms of object kinds they depict and succeed at retrieving the associated labels without being prompted by nonverbal referential signals (Mani & Plunkett, 2010).

If the screen is not as clearly a communicative device for one-year-olds as it is for older children, they will not see the object images as representations and will not be compelled to give them conceptual descriptions. In this case, a more explicit referential signal may facilitate the interpretation of an image as a representation, or, minimally, as an exemplar of an object kind (Csibra & Shamsudheen, 2015; Pomiechowska & Gliga, 2021; for a proposal of an architectural implementation, see Brody, 2020). We propose that this is how nonverbal communication in the form of pointing to the familiar objects helped the infants in our experiments overcome the lack of spontaneous kind assignments and enabled them to succeed in disambiguating novel words. Consistent with this view, a previous study showed that 16-month-olds deploy mutual-exclusivity inferences when provided with a video of an experimenter uttering novel labels, but not in the absence thereof (Mather & Plunkett, 2011).

We implemented nonverbal communication in our design by showing infants a photograph of a pointing hand moving vertically above the familiar object (see also, Yin & Csibra, 2015). We remain agnostic as to whether infants posited a presence of a communicative agent hiding behind the display (Saxe, Tzelnic, & Carey, 2007) or interpreted the familiar shape of a pointing hand as a communicative signal without considering its physical source. Importantly, however, we believe that interpreting this stimulus as an instance of nonverbal communication was critical in triggering kind-based representation of the target object. This conclusion is supported by the fact that highlighting the familiar object in the control condition of Experiment 1 did not promote kind-based representation, despite the heightened attention it elicited. Furthermore, it is also substantiated by the previous experimental evidence that infants' representations of objects and events are qualitatively different depending on what directs attention: communicative actions, non-communicative actions (e.g., shaking, Okumura et al., 2020; grasping, Pomiechowska & Csibra, under review) or abstract directional cues (e.g., arrows, flashes, Wu & Kirkham, 2010; Tummeltshammer et al., 2014). While all successfully trigger attentional orienting to the cued locations and objects, only communicative signals prompt infants to set up representations that support learning about them (for locations, Wu & Kirkham, 2010; Tummeltshammer et al., 2014; for objects, Okumura et al., 2020; Pomiechowska & Csibra, under review).

To conclude, our results indicate that familiar objects targeted by nonverbal communication in the form of pointing become represented through familiar object kinds. Nevertheless, several follow-up questions remain open. Can nonverbal communication influence the representation of unfamiliar objects, for example, by prompting infants to set empty conceptual placeholders into which they could collect information about the targeted object and/or its kind (e.g., Csibra & Shamsudheen, 2015)? Does nonverbal communication modulate the representation of the targeted object or also the representation of other objects available for inspection? Under what conditions would other types of conceptual representations (e.g., properties, individuals, actions) be called upon by communication and how would that affect word disambiguation (for evidence about action roles, see Yin & Csibra, 2015)?

4.2. Kind-Based Object Representation Enables Word Disambiguation via Mutual Exclusivity

While it has been widely assumed that novel word disambiguation strategies relying on mutual exclusivity become available only during the second year of life, our findings indicate that they are in place by 12 months of age. We suggest that what the participants in the past studies were missing to successfully carry out mutual exclusivity inferences was the appropriate representation of the familiar objects they were shown and not computational routines, lexical knowledge and principles guiding inferences about meaning, or extensive word-learning experience. Critically, to disambiguate a novel word by discarding familiar objects as its potential referents, it is necessary to represent these objects in a conceptual format that incorporates or provides access to lexical information. We believe that infants' success in interpreting novel words in the present experiment was made possible by the recruitment of kind concepts to represent familiar objects.

How did infants in the current study succeed to disambiguate novel words? The disambiguation process could have proceeded in one of the following two ways. Under one view, conceptualization triggered the retrieval of lexical information associated with the activated concept, thus making the familiar kind label available for comparison with the word provided at test. When the labelling phrase contained a word matching the activated description, infants concluded that it referred to the familiar object. Conversely, when a mismatching word was provided, they excluded the familiar object from consideration and looked for another available object as the potential referent. Furthermore, if one-year-olds, who still struggle with segmenting words from speech (Gervain & Mehler, 2010), retrieved the phonological form of the label upon setting up a conceptual representation of the familiar object, this might have prepared them to segment and process the test utterance because the incoming speech could have been compared against the activated phonological representation (Yeung & Werker, 2009).

Alternatively, lexical access might not have been contingent upon categorization of the familiar triggered by the observation of pointing. Instead, it could have occurred only at test when infants were prompted to look for referents of linguistic expressions. In such a scenario, pointing simply activated an appropriate conceptual description of the highlighted item. Then, the subsequently heard label was independently translated into a conceptual symbol, which either matched the already activated concept linked to the object image or the representation of the exemplar itself (familiar-label condition), or did not match it (novel-label condition).

Whether or not the inferences involved comparing labels or internal kind-symbols, their computational structure must have contained some form of exclusion. Our study cannot inform us on the nature of principles that underlie the process of exclusion applied by infants, i.e., whether it is based on lexical principles (Markman & Watchel, 1988; Golinkoff et al., 1994; Clark, 1990) and heuristics (Byers-Heinlein & Werker, 2009, 2013; Lewis et al., 2020), or on pragmatic assumptions about the speaker's communicative intentions (Diesendruck & Markson, 2001; Tomasello & Barton, 1994). Neither can we determine how exclusion computations are implemented to yield referent disambiguation. It could be an outcome of infants working through a disjunctive syllogism (Halberda, 2003, 2006), involving a representation of alternatives and logical operator of negation, but it could also result from the application of elimination processes that do not rely on propositional logic (Mody & Carey, 2016; Carey & Leahy, 2020).

A recent meta-analysis concluded that developmental changes in lexical knowledge and experience with lexical regularities are key determinants of children's ability to disambiguate new words using mutual exclusivity inferences (Lewis et al., 2020). The current results single out two additional factors that have not been considered in the past theorizing but might significantly contribute to the children's referent disambiguation skills: conceptual knowledge and its deployment in object representation. For making an inference by exclusion, it is critical to represent the to-be-excluded object in a format that is conducive for assessing label-to-referent fits. Therefore, the development of conceptual and category knowledge, which supplies word meanings, on one hand, and the appreciation of referential communication that warrants its deployment in representation, on the other hand, may jointly provide the lower age limit at which mutual exclusivity inferences can be observed.

5. Conclusions

Twelve-month-olds do not necessarily think about familiar objects in terms of their kinds, but they can be explicitly prompted to do so by nonverbal communication. This, in turn, determines the range of inferences that they can carry out. In particular, recruiting kind concepts to represent familiar objects allowed infants in the present study to disambiguate novel words using mutual exclusivity inferences. This early success calls for a reconsideration of the developmental trajectories contended by the previous research, which suggested that the interpretation of ambiguous novel words through mutual exclusivity emerges only during the second year of life, potentially originating in experience with word learning. While our study makes no claim about when exactly the inferential apparatus necessary for such reference disambiguation emerges, it offers evidence that it is in place before infants start to produce language on their own or develop a sizeable receptive lexicon. Mechanisms of inferential referent disambiguation might be fundamental in helping infants to learn new words as soon as they gain an understanding of the forms that referential communication may take and that warrant the deployment of one's conceptual apparatus in comprehension.

Availability of materials and data

The stimuli used in the current experiments and the data analyzed in this manuscript are available in the following OSF repository: https://osf.io/c87sb/.

Acknowledgements

We thank families in London and Budapest for participating in our research; Borbála Széplaki-Köllőd and Sinead Rocha for recording the speech stimuli; Iti Arora, Viktoria

Csink, Dorottya Mészégető, Mária Tóth, Ágnes Volein for help with testing and participant recruitment.

Funding

This work was partially funded by a program grant from the UK Medical Research Council (MR/K021389/1).

Author contributions

B. Pomiechowska, G.Brody, G. Csibra, T. Gliga: Conceptualization, Design & Methodology, Investigation, Data Analysis, Writing – Review & Edition; B. Pomiechowska: Writing – Original Draft, Software, Visualization;

References

Au, T. K. F., & Glusman, M. (1990). The principle of mutual exclusivity in word learning: To honor or not to honor? *Child Development*, *61*(5), 1474-1490. https://doi.org/10.1111/j.1467-8624.1990.tb02876.x

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology*, *64*(1), 3-26. https://doi.org/10.1006/jecp.1996.2332

Bergelson, E., & Aslin, R. (2017). Semantic specificity in one-year-olds' word comprehension. *Language Learning and Development*, 13(4), 481-501. https://doi.org/10.1080/15475441.2017.1324308

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*, 109(9), 3253-3258. https://doi.org/10.1073/pnas.1113380109

Bion, R. A., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, 126(1), 39-53. https://doi.org/10.1016/j.cognition.2012.08.008

Bloom, P. (2002). How children learn the meanings of words. MIT press.

Bohn, M., Call, J., & Tomasello, M. (2015). Communication about absent entities in great apes and human infants. *Cognition*, *145*, 63-72. https://doi.org/10.1016/j.cognition.2015.08.009

Bohn, M., Zimmermann, L., Call, J., & Tomasello, M. (2018). The social-cognitive basis of infants' reference to absent entities. *Cognition*, *177*, 41-48. https://doi.org/10.1016/j.cognition.2018.03.024

Bonatti, L., Frot, E., Zangl, R., & Mehler, J. (2002). The human first hypothesis: Identification of conspecifics and individuation of objects in the young infant. *Cognitive Psychology*, 44(4), 388-426. https://doi.org/10.1006/cogp.2002.0779

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433-436.

Brody, G. (2020). *Indexing Objects In Vision and Communication*. [Unpublished doctoral dissertation]. Central European University.

Brown, R. (1958). How shall a thing be called? *Psychological Review*, 65, 14–21. https://doi.org/10.1037/h0041727

Byers-Heinlein, K., & Werker, J. F. (2013). Lexicon structure and the disambiguation of novel words: Evidence from bilingual infants. *Cognition*, *128*(3), 407-416. https://doi.org/10.1016/j.cognition.2013.05.010

Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: infants' language experience influences the development of a word-learning heuristic. *Developmental Science*, *12*(5), 815-823. https://doi.org/10.1111/j.1467-7687.2009.00902.x

Byers-Heinlein, K., Chen, K. H., & Xu, F. (2014). Surmounting the Tower of Babel: Monolingual and bilingual 2-year-olds' understanding of the nature of foreign language words. *Journal of Experimental Child Psychology*, *119*, 87-100. https://doi.org/10.1016/j.jecp.2013.09.011

Callanan, M. A. (1985). How parents label objects for young children: The role of input in the acquisition of category hierarchies. *Child Development*, 56, 508–523. https://doi.org/10.2307/1129738

Carey, S. (2009) *The origin of concepts*. Oxford University Press.

Cesana-Arlotti, N., Martín, A., Téglás, E., Vorobyova, L., Cetnarski, R., & Bonatti, L. L. (2018). Precursors of logical reasoning in preverbal human infants. *Science*, 359(6381), 1263-1266.

Clark, E. V. (1990). On the pragmatics of contrast. *Journal of Child Language*, 17, 417–443.

Clark, E. V. (1997). Conceptual perspective and lexical choice in acquisition. *Cognition*, 64(1), 1-37. https://doi.org/10.1016/S0010-0277(97)00010-3

Clark, E. V. (1988). On the logic of contrast. *Journal of Child language*, *15*(2), 317-335.

Csibra, G., & Shamsudheen, R. (2015). Nonverbal generics: Human infants interpret objects as symbols of object kinds. *Annual Review of Psychology*, 66, 689-710. https://doi.org/10.1146/annurev-psych-010814-015232

Davidson, D., & Tell, D. (2005). Monolingual and bilingual children's use of mutual exclusivity in the naming of whole objects. *Journal of Experimental Child Psychology*, *92*(1), 25-45. https://doi.org/10.1016/j.jecp.2005.03.007

Dewar, K., & Xu, F. (2009). Do early nouns refer to kinds or distinct shapes? Evidence from 10-month-old infants. *Psychological Science*, *20*(2), 252-257. https://doi.org/10.1111%2Fj.1467-9280.2009.02278.x

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology*, 37(5), 630. https://psycnet.apa.org/doi/10.1037/0012-1649.37.5.630

Ferguson, B., & Waxman, S. R. (2016). What the [beep]? Six-month-olds link novel communicative signals to meaning. *Cognition*, 146, 185-189. https://doi.org/10.1016/j.cognition.2015.09.020

Ferry, A. L., Hespos, S. J., & Waxman, S. R. (2013). Nonhuman primate vocalizations support categorization in very young human infants. *Proceedings of the National Academy of Sciences*, 110(38), 15231-15235. https://doi.org/10.1073/pnas.1221166110

Ferry, A. L., Hespos, S. J., & Waxman, S. R. (2010). Categorization in 3-and 4-month-old infants: an advantage of words over tones. *Child Development*, *81*(2), 472-479. https://doi.org/10.1111/j.1467-8624.2009.01408.x

Futó, J., Téglás, E., Csibra, G., & Gergely, G. (2010). Communicative function demonstration induces kind-based artifact representation in preverbal infants. *Cognition*, 117(1), 1-8. https://doi.org/10.1016/j.cognition.2010.06.003

Friedrich, M., & Friederici, A. D. (2005a). Lexical priming and semantic integration reflected in the event-related potential of 14-month-olds. *NeuroReport*, 16, 653–656.

Friedrich, M., & Friederici, A. D. (2005b). Phonotactic knowledge and lexical-semantic processing in one-year-olds: Brain responses to words and nonsense words in picture contexts. *Journal of Cognitive Neuroscience*, 17, 1785–1802. https://doi.org/10.1162/089892905774589172

Garrison, H., Baudet, G., Breitfeld, E., Aberman, A., & Bergelson, E. (2020). Familiarity plays a small role in noun comprehension at 12–18 months. *Infancy*, 25(4), 458-477. https://doi.org/10.1111/infa.12333

Geeraerts, D., Grondelaers, S., & Bakema, P. (2012). *The structure of lexical variation: Meaning, naming, and context.* Walter de Gruyter.

Gliga, T., & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science*, 20(3), 347-353. https://doi.org/10.1111%2Fj.1467-9280.2009.02295.x

Golinkoff, R. M., Jacquet, R. C., Hirsh-Pasek, K., & Nandakumar, R. (1996). Lexical principles may underlie the learning of verbs. *Child Development*, 67(6), 3101-3119. https://doi.org/10.1111/j.1467-8624.1996.tb01905.x

Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: the case for a developmental lexical principles framework. *Journal of Child Language*, 21(1), 125-155. https://doi.org/10.1017/S0305000900008692

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Journal of Developmental Psychology*, 28 (1), 99–108. https://psycnet.apa.org/doi/10.1037/0012-1649.28.1.99

Graham, S. A., Poulin-Dubois, D., & Baker, R. K. (1998). Infants' disambiguation of novel object words. *First Language*, 18(53), 149-164. https://doi.org/10.1177%2F014272379801805302

Halberda, J. (2006). Is this a dax which I see before me? Use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cognitive Psychology*, 53(4), 310-344. https://doi.org/10.1016/j.cogpsych.2006.04.003

Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, 87(1), B23-B34. https://doi.org/10.1016/S0010-0277(02)00186-5

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language,* 27(3), 689-705. https://doi.org/10.1017/S0305000900004414

Hernik, M., & Broesch, T. (2019). Infant gaze following depends on communicative signals: An eye-tracking study of 5-to 7-month-olds in Vanuatu. *Developmental Science*, *22*(4), e12779. https://doi.org/10.1111/desc.12779

Hodges, J. R., Spatt, J., & Patterson, K. (1999). "What" and "how": evidence for the dissociation of object knowledge and mechanical problem-solving skills in the human brain. *Proceedings of the National Academy of Sciences*, 96(16), 9444-9448. https://doi.org/10.1073/pnas.96.16.9444

Horst, J. S., Samuelson, L. K., Kucker, S. C., & McMurray, B. (2011). What's new? Children prefer novelty in referent selection. *Cognition*, 118(2), 234-244. https://doi.org/10.1016/j.cognition.2010.10.015

Horton, M. S., & Markman, E. M. (1980). Developmental differences in the acquisition of basic and superordinate categories. *Child Development*, 708-719. https://doi.org/10.2307/1129456

Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias. *Infancy*, 15(2), 125-150. https://doi.org/10.1111/j.1532-7078.2009.00009.x

Kibbe, M. M., & Leslie, A. M. (2019). Conceptually rich, perceptually sparse: Object representations in 6-month-old infants' working memory. *Psychological Science,* 30(3), 362-375. https://doi.org/10.1177%2F0956797618817754

Lakoff, G. & Johnson, M. (1980). *Metaphors we live by*, University of Chicago Press, Chicago, IL.

Lewis, M., Cristiano, V., Lake, B. M., Kwan, T., & Frank, M. C. (2020). The role of developmental change and linguistic experience in the mutual exclusivity effect. *Cognition*, 198, 104191. https://doi.org/10.1016/j.cognition.2020.104191

Lin, E. L., Murphy, G. L., and Shoben, E. J. (1997). The effect of prior processing episodes on basic-level superiority. *Quarterly Journal of Experimental Psychology*, 50A, 25–48. https://doi.org/10.1080/713755686

Mani, N., & Plunkett, K. (2010). In the infant's mind's ear: Evidence for implicit naming in 18-month-olds. *Psychological Science*, 21(7), 908-913. https://doi.org/10.1177%2F0956797610373371

Mather, E. (2013). Novelty, attention, and challenges for developmental psychology. *Frontiers in Psychology*, 4, 491. https://doi.org/10.3389/fpsyg.2013.00491

Mather, E., & Plunkett, K. (2010). Novel labels support 10-month-olds' attention to novel objects. *Journal of Experimental Child Psychology*, 105(3), 232-242. https://doi.org/10.1016/j.jecp.2009.11.004

Mather, E., & Plunkett, K. (2011). Mutual exclusivity and phonological novelty constrain word learning at 16 months. *Journal of Child Language*, 38(5), 933.

Mather, E., & Plunkett, K. (2012). The role of novelty in early word learning. *Cognitive Science*, 36(7), 1157-1177.

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, 14(1), 57-77. https://doi.org/10.1207/s15516709cog1401_4

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. https://doi.org/10.1016/0010-0285(88)90017-5

Merriman, W. E., Bowman, L. L., & MacWhinney, B. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development,* i-129. https://doi.org/10.2307/1166130

Mervis, C. B., & Bertrand, J. (1994). Acquisition of the novel name–nameless category (N3C) principle. *Child Development*, 65(6), 1646-1662. https://doi.org/10.1111/j.1467-8624.1994.tb00840.x

Mody, S., & Carey, S. (2016). The emergence of reasoning by the disjunctive syllogism in early childhood. *Cognition*, 154, 40-48. https://doi.org/10.1016/j.cognition.2016.05.012

Murphy, G. (2004). *The big book of concepts*. MIT press.

Murphy, G. L. and Brownell, H. H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 70–84.

Norman, J. (2002). Two visual systems and two theories of perception: An attempt to reconcile the constructivist and ecological approaches. *Behavioral and brain sciences*, *25*(1), 73.

Okumura, Y., Kanakogi, Y., Kobayashi, T., & Itakura, S. (2020). Ostension affects infant learning more than attention. *Cognition*, *195*, 104082. https://doi.org/10.1016/j.cognition.2019.104082

Parise, E., & Csibra, G. (2012). Electrophysiological evidence for the understanding of maternal speech by 9-month-old infants. *Psychological Science*, 23(7), 728-733. https://doi.org/10.1177%2F0956797612438734

Perszyk, D. R., & Waxman, S. R. (2018). Linking language and cognition in infancy. *Annual Review of Psychology*, 69. https://doi.org/10.1146/annurev-psych-122216-011701

Pickron, C. B., Iyer, A., Fava, E., & Scott, L. S. (2018). Learning to individuate: The specificity of labels differentially impacts infant visual attention. *Child Development, 89*(3), 698-710. https://doi.org/10.1111/cdev.13004

Pomiechowska, B., & Gliga, T. (2021). Nonverbal category knowledge limits the amount of information encoded in object representations: EEG evidence from 12-month-old infants. *Royal Society Open Science.*

Pomiechowska, B., Csibra, G., (under review). Nonverbal action interpretation guides novel word disambiguation in 12-month-olds. https://doi.org/10.31234/osf.io/tkhsp

Pomiechowska, B., & Gliga, T. (2019). Lexical acquisition through category matching: 12-month-old infants associate words to visual categories. *Psychological Science*, 30(2), 288-299. https://doi.org/10.1177%2F0956797618817506

Revencu, B., & Csibra, G. (2020, December 29). For 19-month-olds, what happens on-screen stays on-screen. https://doi.org/10.31219/osf.io/bfsyu

Revencu, B., & Csibra, G. (2020, December 11). Opening the black box of early depiction interpretation: From whether to how in the Theory-of-Puppets debate. Commentary on Kominsky, Lucca, Thomas, Frank, & Hamlin (submitted). https://doi.org/10.31234/osf.io/u3em8

Rogers, T. T., & Patterson, K. (2007). Object categorization: Reversals and explanations of the basic level advantage. *Journal of Experimental Psychology: Genera*l, 136(3), 451. https://psycnet.apa.org/doi/10.1037/0096-3445.136.3.451

Rosch, E. (1978). Principles of categorization. In E. Rosch and B. B. Lloyd (Eds.), *Cognition and Categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.

Rosch, E., Mervis, C. B., Gray, W., Johnson, D., and Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382–439. https://doi.org/10.1016/0010-0285(76)90013-X

Saxe, R., Tzelnic, T., & Carey, S. (2007). Knowing who dunnit: Infants identify the causal agent in an unseen causal interaction. *Developmental psychology*, *43*(1), 149. https://psycnet.apa.org/doi/10.1037/0012-1649.43.1.149

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*(9), 668-671. https://doi.org/10.1016/j.cub.2008.03.059

Shamsudheen, R. (2000). *Ostensive Communication Aids Kind-Based Representation of Objects and Acquisition of Generic Knowledge in Infancy* [Unpublished doctoral dissertation]. Central European University.

Spaepen, E., & Spelke, E. (2007). Will any doll do? 12-month-olds' reasoning about goal objects. *Cognitive Psychology*, 54(2), 133-154. https://doi.org/10.1016/j.cogpsych.2006.06.001

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Blackwell.

Spiegel, C., & Halberda, J. (2011). Rapid fast-mapping abilities in 2-year-olds. *Journal of Experimental Child Psychology*, 109(1), 132-140. https://doi.org/10.1016/j.jecp.2010.10.013

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, 13(5), 480-484. https://doi.org/10.1111%2F1467-9280.00485

Surian, L., & Caldi, S. (2010). Infants' individuation of agents and inert objects. *Developmental Science*, 13(1), 143-150. https://doi.org/10.1111/j.1467-7687.2009.00873.x

Tanaka, J. W. and Taylor, M. E. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, 15, 121–149. https://doi.org/10.1016/0010-0285(91)90016-H

Tauzin, T., & Gergely, G. (2019). Variability of signal sequences in turn-taking exchanges induces agency attribution in 10.5-mo-olds. *Proceedings of the National Academy of Sciences*, *116*(31), 15441-15446. https://doi.org/10.1073/pnas.1816709116

Tauzin, T., & Gergely, G. (2018). Communicative mind-reading in preverbal infants. *Scientific reports*, *8*(1), 1-9. https://doi.org/10.1038/s41598-018-27804-4

Thiele, M., Hepach, R., Michel, Ch., & Haun, D. (under review). Observing others'joint attention increases 9-month-old infants' object encoding.

Tomasello, M., & Barton, M. E. (1994). Learning words in nonostensive contexts. *Developmental Psychology*, 30(5), 639. https://psycnet.apa.org/doi/10.1037/0012-1649.30.5.639

Tummeltshammer, K. S., Wu, R., Sobel, D. M., & Kirkham, N. Z. (2014). Infants track the reliability of potential informants. *Psychological Science*, *25*(9), 1730-1738. https://doi.org/10.1177%2F0956797614540178

Van de Walle, G. A., Carey, S., & Prevor, M. (2000). Bases for object individuation in infancy: Evidence from manual search. *Journal of Cognition and Development*, 1(3), 249-280. https://doi.org/10.1207/S15327647JCD0103_1

Wu, R., & Kirkham, N. Z. (2010). No two cues are alike: Depth of learning during infancy is dependent on what orients attention. *Journal of Experimental Child Psychology*, *107*(2), 118-136. https://doi.org/10.1016/j.jecp.2010.04.014

Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition*, 85(3), 223-250. https://doi.org/10.1016/S0010-0277(02)00109-9

Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology*, 30(2), 111-153. https://doi.org/10.1006/cogp.1996.0005

Xu, F., Carey, S., & Quint, N. (2004). The emergence of kind-based object individuation in infancy. *Cognitive Psychology*, 49(2), 155-190. https://doi.org/10.1016/j.cogpsych.2004.01.001

Xu, F., Carey, S., & Welch, J. (1999). Infants' ability to use object kind information for object individuation. *Cognition*, 70(2), 137-166. https://doi.org/10.1016/S0010-0277(99)00007-4