**Moos as cues: One-to-one biases in a non-linguistic and non-communicative domain**

Most objects in children’s daily lives provide them with information in several sensory modalities simultaneously. Learning about these objects involves apprehending these associations. When playing with pets or watching a movie, the child should link a naturalistic “meow” sound to a cat and a “woofing” sound to a dog. Eventually, the child should also link specific spoken words to these same animals, learning that they can be referred to as *cats* and *dogs* or by their typical onomatopoetic sounds*.* Naturalistic animal vocalizations, onomatopoetic sounds, and animal names, therefore provide consistent associations between auditory stimulus and common objects in children’s environments. In this study, we compare children’s efficiency in interpreting the link between animals and these three meaningful sounds. In addition, we investigate whether disambiguation strategies typically characterized as pragmatic or lexically specific (Bloom, 2002; Diesendruck & Markson, 2001; Markman, 1991) can also be used to infer the vocalization produced by novel animals.

Young infants seem to easily create associations between any sound and an object, but older children seem to resist arbitrary non-linguistic associations. While 13-month-olds can link a non-linguistic sound to a novel object, 20-month-olds can only retain the link between words and objects (Woodward & Hoyne, 1999). More recent studies suggest that 12-month-olds already privilege associations between objects and sounds that conform to those commonly used as words in their environment (MacKenzie, Graham, & Curtin, 2011). Infants as young as 6 months of age can use labels but not tones to form object categories (Fulkerson & Waxman, 2007). From these findings, it seems that children’s flexibility in accepting arbitrary associations between sounds and objects decreases with age, with infants increasingly favoring words over other complex auditory stimuli.

However, children must also learn several non-arbitrary and non-linguistic associations between objects and sounds in their environment. Five-month-old infants already recognize meaningful associations between monkey calls and their faces (Vouloumanos, Druhen, Hauser, & Huizink, 2009). While children’s ability to link non-meaningful sounds to objects decreases with age, children become more efficient in recognizing the link between environmental sounds and objects from 15 to 25 months of age (Cummings, Saygin, Bates, & Dick, 2009). Interestingly, children’s accuracy is indistinguishable when environmental sounds (e.g., the sound of a dog barking) and linguistic stimuli (e.g., *dog barking*) are used as cues to object identity.

While much research focuses on the strategies that children use to map words to objects, little is known about their learning of environmental sounds. Typically, word learning is portrayed as an intractable challenge, while associating objects with the sounds they produce might appear trivial. In the case of animal vocalizations, the sounds are causally connected to the entities that produced them. They can be determined by the shape of the vocal track of the animal, are accompanied by synchronous mouth movements, and might be identified by paying attention to the origin of the sound. In contrast, word learning is often characterized as a complex problem of induction. For every possible word, there are many possible referents in the immediate environment (Markman, 1991).

In order to solve the word-leaning puzzle, children are said to be equipped with constraints on the possible meaning of words. The most widely studied of these constraints rules that each word must have only one name (Markman, 1991; Markman & Wachtel, 1988; Markman, Wasow, & Hansen, 2003). Evidence for this default assumption typically comes from disambiguation tasks in which children hear a novel label in the presence of a novel object and one or more familiar objects. In these situations, children tend to select the novel object as the referent for the novel word, presumably because the familiar objects already have names associated with them (Markman, 1991). While this approach to word learning was influential, the origins, scope, and generality of the Mutual Exclusivity (ME) constraint have been widely debated.

A fundamental aspect of this debate concerns whether it is a lexical constraint or the result of inferences about speakers’ communicative intent instead. Lexical accounts characterize ME as a “domain specific mechanism specific to word leaning” (de Marchena, Eigsti, Worek, Ono, & Snedeker, 2011), that “predicts disambiguation only within the domain of word learning (i.e., it is domain-specific)” (Scofield & Behrend, 2003). Pragmatic accounts predict that ME extends to communicative acts more broadly, and reflects assumptions that speakers are cooperative and should use conventional names to refer to familiar objects (Bloom, 2002; Clark, 1990; Diesendruck & Markson, 2001). A third possibility is that the one-to-one expectations observed in word learning and in other communicative domains reflects general biases to find simple regularities in complex domains. This possibility was acknowledged by Markman (Markman, 1991, 1992), and has been embraced by recent computational approaches to word learning (Frank, Goodman, & Tenenbaum, 2009; McMurray, Horst, & Samuelson, unpublished manuscript; Regier, 2003; Regier, et al., 2001).

In summary, there are three main accounts of the scope of the ME constraint. Lexical-specific accounts predict that one-to-one biases are unique to word learning, pragmatic accounts predict that they generalize to communicative acts more broadly, and domain-general accounts predict that they apply to any domain in which consistent one-to-one mappings are observed. To explore the domain-general account, this study investigated whether children show one-to-one biases in a domain that is non-linguistic and non-communicative for them, but in which strong regularities can be found: the vocalizations that animals produce. Importantly, learning about the vocalizations that animals produce should not provide the same challenges that children are supposed to face when learning the meaning of words. Yet, it is a domain in which consistent meaningful associations are observed between objects and auditory cues.

The first question we asked was whether children could appreciate the associations between familiar animals and their characteristic vocalizations. Specifically, we focused on children’s efficiency in recognizing links between animals and three types of auditory stimuli: the natural animal vocalization (e.g., dog barking), the lexicalized animal sound (e.g. *bow-wow*), and the animal name (e.g., *dog*). The second question was whether children showed one-to-one biases for the types of vocalizations that animals produce, similar to their biases in word learning and communicative contexts. In specific, we asked whether children orient to a novel animal, when seeing the picture of a familiar (e.g., dog) and a novel animal (e.g., aardvark) and hearing either a novel animal name or a novel animal vocalization.

**Experiment 1**

The goal of this first study was to compare children’s efficiency in interpreting familiar and novel animal names, onomatopoetic sounds, and animal vocalizations. Our first question was whether these three cues were equally effective in guiding children’s attention to a target animal when both the cue and the animal are familiar. The second question was whether children orient to a novel animal after hearing a novel animal name or a novel animal vocalization, showing disambiguation biases in a non-linguistic and non-communicative domain.

*Method*

*Participants.* Participants were 21 31-month-old children (*M*=31.8 months; range = 30.2-34.3), 10 girls. All were reported by parents to be typically developing and from families where English was the dominant language. Two participants were excluded due to fussiness.

*Visual stimuli.* The visual stimuli were pictures of four familiar animals (horse, dog, cow, sheep) and two novel animals (pangolin, tapir), each centered on a grey background in a 640 x 480 pixel space. The novel animals were real animals, selected because they were uncommon and unlikely to be familiar to the children in the study. All children were reported by parents to have had no exposure to either novel animal.

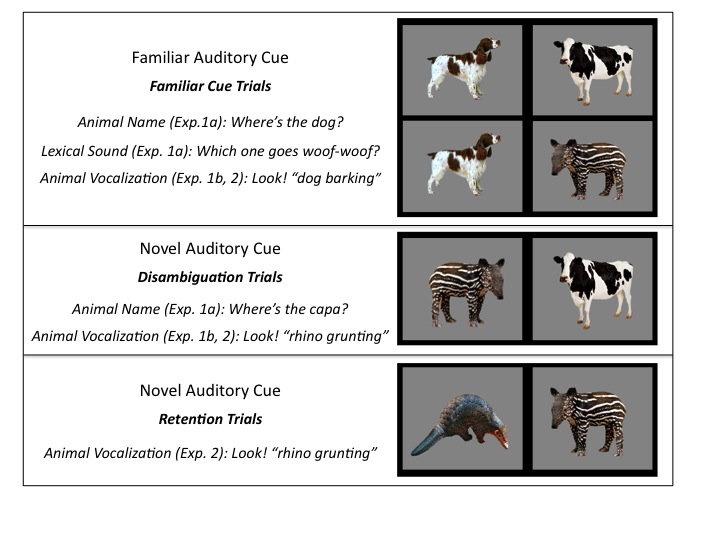
*Auditory stimuli.* The auditory stimuli consisted of auditory cues that were thought to be either Familiar or Novel to 30-month-olds. Familiar cues consisted of *familiar animal names* (horse, dog, cow and sheep)*,* *familiar* *lexical sounds* (neigh, woof-woof, moo and baa)*,* and *familiar natural animal vocalizations* (horse neighing, dog barking, cow mooing and sheep baaing). Novel cues consisted of *novel animal names* (capa, nadu)and *novel natural animal vocalizations* (rhino grunting, gorilla snorting). Trials in which the target cue was an animal name or lexical sound began with a brief carrier frame and concluded with simple questions that served to introduce prosodic variability across trials (e.g., *Where’s the dog?/Which one goes woof? Can you find it?*). The duration of the target cue was 800 ms for lexical sounds and 750 ms for animal names, and the intensity of the phrases was normalized using Praat speech analysis software (Boersma, 2002). Trials with natural animal vocalizations began with a single word recorded by the same native speaker of American English, which served to draw children’s attention (e.g., *Look! “dog barking*”). Familiar animal vocalizations were selected based on representativeness of the familiar animal. Novel animal vocalizations were selected based on unfamiliarity and affordances (i.e., the plausibility that the novel animal could produce the novel sound). All children were reported by parents to have had no exposure to either novel animal’s natural vocalization. The duration of all target animal vocalizations was 2000 ms.

*Books.* To ensure that all children had at least some experience with the familiar animals and familiar auditory cues used in our study, we gave parents two children’s books, both titled *Sounds on the Farm,* at least a week before their visit*.* Parents were instructed to share each book with their child for five to ten minutes in the days leading up to the experiment. The first book was created by our lab and consisted of colorful pictures of each familiar animal and text designed to prompt parents to produce each animal’s lexical sound (e.g., *Wow, look at all those cows! This cow says moo, moo!*). To give children exposure to the natural animal vocalizations, we used a Hear and There BookTM book, which contained buttons that children could press to hear the actual noise that each animal produces.

*Procedure.* Speed and accuracy in identifying the correct target picture was assessed using the looking-while-listening (LWL) procedure (see Fernald, Zangl, Portillo, & Marchman, 2008). On each trial, a pair of pictures was presented on the screen for approximately 4 s, with the auditory stimuli starting after 2 s, followed by 1 s of silence.

Children were tested at two different visits. On the first visit, each child saw 27 trials, consisting of three different trial types (Figure 1). On 8 *Familiar Animal Name* trials, each familiar animal served as the target twice and was paired once with another familiar animal and once with a novel animal. On 8 *Familiar Lexical Sound* trials, each familiar animal served as the target two times and was always paired with another familiar animal. On 6 *Novel Animal Name* trials, each novel animal was labeled three times with a novel animal name (i.e., *capa, nadu*), always paired with a familiar animal. 5 *Filler* trials were interspersed throughout to add variety and maintain children’s attention. Pairings of the novel animal and name, and side of presentation of target object, were counterbalanced across participants. Caregivers wore darkened sunglasses so that they could not influence infants’ looking to the correct picture throughout the 5-min procedure.

On the second visit, children saw 29 trials, consisting of two different trial types (Figure 1). On 16 *Familiar Animal Vocalization* trials, each familiar animal served as the target four times and was paired with another familiar animal twice and each novel animal once. On 8 *Novel Animal Vocalization* trials, each novel animal vocalization served as the target four times and was paired with each familiar animal once. As in the first visit, 5 *Filler* trials were interspersed throughout to add variety and maintain children’s attention. Counterbalancing was the same as in the first visit.



**Figure 1.** Trial types in Experiments 1 and 2 organized by type of cue: Familiar vs. Novel. The target animal for each trial type is on the left. Familiar cues consisted of: a) *familiar animal names* (horse, dog, cow and sheep)*,* b) *familiar* *lexical sounds* (neigh, woof-woof, moo and baa)*,* and c) *familiar natural animal vocalizations* (horse neighing, dog barking, cow mooing and sheep baaing). Novel cues consisted of: a) *novel animal names* (capa, nadu)and b) *novel natural animal vocalizations* (rhino grunting, gorilla snorting).

Participants’ eye movements were video-recorded and coded with a precision of 33 ms by observers who were blind to trial type. Inter- and intra-observer reliability checks were conducted for all coders. For 25% of the subjects, two measures of inter-observer reliability were assessed. The ﬁrst was the proportion of frames (33-ms units) on each trial on which two coders agreed. In this case, agreement was XX%. However, because this analysis included many frames on which the child was maintaining ﬁxation on one picture, we also calculated a more stringent test of reliability. This second measure focused only on shifts in gaze, ignoring steady-state ﬁxations in each trial on which agreement was inevitably high. By this more conservative measure, coders agreed within one frame on XX% of all shifts.

On those trials in which the infant was fixating a picture at the onset of the speech stimulus, accuracy was computed by dividing the time looking to the target object by the time looking to both target and distracter, from 300 to 4300 ms from the onset of the target word. Accuracy before 300 ms was not included because shifts to the target occurring in this window had presumably been initiated before the onset of the noun (Haith, Wentworth, & Canfield, 1993). This analyses window included the entire duration of the trial, and it was longer than that of studies with familiar words (Fernald et al., 1996) because of the longer duration of the animal vocalizations (2 s.) and because of the introduction of novel auditory cues. Analysis windows of up to 10 s. have been used in looking-time experiments with novel words (Mather and Plunkett, 2010). A single analyses window was used for all trial types for consistency, and the entire duration of the trial was used in order to avoid arbitrary decisions. Mean accuracy was then computed for each participant on each trial type.

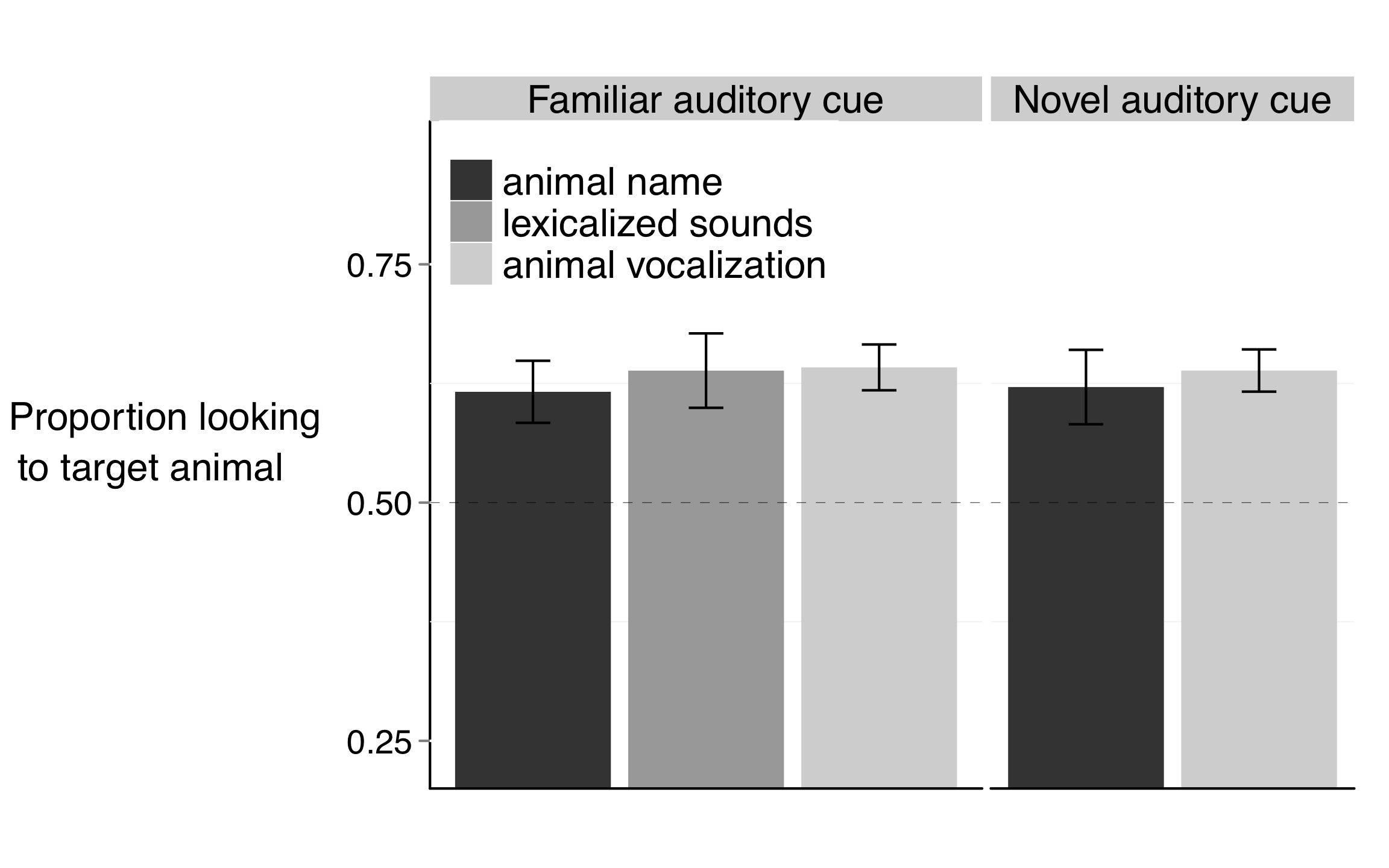
Mean reaction time (RT) for each child was computed from trials when the child stared looking at the distracter image and shifted to the target picture within 300-1800 ms from target-cue onset. Because children vary in the likelihood they will by chance start out on the distracter on a given trial, mean RTs are based on different numbers of trials across participants.

*Results and Discussion*

*Using familiar animal names, lexicalized sounds, and animal vocalization to identify familiar animals:*

Our first question is whether children can use associations between familiar animals and each of three different familiar auditory cues to identify the appropriate animal. Specifically, we evaluate and compare children’s efficiency in recognizing links between animals and their names (e.g., *dog*), their lexicalized sounds (e.g., *bow-wow*), and their natural vocalizations (e.g., *dog barking*).

*Accuracy measures.* Figure 1a presents children’s proportion looking to the familiar animal when hearing each of the three different familiar auditory cues, with looking time averaged over a window from 300 to 4300 ms after the onset of the cue. The three types of auditory cue were equally effective in guiding children’s attention to the target animal. Children looked to the correct animal when hearing the animal name (*M =* 0.62, *t* (18) = 3.57, *p* = 0.002), the lexicalized sound (*M =* 0.64, *t* (18) = 3.54, *p* = 0.002), and the animal vocalization (*M =* 0.64, *t* (18) = 5.90, *p* < 0.001), and performance was indistinguishable across the three conditions (*p* > 0.6). These results show that the animal vocalization can be as good a cue to identify familiar animals as their names or lexicalized sounds.

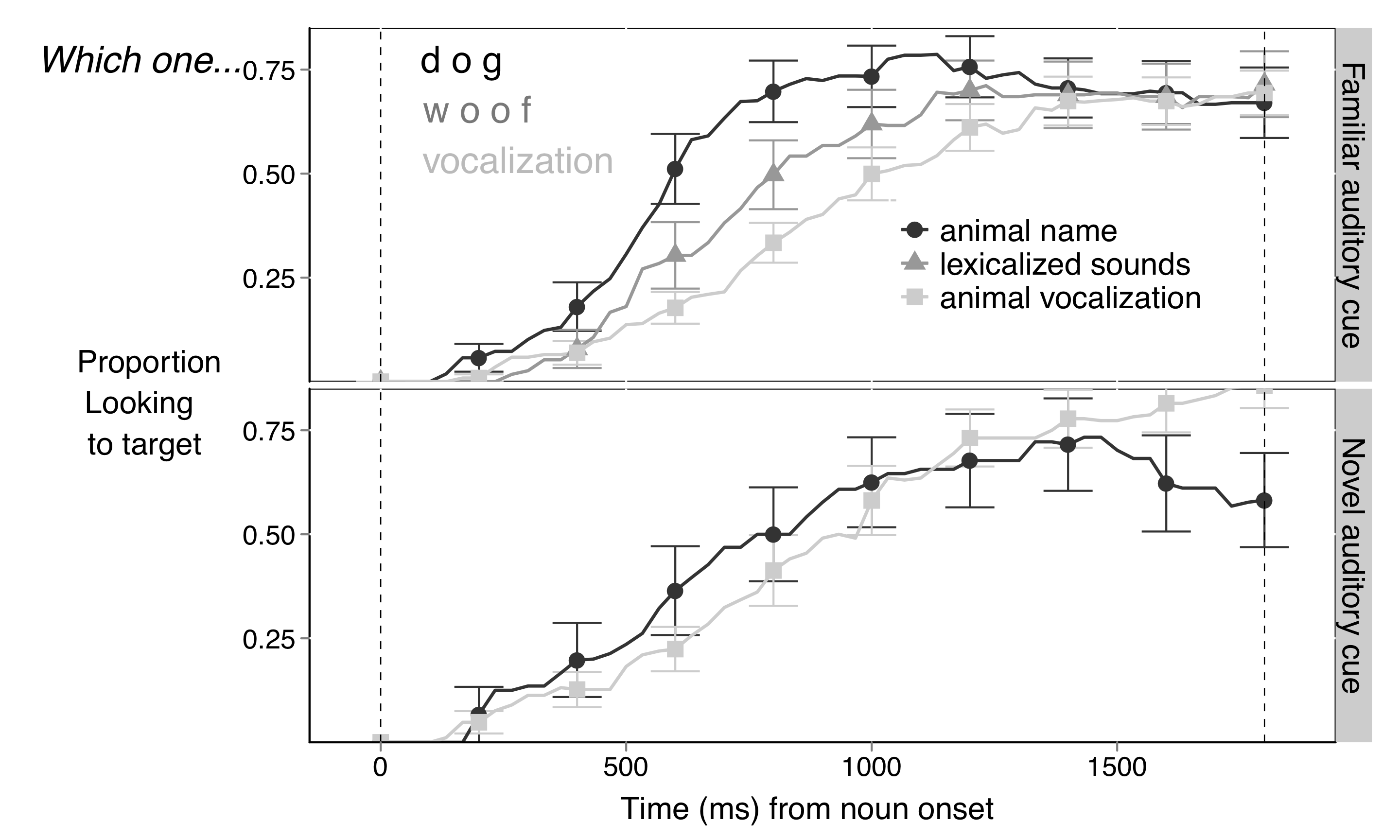


(1.a)

(1.b)

**Figure 1.** Accuracy of responses to familiar and novel auditory cues. 1.a. When hearing familiar animal names, lexicalized sounds, or animal vocalizations, children reliably looked to the target familiar animal. 1.b. When hearing novel animal names or animal vocalizations, children reliably looked at the novel animal. The different auditory cues were equally effective in guiding children’s attention to the target.

*Reaction-time measures:* Figure 2a presents children’s speed in recognizing animals on the different trial types, showing that the animal name can be more rapidly exploited to identify animals than the lexicalized sound or animal vocalization. That is, children were faster to identify the target animal when hearing its name (*M* = 541ms), as compared to its vocalization (*M* = 840 ms, *t* (18) = 4.61, *p* < 0.001) or lexicalized sound (*M* = 801 ms, *t* (18) = 3.40, *p* = 0.002), with the last two conditions not differing from each other (*t* (18) = 0.47, *p* > 0.6).



(2.b)

(2.a)

**Figure 2.** Time course of children’s looking to the target animal after hearing different familiar and novel auditory cues. The curves depict changes in the proportion of looking to the target animal as sounds unfolded, measured from sound onset (in milliseconds). When hearing a familiar auditory cue (2.a), children were faster to orient to the target animal after hearing its name, than when hearing the lexicalized sound or animal vocalization. Children were equally fast to orient to the novel animal after hearing a novel animal name or vocalization (2.b).

*Using novel animal names and animal vocalization to disambiguate novel animals*

Our second question is whether children would orient to a novel animal after hearing a novel animal name or vocalization. Specifically, we evaluate and compare children’s response to a novel animal name or animal vocalization in the presence of both a familiar and a novel animal.

*Accuracy measures:* Figure 1b shows children’s proportion of looking to the novel animal when hearing a novel animal name or a novel animal vocalization. Children reliably looked to the novel animal when hearing the novel animal name (*M* = 0.62) or novel animal vocalization (*M* = 0.63). Performance did not differ between these two conditions (*t* (18) = 0.38, *p >* 0.70). Therefore, children seem to have one-to-one biases for the vocalizations that animals produce already at 30 months of age, the earliest age at which the disambiguation effect has been observed in a domain other than word learning.

*Reaction-time measures:* As seen in Figure 2b, there was no difference in RT between children’ speed in recognizing the novel animal after hearing a novel animal name (*M =* 783 ms) or animal vocalization (*M =* 770 ms, *t* (18) = 0.08, *p* > 0.9).

**Experiment 2**

An issue that has received much attention in recent years concerns the relation between children’s referent selection and retention abilities. While earlier studies tended to conflate disambiguation strategies and children’s word learning, more recent studies suggest that these two abilities might operate at different timescales (Bion, Borovsky, & Fernald, under review; Horst, McMurray, & Samuelson, 2006; McMurray, 2007; McMurray, Aslin, & Toscano, 2009; McMurray, et al., unpublished manuscript; McMurray, Horst, Toscano, & Samuelson, 2009).

Horst and Samuelson (2008) examined both referent selection and retention in four experiments with 2-year-olds. When children were shown a novel object among familiar objects, they selected the novel object when hearing a novel label, as found in previous studies. But surprisingly, on retention trials 5 min later, these children showed no evidence of remembering the names of the novel objects they had previously identified. Similarly, Bion et al. (under review) found that children as old as 30 months of age do not retain the names for novel objects, despite previously looking at them when hearing a novel label in the presence of a familiar object. Importantly, when no familiar object is present, children retained the names of the objects already at 18 months of age.

Experiment 2 aimed at replicating and expanding on two findings from Experiment 1: First, children identified familiar animals based on the vocalizations they produce. Second, children used novel vocalizations to disambiguate novel animals. In addition, we also ask whether children can retain the link created through disambiguation between a novel animal and a novel animal vocalization.

*Method*

*Participants.* Participants were 22 31-month-old children (*M*=31.1 months; range = 27.4-32.5), 12 girls. All were reported by parents to be typically developing and from families where English was the dominant language.

*Visual stimuli.* The visual stimuli were the same as in Experiment 1, except for the novel animals (aardvark and capybara), which replaced the novel animals (pangolin and tapir) used in Experiment 1. All children were reported by parents to have had little to no exposure to the novel animals.

*Auditory stimuli.* The auditory stimuli consisted of only the natural animal vocalizations and they were the same as in Experiment 1.

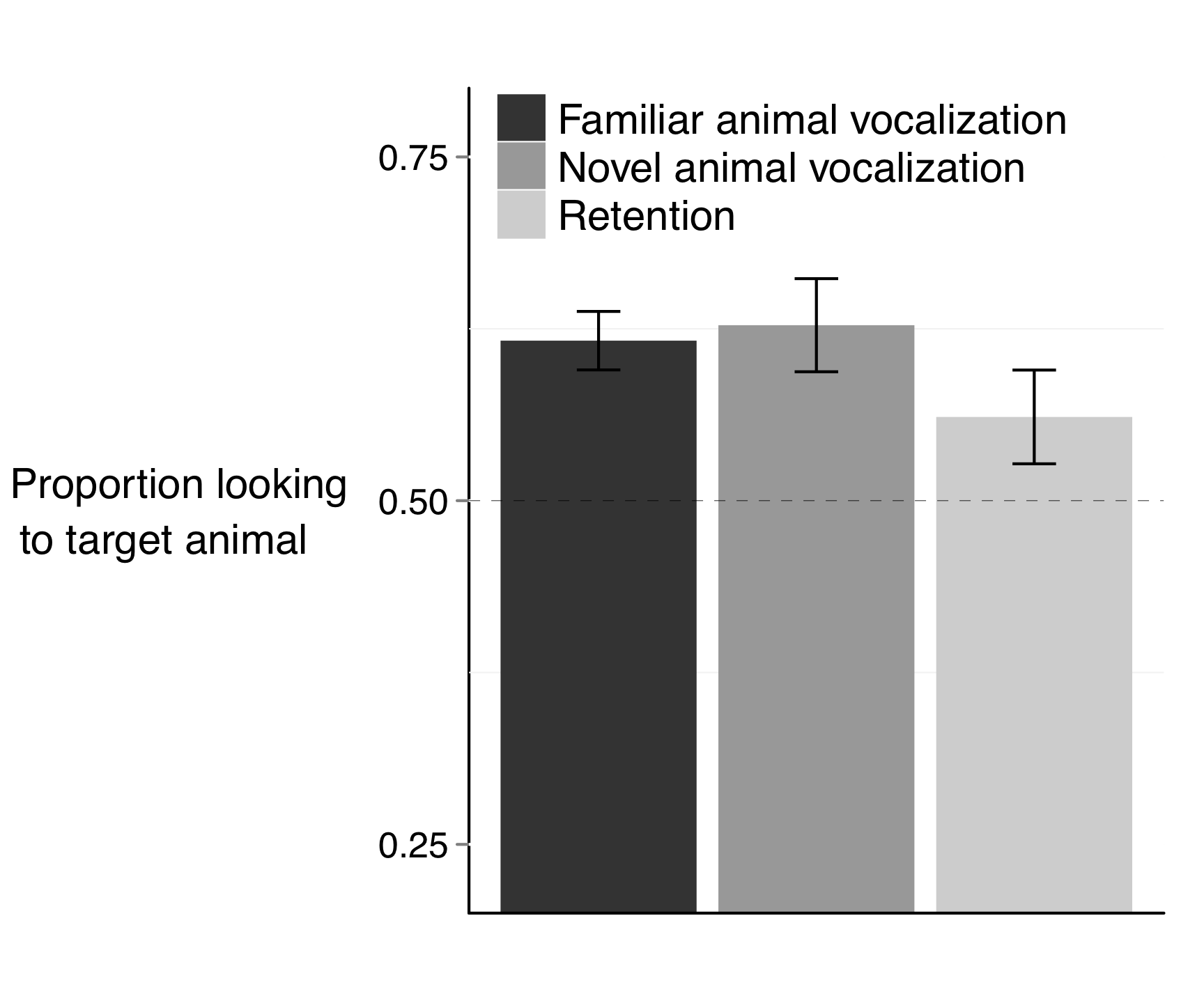
*Books.* As in Experiment 1, we sent home a children’s book to ensure that all participants had at least some exposure to the familiar animals and auditory cues. Since, in Experiment 2, we were interested in the natural animal vocalizations and not the names/lexical sounds, only the Hear and ThereTM *Sounds on the Farm* book was used. Instructions given to the parents were the same as in Experiment 1, and the book was sent home a week before the visit.

*Procedure.* Experiment 2 consisted of one visit. Each child saw 35 trials, consisting of three trial types (Figure 1). The 16 *Familiar Animal Vocalization* trials and 8 *Novel Animal Vocalization* trials were identical in structure to Experiment 1. In addition, on 6 *Retention* trials, the two novel animals were presented side by side, with each serving as the target three times. The same coding and speed/accuracy measures were used as in Experiment 1.

*Results and Discussion:*

*Retention of the link between a novel animal and a novel vocalization:*

*Accuracy measures:* Figure 4 presents children’s proportion looking to the target animal after hearing a familiar or a novel animal vocalization over a window from 300 to 4300 ms after the onset of the vocalization. When children heard a familiar animal vocalization, they oriented to the target familiar animal (*M =* 0.62, *t* (21) = 5.47, *p* < 0.001). When children heard a novel animal vocalization, they looked at a novel animal instead (*M =* 0.63, *t* (21) = 3.76, *p* = 0.001). These results replicate three main findings from Experiment 1.

****

**Figure 1.** Accuracy of responses to familiar and novel animal vocalizations. Children reliably looked to the target animal after hearing a familiar animal vocalization and a novel animal vocalization. Children were marginally successful in retaining the map between a novel animal and a novel vocalization. Performance was indistinguishable across the three conditions.

Can children remember the association between the novel animal and the novel vocalization, when this association is created through a disambiguation strategy? In order to test this hypothesis, two novel animals were paired with each other, as children heard the animal vocalization previously associated with one of the two animals.

Children remembered the link between the novel animal and the novel vocalization (Figure 4, *M =* 0.56, *t* (21) = 1.79, *p* = 0.089). Performance on retention trials was marginally above chance, and was indistinguishable from performance on trials in which children had to identify a familiar animal (*t* (21) = 1.38, *p* = 0.177) or disambiguate a novel animal (*t* (21) = 1.39, *p* = 0.176).

Three findings emerged from this analyses of accuracy, the first two replicated findings from Experiment 1: First, children oriented to a familiar animal after hearing a familiar animal vocalization. Second, children oriented to a novel animal after hearing a novel animal vocalization. In addition, we also found that children were marginally successful in retaining the link between a novel animal vocalization and a novel animal.

*RT measures:* Reaction time measures were indistinguishable between the three experimental conditions in this experiment. Children were equally fast to orient to a familiar animal after hearing a familiar animal vocalization (*M =* 730 ms), to orient to a novel animal after hearing a novel animal vocalization (*M =* 809 ms), and on retention trials to orient to the novel animal previously paired with a novel animal vocalization (*M* = 824 ms, all *p* > 0.25).

*General Discussion.*

This study investigated children’s interpretation of familiar and novel animal vocalizations. Our first finding was that children reliably looked at a familiar animal when hearing a familiar animal vocalization, with accuracy comparable to their interpretation of the animal name and its onomatopoetic sound, but with reaction times slower than their processing of familiar animal names. Our second finding was that children reliably looked at a novel animal when hearing a novel animal vocalization, with accuracy comparable to their disambiguation of a novel animal name. Our third finding is that children were only marginally successful in retaining the link between the novel animal and its vocalization, supporting dissociation between performance on disambiguation trials and subsequent retention (McMurray, Horst, et al., 2009).

Previous studies compared the processing of familiar words and environmental sounds, including but not analyzing separately performance for animal names and vocalizations. Normally developing children did not differ in their processing of environmental and verbal sounds (Cummings, et al., 2009), but autistic children and children with developmental language impairment showed an advantage in their processing of animal sounds (Cummings & Čeponienė, 2010; McCleery, et al., 2010). An advantage in the processing of environmental sounds was also found in studies with adults on crossmodal priming (Chen & Spence, 2011). Chen and Spence interpreted this finding as evidence for a dual processing route model, in which environmental sounds are directly connected to semantic representations, but words are linked to the semantic representations through the lexical system. Recent brain imaging studies, however, seem to suggest that largely overlapping networks process verbal and non-verbal meaningful semantic information (Cummings, et al., 2006; Dick, et al., 2007; Liu & Holt). The fact that our children were faster to process animal names as compared to their vocalization might be due to the fact words were produced within familiar frames, because words were shorter than the animal vocalizations, or because children hear animal names more frequently than animal vocalizations. Most of the children in our study experienced the sounds produced by the animals mainly through multimedia material, rather than interacting with real animals.

Little is known about children’s and adults’ processing of onomatopeic sounds. Hashimoto et al. (Hashimoto, et al., 2006) compared brain responses to nouns, animal sounds, and onomatopoetic sounds, and found that onomatopoeic sounds were processed by extensive brain regions involved in the processing of both verbal and nonverbal sounds. The authors conclude that onomatopoeic sounds might serve as a bridge between nouns and animal sounds. A similar hypothesis is proposed by Cummings et al., who suggests that onomatopoeic sounds might provide young children with information about intermodal associations, provide a link between non-arbitrary environmental sounds and arbitrary word-object associations (Cummings, et al., 2009). Fernald and Morikawa (1993) reported that 52% of Japanese mothers use onomatopoeic sounds to label target objects, while only 1 in 30 American mothers did so. It would be interesting to investigate how this cross-cultural variation might impact early word learning.

Our third finding was that children looked at a novel animal when hearing a novel animal vocalization, with accuracy comparable to their disambiguation of novel animal names. In the book *How Children Learn the Meaning of Words*, Paul Bloom discusses different theories explaining children’s disambiguation biases. These biases could be a specifically lexical phenomenon that applies only to words (lexical account), a product of children’s theory of mind restricted to communicative situations (pragmatic account), or a special case of a general principle of learning that exaggerates regularities across domains (domain-general account). As a possible way to decide among these three theories, he suggests that it would be interesting to know if children show disambiguation biases when learning the sounds that different animals make (p., 69, Bloom, 2002).

Children older than three years of age show disambiguation biases in different communicative domains. For example, they expect faces to map to individual voices (Moher, Feigenson, & Halberda, 2010), and they also expect speakers to use consistent facts to refer to objects (Diesendruck & Markson, 2001). Recent studies suggest that different strategies might be used to make inferences about speakers’ communicative intent and the meaning of a novel word. Autistic children who are known to have pragmatic deficits show disambiguation biases and select a novel word when hearing a novel object (Preissler & Carey, 2005). Disambiguation biases for words are correlated with vocabulary, and disambiguation biases for facts are correlated with social-pragmatic skills (de Marchena, et al., 2011). As these authors acknowledge, these findings suggest that disambiguation biases for words might not be motivated by pragmatic inferences, but they do not rule out domain-general accounts.

Only one study compared disambiguation biases for words and nonlinguistic stimuli. In contexts in which learning could benefit from disambiguation biases, Yoshida et al. (Yoshida, Rhemtulla, & Vouloumanos, in press) found an advantage for the learning of speech over nonspeech sounds. Again, these findings do not rule out a domain-general account, since participants had no reason to expect that the mapping between tones and objects should be mutually exclusive. As pointed out by the authors, perhaps similar strategies might be applied to nonlinguistic sounds if they were meaningfully related to the object.

Our study was the first one to demonstrate that young children show disambiguation biases in a nonlinguistic domain. It was also the youngest age at which children were shown to show disambiguation biases in a situation other than word learning. The similar comparable behavior for the disambiguation of names and vocalizations does not necessarily mean that children are using the same strategy to motivate their responses. However, we emphasize that crossituational models of word learning that were used to explain disambiguation biases, can also explain our findings with animal vocalizations. This pattern of responses emerges by looking for regularities in complex learning domains that have consistent crossmodal mappings (Frank, et al., 2009; McMurray, Horst, et al., 2009; Regier, 2003; Yu & Smith, 2007).

Interestingly, other predictions of word learning models relating disambiguation biases and retention were also supported by our study. McMurray and colleagues propose that referent selection and word learning are related, but should not be conflated (Horst, et al., 2006; McMurray, Aslin, et al., 2009). Referent selection requires that children give their best guess in a specific ambiguous situation, but word learning operates over a much longer time scale. Although disambiguation can be viewed as the *product* of learning that has occurred up to that point, for younger children it does not necessarily result in learning a new word. We have shown in this study that children looked at a novel animal when hearing a novel animal vocalization, and yet showed only fragile evidence of retaining this link over time.

Children’s learning about objects in their environment involves more than learning their names. Nelson argued throughout her work that before object names are learned, sounds and actions might form the basis on which objects are conceptualized (Nelson, 1973). For example, children might see barking as a defining feature of dogs, and often say bow-wow in response to the picture of a dog, even before they learn the animal name (Nelson, 1973). Learning the meaning of an object therefore requires learning several cross-modal associations, including learning about the object’s texture, smell, the sounds it makes, and eventually its name. Independently of how these associations are learned, they are meaningfully and systematically associated with the object. As a consequence, children might expect these associations to confirm to their prior expectations and experiences with the world, and be more likely to map novel stimuli to novel objects. For example, a child familiar with the smell of freshly baked cookies might expect their mom to be cooking another kind of food when hearing the unfamiliar smell of lasagna coming from the oven.

Children use all the knowledge that they have available in order to make sense of a constantly changing world. In the case of animal vocalizations, they observed constant mappings and expect novel sounds to map to novel animals. They probably also rely on other cues, such as the shape of the vocal track of animals, their location and size, or previous associations. In the case of words, children learn to rely on different kinds of cues, including semantics (Goodman, McDonough, & Brown, 1998), cross-situational statistics (Goodman, et al., 1998), syntax (Brown, 1957), and pragmatic and social cues (Baldwin, 1991, 1993), and disambiguation (Markman, 1991). As children grow older, these different sources of information might converge as evidence that a novel word should refer to a novel object. They might be motivated to select a novel object after hearing a novel label because of several converging cues. Children can rely on their knowledge about the world and about previous experiences with words in order to figure out what speakers are talking about – a task we continue to do throughout our lives in learning knew words and interpreting complex sentences.

*References*

Baldwin, D. (1991). Infants' contribution to the achievement of joint reference. *Child Development, 62*(5), 874-890.

Baldwin, D. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language, 20*, 395-395.

Bion, R., Borovsky, A., & Fernald, A. (under review). Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30 months.

Bloom, P. (2002). *How children learn the meanings of words*: The MIT Press.

Brown, R. W. (1957). Linguistic determinism and the part of speech. *The Journal of Abnormal and Social Psychology, 55*(1), 1.

Chen, Y. C., & Spence, C. (2011). Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. *Journal of Experimental Psychology: Human Perception and Performance, 37*(5), 1554-1568.

Clark, E. (1990). On the pragmatics of contrast. *Journal of Child Language, 17*(2), 417-431.

Cummings, A., & Čeponienė, R. (2010). Verbal and nonverbal semantic processing in children with developmental language impairment. *Neuropsychologia, 48*(1), 77-85.

Cummings, A., Čeponienė, R., Koyama, A., Saygin, A., Townsend, J., & Dick, F. (2006). Auditory semantic networks for words and natural sounds. *Brain research, 1115*(1), 92-107.

Cummings, A., Saygin, A. P., Bates, E., & Dick, F. (2009). Infants' recognition of meaningful verbal and nonverbal sounds. *Language Learning and Development, 5*(3), 172-190.

de Marchena, A., Eigsti, I., Worek, A., Ono, K., & Snedeker, J. (2011). Mutual Exclusivity in Autism Spectrum Disorders: Testing the Pragmatic Hypothesis. *Cognition*.

Dick, F., Saygin, A. P., Galati, G., Pitzalis, S., Bentrovato, S., D'Amico, S., et al. (2007). What is involved and what is necessary for complex linguistic and nonlinguistic auditory processing: evidence from functional magnetic resonance imaging and lesion data. *Journal of cognitive neuroscience, 19*(5), 799-816.

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology, 37*(5), 630-641.

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development, 64*(3), 637-656.

Frank, M., Goodman, N., & Tenenbaum, J. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science, 20*(5), 578.

Fulkerson, A. L., & Waxman, S. R. (2007). Words (but not tones) facilitate object categorization: Evidence from 6-and 12-month-olds. *Cognition, 105*(1), 218-228.

Goodman, J. C., McDonough, L., & Brown, N. B. (1998). The role of semantic context and memory in the acquisition of novel nouns. *Child Development, 69*(5), 1330-1344.

Hashimoto, T., Usui, N., Taira, M., Nose, I., Haji, T., & Kojima, S. (2006). The neural mechanism associated with the processing of onomatopoeic sounds. *Neuroimage, 31*(4), 1762-1770.

Horst, J. S., McMurray, B., & Samuelson, L. K. (2006). Online processing is essential for learning: Understanding fast mapping and word learning in a dynamic connectionist architecture. *Proc. of CogSci’06*.

Liu, R., & Holt, L. L. Neural changes associated with nonspeech auditory category learning parallel those of speech category acquisition. *Journal of cognitive neuroscience, 23*(3), 683-698.

MacKenzie, H., Graham, S. A., & Curtin, S. (2011). Twelve‐month‐olds privilege words over other linguistic sounds in an associative learning task. *Developmental Science, 14*(2), 249-255.

Markman, E. (1991). *Categorization and naming in children: Problems of induction*: The MIT Press.

Markman, E. (1992). Constraints on word learning: Speculations about their nature, origins, and domain specificity. *Modularity and constraints in language and cognition, 25*, 59-101.

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

Markman, E., Wasow, J., & Hansen, M. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology, 47*(3), 241-275.

McCleery, J. P., Ceponiene, R., Burner, K. M., Townsend, J., Kinnear, M., & Schreibman, L. (2010). Neural correlates of verbal and nonverbal semantic integration in children with autism spectrum disorders. *Journal of Child Psychology and Psychiatry, 51*(3), 277-286.

McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science, 317*(5838), 631.

McMurray, B., Aslin, R. N., & Toscano, J. C. (2009). Statistical learning of phonetic categories: insights from a computational approach. *Developmental Science, 12*(3), 369-378.

McMurray, B., Horst, J., & Samuelson, L. (unpublished manuscript). Using your lexicon at two timescales: Investigating the interplay of word learning and recognition.

McMurray, B., Horst, J. S., Toscano, J. C., & Samuelson, L. K. (2009). . Integrating Connectionist Learning and Dynamical Systems Processing: Case Studies in Speech and Lexical Development. *Toward a Unified Theory of Development Connectionism and Dynamic System Theory Re-Consider, 1*(9), 218-251.

Moher, M., Feigenson, L., & Halberda, J. (2010). A One‐to‐One Bias and Fast Mapping Support Preschoolers’ Learning About Faces and Voices. *Cognitive Science, 34*(5), 719-751.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, 1-135.

Preissler, M. A., & Carey, S. (2005). The role of inferences about referential intent in word learning: Evidence from autism. *Cognition, 97*(1), B13-B23.

Regier, T. (2003). Emergent constraints on word-learning: A computational perspective. *TRENDS in Cognitive Sciences, 7*(6), 263-268.

Regier, T., Corrigan, B., Cabasaan, R., Woodward, A., Gasser, M., & Smith, L. (2001). *The emergence of words*.

Scofield, J., & Behrend, D. A. (2003). Two-year-olds differentially disambiguate novel words and facts. *Unpublished manuscript, University of Arizona*.

Vouloumanos, A., Druhen, M. J., Hauser, M. D., & Huizink, A. T. (2009). Five-month-old infants' identification of the sources of vocalizations. *Proceedings of the National Academy of Sciences, 106*(44), 18867-18872.

Woodward, A., & Hoyne, K. (1999). Infants' learning about words and sounds in relation to objects. *Child Development, 70*(1), 65-77.

Yoshida, K., Rhemtulla, M., & Vouloumanos, A. (in press). Exclusion constraints facilitate statistical word learnin. *Cognitive Science*.

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science, 18*(5), 414.