**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 stimuli 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](#_ENREF_7); [Diesendruck & Markson, 2001](#_ENREF_19); [Markman, 1991](#_ENREF_31)) can also be used to infer the vocalization produced by novel animals.

The question of whether words are a special kind of stimuli for infants is not new. Several studies have found advantages for speech over tones in object individuation and categorization in young infants ([Fulkerson & Waxman, 2007](#_ENREF_23); [Xu, 2002](#_ENREF_59)). Focusing on associations between objects and sounds, objects and tones, or objects and gestures, several studies found that younger infants accept many forms as object labels, but older infants seem to favor words ([Namy & Waxman, 1998](#_ENREF_43); [Woodward & Hoyne, 1999](#_ENREF_58)). And a different line a research found that infants prefer to hear words over some nonlinguistic analogues ([A Vouloumanos & Werker, 2004](#_ENREF_52), [2007a](#_ENREF_53), [2007b](#_ENREF_54)), and that the neonate brain responds differently to speech as compared to backwards speech ([Peña et al., 2003](#_ENREF_45)). These studies found advantages for speechover nonlinguistic analogues in categorization, individuation, crossmodal association, and speech preferences. But they focused on arbitrary nonlinguistic cues that are not consistently associated to objects in children's environments.

Other studies approached the question of whether speech is special from a different perspective, comparing the processing of words and *meaningful* environmental sounds, such as animal vocalizations or the sounds that objects make (e.g., a cow mooing or a car starting). Studies with adults found similarities and differences in both behavioral and neural responses to crosmodal semantic associations between words and environmental sounds. In a task involving picture detection, Chen and Spence ([2011](#_ENREF_10)) found priming from environmental sounds but not from words. The authors propose that recognition of environmental sounds is faster because words go though a lexical stage of processing, while environmental sounds activate semantic representations directly. In contrast, in a task in which participants had to decide whether a sound and picture matched, Lupyan and Thompson-Schill ([2012](#_ENREF_30)) found advantages for words as compared to environmental sounds. This finding was seen as evidence that words activate conceptual information more quickly and accurately and in a more categorical way than nonverbal sounds.

In an ERP imaging study, Cummings et al. ([2006](#_ENREF_12)) looked at semantic integration of verbal and non-verbal sounds and objects, and found that largely overlapping neural networks processed verbal and non-verbal meaningful sounds. In another study focusing on three different meaningful sound types that differ in arbitrariness, Hashimoto et al. ([2006](#_ENREF_25)) found different neural mechanisms for the processing of animal names and vocalizations, with onomatopoetic words activating both areas (Hashimoto, et al., 2006). Research on environmental sounds is still at its beginning stages, and it is hard to conciliate the different findings. Different tasks or timing of stimuli can lead to different findings, and different theoretical commitments can lead to different interpretations. For example, environmental sounds are often treated as encompassing both the sounds of living and man-made objects (e.g. cow mooing and car starting) despite evidence that these sounds are treated differently by the adult brain ([Murray, Camen, Andino, Bovet, & Clarke, 2006](#_ENREF_42)). But this new line of research is promising in that it gives us a new way to look at old questions such as whether language emerges from the interaction of domain general cognitive processes or domain-specific mechanism ([Bates & MacWhinney, 1989](#_ENREF_3)), serving as stimuli to compare the processing of speech and non-speech sounds in research on autism, developmental language impairment, or cochlear implants ([Cummings & Čeponienė, 2010](#_ENREF_13); [McCleery et al., 2010](#_ENREF_35)).

From a developmental perspective, it is also important to understand how children process words and meaningful nonlinguistic sounds. Very few studies looked at this question. In a study using a preferential looking paradigm, Cummings, Saygin, Bates, and Dick ([2009](#_ENREF_15)) found that 15- and 25-mo-olds can use words and environmental sounds to guide their attention to familiar objects, getting better with as they get older. Vouloumanos and Werker found that 5-month-olds can match some animals to their vocalizations ([A. Vouloumanos, Druhen, Hauser, & Huizink, 2009](#_ENREF_55)). And studies with children with autism and developmental language impairment found more severe deficits for the processing of words than environmental sounds, confirming the linguistic nature of these disorders (McCleery et al., 2010; Cummings & Čeponienė, 2010). Our first question builds on these earlier studies, focusing on three particular types of meaningful auditory stimuli that are consistently associated to common objects in children's environment but that differ in arbitrariness. The first question we ask in this research is whether 30 month-olds could use familiar animal names (e.g., *dog*), onomatopoeic sounds (e.g. *bow-wow*), and animal vocalizations (e.g., dog barking), to identify a familiar animal, and whether one of these three sounds is a particularly effective cue in guiding children's attention to an animal in their environment. Towards this end, we use the looking-while-listening task, a paradigm that has been widely used to test word recognition in young infants ([Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998](#_ENREF_21)).

In addition to a question about the processing of these sounds, we also focus on how linguistic and nonlinguistic sounds are learned by young children. 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](#_ENREF_31)).

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](#_ENREF_31); [Markman & Wachtel, 1988](#_ENREF_33); [Markman, Wasow, & Hansen, 2003](#_ENREF_34)). 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](#_ENREF_31)). 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](#_ENREF_17)), that “predicts disambiguation only within the domain of word learning (i.e., it is domain-specific)” ([Scofield & Behrend, 2003](#_ENREF_49)). 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](#_ENREF_7); [Clark, 1990](#_ENREF_11); [Diesendruck & Markson, 2001](#_ENREF_19)). 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 third possibility was acknowledged by Markman ([Markman, 1991](#_ENREF_31), [1992](#_ENREF_32)), and has been embraced by recent computational approaches to word learning ([Frank, Goodman, & Tenenbaum, 2009](#_ENREF_22); [B. McMurray, Horst, & Samuelson, unpublished manuscript](#_ENREF_38); [Regier, 2003](#_ENREF_47); [Regier et al., 2001](#_ENREF_48)).

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 second question in this study therefore was whether children show one-to-one biases for the types of vocalizations that animals produce, similar to their biases in word learning in communicative contexts. Specifically, we ask whether children would 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 would 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.

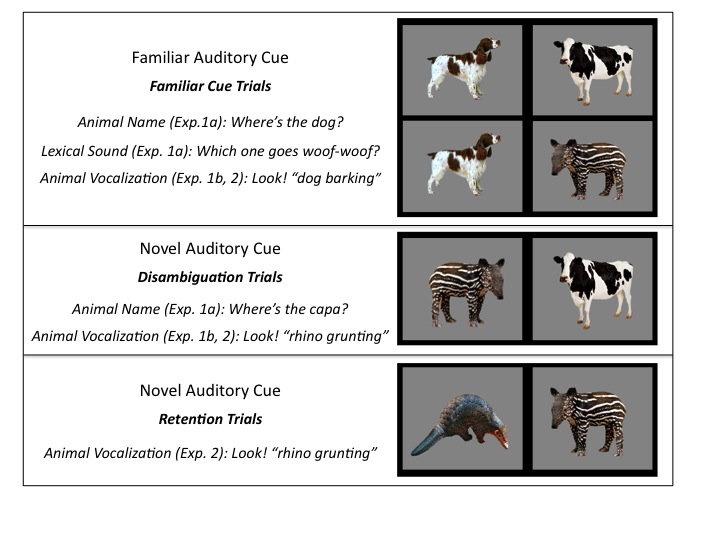
*Auditory stimuli.* The auditory stimuli consisted of auditory cues that were thought to be either Familiar or Novel to 30-month-olds. Familiar cues consisted of *familiar animal names* (horse, dog, cow and sheep)*,* *familiar* *lexical sounds* (neigh, woof-woof, moo and baa)*,* and *familiar natural 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).

*Measures of processing efficiency*

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 98%. 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 94% of all shifts.

*Accuracy:* 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.

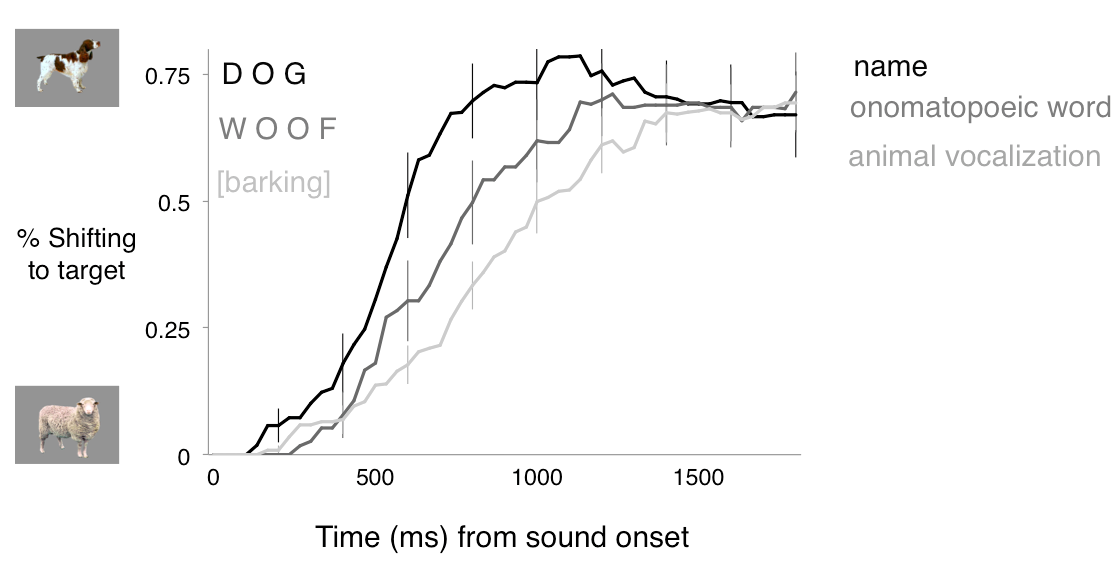
*Reaction time:* We calculated reaction time (RT) on those trials on which participants were looking at the distractor animal at the beginning of the sound. RT on each trial was the latency of the first shift to the correct animal within a 300- to 1,800-ms window from sound onset, as typically done in studies using this procedure ([Fernald et al., 1998](#_ENREF_21)).

*Results and Discussion*

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

Our first question is whether a familiar animal name, onomatopoeic word, and animal vocalization is a particularly effective cue in guiding children's attention to an animal in their environment. We 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*).

Figure 1a 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).



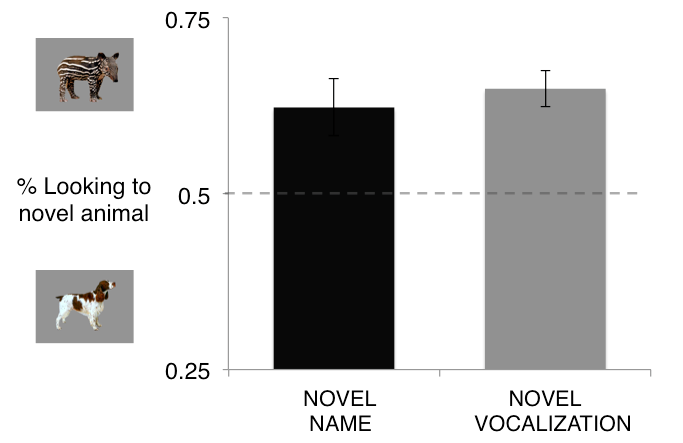
**Figure 1.** Time course of children’s looking to the target animal after hearing different familiar and novel auditory cues. The curves depict changes in the proportion of looking to the target animal as sounds unfolded, measured from sound onset (in milliseconds). When hearing a familiar auditory cue, children were faster to orient to the target animal after hearing its name, than when hearing the lexicalized sound or animal vocalization.

Crucially, children were able to use the associations between familiar animals and each of three different familiar auditory cues to identify the appropriate animal. When accuracy is computed 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.

*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.

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.



**Figure 1.** Accuracy of responses to novel auditory cues. 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.

**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 should not be conflated ([Bion, Borovsky, & Fernald, under review](#_ENREF_5); [J. S. Horst, B. McMurray, & L. K. Samuelson, 2006](#_ENREF_27); [B. McMurray, 2007](#_ENREF_36); [B. McMurray, Aslin, & Toscano, 2009](#_ENREF_37); [B. McMurray et al., unpublished manuscript](#_ENREF_38); [B. McMurray, Horst, Toscano, & Samuelson, 2009](#_ENREF_39)).

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. Using a looking-time task, Bion et al. (under review) replicated these findings in a study with 18-, 24-, and 30-month-old infants.

Experiment 2 asks whether children can retain the link created through disambiguation between a novel animal and a novel animal vocalization.. In addition, we aim at replicating the findings from Experiment 1, showing that children can identify familiar animals based on the vocalizations they produce, and use novel vocalizations to disambiguate novel animals.

*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.

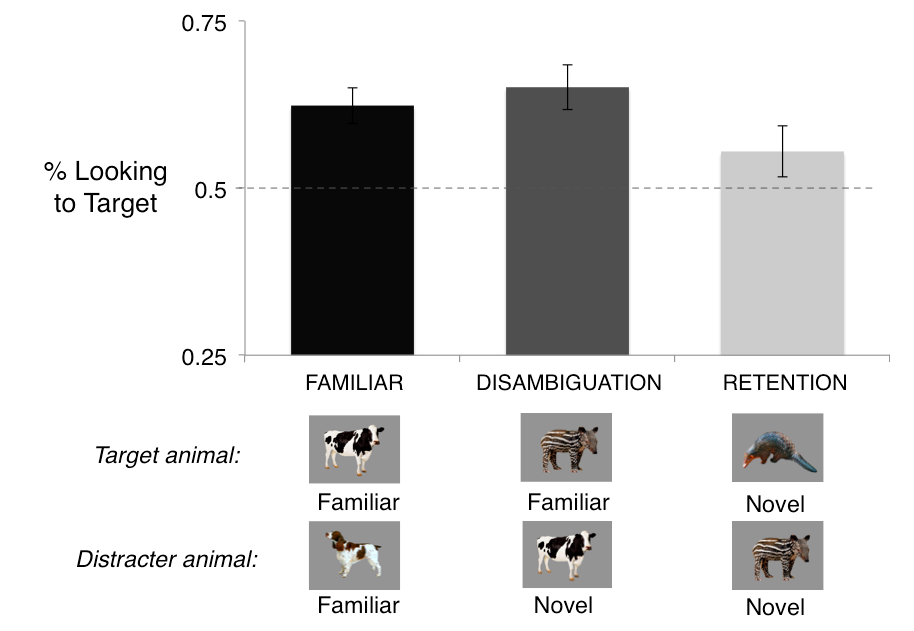
*Books.* As in Experiment 1, we sent home a children’s book to ensure that all participants had at least some exposure to the familiar animals and auditory cues. Since, in Experiment 2, we were interested in the natural animal vocalizations and not the names/lexical sounds, only the Hear and ThereTM *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:*

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 3.** 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 were marginally successful in retaining 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 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.

*General Discussion.*

Three main findings emerged in this research. The first finding was that 30-month-olds responded fastest to the familiar animal name and slowest to the familiar animal vocalization, with onomatopoeic sounds somewhere in between. Yet, children could identify the associations between familiar animals and their characteristic vocalizations, onomatopoeic sounds, and names. The second finding was that children showed one-to-one biases for the types of vocalizations that animals produce, similar to their biases in word learning. The third finding was that these one-to-one biases do not necessarily lead to learning, as children were only marginally successful in retaining the link between novel animals and their vocalizations ([see B. McMurray, Horst, et al., 2009](#_ENREF_39)).

Children reliably looked at a familiar animal when hearing a familiar animal vocalization, but they were fastest to guide their attention to a familiar animal when hearing the animal name. Previous studies found no differences between environmental sounds and words, advantages for words, and advantages for environmental sounds, depending on the population and paradigm being used. Studies with normally developing children also found comparable accuracy for the processing of familiar environmental sounds and words ([Cummings et al., 2009](#_ENREF_15)). In contrast, studies with language impaired and autistic children suggest an advantage for the processing of environmental sounds in these populations ([Cummings & Čeponienė, 2010](#_ENREF_13); [McCleery et al., 2010](#_ENREF_35)), with similar benefits in crossmodal priming studies with adults ([Chen & Spence, 2011](#_ENREF_10)). It is therefore not clear whether environmental sounds and words are processed differently. Recent brain imaging studies suggest that largely overlapping networks process verbal and non-verbal meaningful semantic information ([A. Cummings et al., 2006](#_ENREF_14); [Dick et al., 2007](#_ENREF_18); [Liu & Holt](#_ENREF_29)). In contrast, dual processing models propose that environmental sounds are directly connected to semantic representations, but words are linked to semantic representations through the lexical system ([Chen & Spence, 2011](#_ENREF_10)).

In our study, we found an advantage for words over other meaningful sounds. One of the reasons for this advantage for words might be due to differences in the processing of words and nonverbal sounds. Some language development theories argue that words are special because they refer to objects in the world ([Waxman & Gelman, 2009](#_ENREF_56)), and other theories argue that words are special because they activate conceptual information more quickly and accurately and in a more categorical way than nonverbal sounds. And it is also possible that words and nonverbal sounds might also be processed by different brain regions, with words being more rapidly accessed. A second explanation for the advantage for words might be due to differences in frequency. At least in our particular sample, it is save to assume that children have heard the word cowmany more times than they have heard an actual cow mooing. Frequency effects have been robustly demonstrated in the processing of words, with adults being faster to recognize words that they hear more frequently ([Dahan, Magnuson, & Tanenhaus, 2001](#_ENREF_16)). And third, words are very effective at presenting a lot of information in a short period of time. When a dog and a sheep are pairedwith each other, the first phoneme of the words is enough to determine the animal that is likely to be talked about next. In contrast, the animal vocalizations in our study were longer, and they might have later disambiguation points, something that could be determine in a future gating study.

Little is known about children’s and adults’ processing of onomatopoeic sounds. Hashimoto et al. ([Hashimoto et al., 2006](#_ENREF_25)) 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. Cummings et al. argues that onomatopoeic sounds might provide young children with information about intermodal associations, bridging their understanding of non-arbitrary environmental sounds and arbitrary word-object associations ([Cummings et al., 2009](#_ENREF_15)). Fernald and Morikawa ([1993](#_ENREF_20)) reported that 52% of Japanese mothers used onomatopoeic sounds to label target objects, while only 1 in 30 American mothers did so.

Our second 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 (i.e., lexical account), a product of children’s theory of mind restricted to communicative situations (i.e., pragmatic account), or a special case of a general principle of learning that exaggerates regularities across domains (i.e., 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](#_ENREF_7)).

Previous studies contrasted lexical-specific and pragmatic accounts, with little attention to domain-general explanations. Diesendruk and Markson found that children expect speakers to use consistent facts to refer to objects, and they select a novel object when hearing a novel fact ([Diesendruck & Markson, 2001](#_ENREF_19)). Yet, 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](#_ENREF_46)). Disambiguation biases for words are correlated with vocabulary, and disambiguation biases for facts are correlated with social-pragmatic skills ([de Marchena et al., 2011](#_ENREF_17)). As these authors acknowledge, these findings suggest that disambiguation biases for words might not be motivated uniquelly by pragmatic inferences, but they do not rule out domain-general accounts.

Few studies looked at disambiguation biases in non-linguistic domains. Three-year-olds expect faces to map to individual voices ([Moher, Feigenson, & Halberda, 2010](#_ENREF_41)), showing that one-to-one biases might extend to other communicative domains. In contrast, studies with adults found an advantage for words over non-linguistic stimuli in a task that could benefit from disambiguation biases ([Yoshida, Rhemtulla, & Vouloumanos, in press](#_ENREF_60)). The findings with adults do not rule out a domain-general account, since participants had no reason to expect that the mapping between random non-linguistic sounds 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 to demonstrate that young children show disambiguation biases in a nonlinguistic and non-communicative domain. This is also the youngest age at which disambiguation biases were shown in a domain other than word learning. These findings seem to favor domain-general accounts that see disambiguation biases as the natural consequence of a system that attempts to find regularities in complex learning tasks that involve consistent mappings. Previous connectionist and Bayesian models of word learning showed that disambiguation biases emerge as children are exposed to consistent mappings between words and objects, without the need of built-in constraints on the meaning of words ([Frank et al., 2009](#_ENREF_22); [B. McMurray, Horst, et al., 2009](#_ENREF_39); [Regier, 2003](#_ENREF_47); [Yu & Smith, 2007](#_ENREF_61)). In principle, these same biases would emerge if these models attempted to map animal vocalizations to animals and consistent co-occurrences were present in the environment.

The question that remains is whether the finding from our studies and from previous studies could be explained by a single disambiguation mechanism. We showed that children can disambiguate stimuli other than words and facts, suggesting at least the existence of a domain general mechanism that leads to disambiguation. A preference for parsimony provides some week evidence in favor of a single mechanism. But as pointed out by recent computational work, it is possible that different mechanisms jointly contribute to disambiguation behavior, explaining findings across different populations and contexts ([Lewis & Frank, under review](#_ENREF_28)). It is possible therefore that the same behavior - selecting a novel object when hearing a novel auditory stimulus - might result from different computational mechanisms or motivations depending on the task at hand, children's age, or the particular stimuli set and assumptions about the people involved in the interaction.

Our third finding was that one-to-one biases for animal vocalizations do not necessarily lead to retention of the link between a novel animal and a novel vocalization. Importantly, this finding supports the prediction of recent cross-situational models of early word learning ([J. S. Horst et al., 2006](#_ENREF_27); [B. McMurray, Aslin, et al., 2009](#_ENREF_37)). McMurray and colleagues propose that referent-selection requires that children give their best guess in a specific ambiguous situation, but 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. These claims are corroborated by studies on early word learning using online and offline measures of retention (Horst, et al., 2006, Bion et al., 2012).

The findings of this study add to a recent body of work that encourages us to think differently about disambiguation biases. These studies have emphasized the role of experience in the emergence of disambiguation biases, showing that the tendency to select a novel object when hearing a novel word is not robustly present across populations ([Byers Heinlein, 2010](#_ENREF_9); [Yurovsky, Bion, Smith, & Fernald, 2012](#_ENREF_62)). For example, bilingual children, children from lower socioeconomic status, children who receive less language input, and children with less structured vocabularies or smaller vocabulary sizes, take longer to show evidence of disambiguation biases ([R. A. Bion, A. Borovsky, & A. Fernald, 2012](#_ENREF_6); [Weisleder, Hurtado, & Fernald, in preparation](#_ENREF_57); [Yurovsky et al., 2012](#_ENREF_62)). Other studies have problematized the relation between disambiguation biases and word learning, showing that success in referent selection does not necessarily mean that the link between the novel word and novel object will be retained ([R. Bion, A. Borovsky, & A. Fernald, 2012](#_ENREF_4); [J. Horst, B. McMurray, & L. Samuelson, 2006](#_ENREF_26); [B. McMurray, Aslin, et al., 2009](#_ENREF_37); [Bob McMurray, Horst, & Samuelson, 2012](#_ENREF_40)). And our study adds to studies taking a fresh look at an old question: the scope of disambiguation biases ([Moher et al., 2010](#_ENREF_41); [Suanda & Namy, 2012](#_ENREF_50), [2013](#_ENREF_51)). Taken together, these recent studies challenge some widespread assumptions about the emergence, importance, and scope of disambiguation biases.

Children’s learning about objects in their environment involves more than learning their names. Before object names are learned, sounds and actions might form the basis on which objects are conceptualized ([Nelson, 1973](#_ENREF_44)). For example, children might see barking as a defining feature of dogs, and may 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 the object’s texture, smell, as well as its sounds and names. Children do not have explicit constraints that freshly baked cookies should have only one smell. Yet, they might recognize and get excited about the familiar smell coming from the kitchen, and might assume their mothers are baking something new when smelling something unfamiliar.

Children use different types of knowledge in order to make sense of a constantly changing world. They might identify animal vocalizations based on the shape of the vocal tract of the animal, its location and size, and their previous knowledge about animal vocalizations. Importantly, these cues normally converge in helping children identify an animal in the environment. The same is true for their identification of referents for words. Children can identify the referent for a word based on semantics ([Goodman, McDonough, & Brown, 1998](#_ENREF_24)), cross-situational statistics ([Goodman et al., 1998](#_ENREF_24)), syntax ([Brown, 1957](#_ENREF_8)), and pragmatic and social cues ([Baldwin, 1991](#_ENREF_1), [1993](#_ENREF_2)), and – why not - disambiguation biases ([Markman, 1991](#_ENREF_31)). As children grow older, these different sources of information provide converging evidence that a novel word should refer to a novel object. Children can rely on their knowledge about the world, speakers, and on their previous experiences with words in order to figure out what speakers are talking about – a task we continue to do throughout our lives when 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.

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. *The crosslinguistic study of sentence processing*, 3–73.

Bion, R., Borovsky, A., & Fernald, A. (2012). Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*.

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.

Bion, Ricardo AH, Borovsky, Arielle, & Fernald, Anne. (2012). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30months. *Cognition*.

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.

Byers Heinlein, K. (2010). *Bilingualism in infancy: a window on language acquisition.* (PhD), University of British Columbia.

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

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

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, AP, 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.

Dahan, Delphine, Magnuson, James S, & Tanenhaus, Michael K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology, 42*(4), 317-367.

de Marchena, A, Eigsti, IM, Worek, A, Ono, KE, & 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., . . . Pizzamiglio, L. (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.

Fernald, A., Pinto, J., Swingley, D., Weinberg, A., & McRoberts, G. (1998). Rapid gains in speed of verbal processing by infants in the second year. *Psychological Science, 9*, 72-75.

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

Fulkerson, Anne L, & Waxman, Sandra 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., McMurray, B., & Samuelson, LK. (2006). Online processing is essential for learning: Understanding fast mapping and word learning in a dynamic connectionist architecture. *Proc. of CogSci’06*.

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*.

Lewis, M., & Frank, M. (under review). Modeling disambiguation in word learning via multiple probabilistic constraints.

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.

Lupyan, Gary, & Thompson-Schill, Sharon L. (2012). The evocative power of words: Activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology-General, 141*(1), 170.

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

Markman, E., Wasow, JL, & Hansen, MB. (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.

McMurray, Bob, Horst, Jessica S, & Samuelson, Larissa K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review, 119*(4).

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.

Murray, Micah M, Camen, Christian, Andino, Sara L Gonzalez, Bovet, Pierre, & Clarke, Stephanie. (2006). Rapid brain discrimination of sounds of objects. *The Journal of neuroscience, 26*(4), 1293-1302.

Namy, Laura L, & Waxman, Sandra R. (1998). Words and gestures: Infants' interpretations of different forms of symbolic reference. *Child Development, 69*(2), 295-308.

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

Peña, M, Maki, A, Kovac i , D, Dehaene-Lambertz, G, Koizumi, H, Bouquet, F, & Mehler, J. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences, 100*(20), 11702.

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*.

Suanda, Sumarga H, & Namy, Laura L. (2012). The Organization of Words and Symbolic Gestures in 18‐Month‐Olds’ Lexicons: Evidence From a Disambiguation Task. *Infancy*.

Suanda, Sumarga H, & Namy, Laura L. (2013). Young word learners’ interpretations of words and symbolic gestures within the context of ambiguous reference. *Child Development, 84*(1), 143-153.

Vouloumanos, A, & Werker, JF. (2004). Tuned to the signal: the privileged status of speech for young infants. *Developmental Science, 7*(3), 270-276.

Vouloumanos, A, & Werker, JF. (2007a). Listening to language at birth: evidence for a bias for speech in neonates. *Developmental Science, 10*(2), 159-171.

Vouloumanos, A, & Werker, JF. (2007b). Why voice melody alone cannot explain neonates' preference for speech. *Developmental Science, 10*(2), 169.

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.

Waxman, Sandra R, & Gelman, Susan A. (2009). Early word-learning entails reference, not merely associations. *TRENDS in Cognitive Sciences, 13*(6), 258-263.

Weisleder, A., Hurtado, N., & Fernald, A. (in preparation). Early language experience predicts use of 'mutual exclusivity' bias in low-SES Latino children.

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

Xu, Fei. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition, 85*(3), 223-250.

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.

Yurovsky, Daniel, Bion, Ricardo AH, Smith, Linda B, & Fernald, Anne. (2012). *Mutual exclusivity and vocabulary structure.* Paper presented at the Proceedings of the 34th Annual Conference of the Cognitive Science Society.