**Disambiguation biases in a non-linguistic and non-communicative domain:**

**Young children map novel animal vocalizations to unfamiliar animals**

Authors:

Ricardo A. H. Bion

Anne Fernald

Affiliations:

Stanford University - Department of Psychology,

450 Serra Mall, Stanford, CA 94305

Corresponding author:

Ricardo A. H. Bion

Stanford University - Department of Psychology,

450 Serra Mall, Stanford, CA 94305

E-mail: ricardoh@stanford.edu

Phone: (650) 704-9141

**Abstract**

When hearing a novel name, children tend to select a novel object rather than a familiar one, a bias known as disambiguation. This bias is often assumed to reflect children’s expectations about the nature of words or expectations about the communicative intention of speakers. This study investigated whether similar biases emerge in a domain that is non-linguistic and non-communicative for children, but in which strong regularities can be found: the vocalizations that animals produce. In Experiment 1, we used online processing measures to compare 32-month-olds' processing of animal vocalization and animal names. When hearing a familiar animal vocalization or name, children looked at the correct animal, but they were faster to identify the animal when hearing its name than its vocalization. When hearing a novel animal vocalization or name, children looked at a novel animal rather than at a familiar one, being equally fast when hearing the novel name or novel vocalization. In Experiment 2, we demonstrate that this bias to look at novel animals when hearing novel vocalizations can be used by children to learn about new animals and their vocalizations. This is the earliest age at which disambiguation biases have been observed in a domain other than word learning. We characterize disambiguation biases as resulting from domain-general processing and learning mechanisms, rather than from lexical or communicative constraints.

Keywords: disambiguation, mutual exclusivity, environmental sounds, retention, word learning.

**Disambiguation biases in a non-linguistic and non-communicative domain:**

**Young children map novel animal vocalizations to unfamiliar animals**

When children encounter people and animals in daily life, they experience them through multiple sensory modalities simultaneously. When playing with pets, for example, the child learns to associate the physical features and actions of dogs and cows with the barking or mooing sounds typically produced by these animals. Eventually, the child will also link specific spoken words to each type of animal, learning that they are associated with the object names *dog* or *cow,* as well as withonomatopoeic words resembling their characteristic vocalizations, such as *woof-woof* or *mooo.* Thus, just as familiar object names, including animal names, are consistently associated with particular kinds of animate and inanimate objects, animal vocalizations and onomatopoeic words for vocalizations also provide consistent associations between auditory stimuli and different types of animal. In this study, we compare young children’s efficiency in using these different classes of sounds as cues to identifying particular animals. In addition, we investigate whether children can use disambiguation strategies, which have frequently been characterized as pragmatic or lexically-specific in nature (Bloom, 2002; Diesendruck & Markson, 2001; Markman, 1991), in order to infer which of two animals is associated with a novel vocalization.

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

Other research has approached the question of whether speech is special from a different perspective, comparing how children process spoken words as compared to non-arbitraryenvironmental sounds, such as animal vocalizations or the sounds produced by inanimate objects (e.g., cat meowing, car starting). Studies with adults found similarities and differences in both behavioral and neural responses to cross-modal semantic associations between words and environmental sounds. In a picture detection task, Chen and Spence (2011) found priming from environmental sounds but not from words. These authors propose that recognition of environmental sounds is faster because words must also be processed at a lexical stage, 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) found advantages for words as compared to environmental sounds. This finding was interpreted 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 with adults, Cummings, Čeponienė, Koyama, Saygin, Townsend, and Dick (2006) 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 sound types that varied in arbitrariness, Hashimoto et al. (2006) found different neural mechanisms for the processing of animal names and vocalizations, with onomatopoeic words activating both areas. Because research on environmental sounds is still just beginning, it is hard to reconcile these somewhat discrepant findings. Variations in tasks or timing of stimuli can influence results, 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, bell ringing) despite evidence that these sounds are treated differently by the adult brain (Murray, Camen, Andino, Bovet, & Clarke, 2006). But this line of research provides promising new ways to examine the perennial question of whether language emerges from the interaction of domain-general cognitive processes or domain-specific mechanism (Bates & MacWhinney, 1989). For example, recent research comparing processing of speech and non-speech sounds is leading to new insights relevant to autism, developmental language impairment, and cochlear implants (Cummings & Čeponienė, 2010; McCleery, Čeponienė, Burner, Townsend, Kinnear, Schreibman, 2010).

From a developmental perspective, it is also important to understand how children process words and non-arbitrary nonlinguistic sounds. Very few studies have looked at this question. In a study using a preferential-looking paradigm, Cummings, Saygin, Bates, and Dick (2009) found that 15- and 25-mo-olds can use words and environmental sounds to guide their attention to familiar objects, improving as they get older. Vouloumanos, Druhen, Hauser, & Huizink (2009) found that 5-month-olds can match some animals to the vocalizations they produce. And studies with children with autism and developmental language impairment found more severe deficits for the processing of words than environmental sounds (McCleery et al., 2010; Cummings & Čeponienė, 2010).

We build on these earlier studies using the looking-while-listening paradigm, which has been widely used to assess real-time interpretation of spoken words by infants and young children (Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998; Fernald, Zangl, Portillo, & Marchman, 2008). One major goal in this research is to investigate how two-year-old children process different types of auditory stimuli that are consistently associated with familiar animals, but that vary in level of arbitrariness. First we ask whether 32-month-olds can use onomatopoeic sounds (e.g. *bow-wow*), and animal vocalizations (e.g., dog barking) to identify familiar animals, as well as familiar animal names (e.g., *dog*). By using real-time processing measures, we can also determine whether these three sounds are equally effective as acoustic cues in guiding children's attention to a particular animal in the visual scene. The use of looking to visual stimuli, rather than object-choice responses, reduce the task demands of procedures requiring more complex responses such as reaching or pointing, and yield continuous rather than categorical measures of attention on every trial, capturing differences in processing that might not be detected by offline tasks that rely on categorical responses (Fernald et al., 1998; Bion, Borovsky, and Fernald, 2013).

A second major goal in this research is to investigate how young children learn to link nonlinguistic sounds to animate objects. Do two-year-olds make similar inferences when mapping a novel word and a novel vocalization to an unfamiliar animal? Typically, word learning is portrayed as an intractable challenge, while associating animals with the sounds they produce might appear trivial. The acoustic structure of vocalizations is influenced by the size and shape of the vocal tract and other physical features, linking sounds to their source in a non-arbitrary way. And the fact that many animal vocalizations are accompanied by synchronous physical movements might provide children with additional non-arbitrary cues to the source of the sound. Even in the absence of additional visual cues, it is often possible to pinpoint the source of a sound with reasonable accuracy. In contrast, because the acoustic structure of a word is in most cases arbitrary with respect to potential referents, and it is produced by a speaker and not by the object itself, learning to associate speech sounds with objects is often characterized as a complex problem of induction (Markman, 1991).

To solve the word-learning puzzle, children are said to be equipped with constraints on the possible meanings of words. The most widely studied of these constraints rules that each object must have only one name (Markman, 1991). Evidence for this default assumption 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. Debate about the origins, scope, and generality of the Mutual Exclusivity (ME) constraint has focused on whether this response bias provides evidence for a lexical constraint, or whether it results instead from inferences about speakers’ communicative intent. Lexical accounts characterize ME as a “domain specific mechanism specific to word leaning” (de Marchena, Eigsti, Worek, Ono, & Snedeker, 2011), which “predicts disambiguation only within the domain of word learning (i.e., it is domain-specific)” (Scofield & Behrend, 2003). Pragmatic accounts propose that ME extends to communicative acts more broadly, reflecting assumptions that speakers are cooperative and should use conventional names to refer to familiar objects (Bloom, 2002; Clark, 1990). A third possibility is that the bias toward one-to-one mappings observed in word learning and other communicative domains reflect general biases to find simple regularities in complex domains, a perspective embraced by recent computational approaches to word learning (Frank, Goodman, & Tenenbaum, 2009; McMurray, Horst, & Samuelson, 2013; Regier et al., 2001).

Thus lexical and pragmatic accounts of the scope of the ME constraint assert that one-to-one biases are either unique to word learning or that they generalize to communicative acts more broadly, while domain-general accounts predict that they would apply to any domain in which consistent one-to-one mappings are observed.

To explore the possibility that one-to-one biases in sound-object mappings are not limited to interpreting communicative acts, we investigated whether children would show responses comparable to the ‘mutual exclusivity’ bias in a domain that is neither linguistic nor communicative for them, but in which consistent associations are observed between objects and auditory cues. The question of interest was whether children would show one-to-one biases in linking novel, non-speech vocalizations to unfamiliar animals, similar to their biases in word learning contexts. When presented with paired pictures of a familiar animal (e.g., dog) and an unfamiliar animal (e.g., aardvark) in a disambiguation task, will 2-year-olds orient to the novel animal not only when they hear a novel animal name, but also when they hear a novel animal vocalization?

**Experiment 1**

This first experiment asks two questions. First, we ask whether 32-month-olds can use familiar animal names (e.g., *dog*), onomatopoeic words (e.g. *bow-wow*), and animal vocalizations (e.g., dog barking) to identify familiar animals. These three sound types differ in arbitrariness within a continuum, with speech as the most arbitrary and vocalizations as the least arbitrary sound. We ask whether they are equally effective as acoustic cues in guiding children's attention to animals in a visual scene. Children will hear a familiar animal name, an onomatopoeic word, or an animal vocalization, while looking at the picture of two familiar animals, one that matches and one that does not match the sound stimuli. We compare children's proportion of looking to the matching animal when hearing these different sound types. One of three patterns of result is most likely to emerge: The first is that children are faster to identify animal names than onomatopoeic words, and faster to identify onomatopoeic sounds than animal vocalizations. This pattern of results could be predicted by computational models that see frequency as a crucial determinant of speed of processing (e.g., McMurray et al, 2009) either at the token level (children in urban environments are likely to hear the names of animals more often than their vocalizations) or at a higher level resulting in more practice interpreting speech (high SES children might hear thousands of words daily, but not nearly as many animal sounds). It could also be predicted by developmental accounts that see speech as having a privileged status in early development either in getting children's attention (Vouloumanos & Werker, 2004, 2007a, 2007b), or due to the fact that it refer to objects (Waxman et al., 2009). The second pattern of result is that children are faster to identify animal vocalizations than onomatopoeic sounds, and faster to identify onomatopoeic sounds than animal names. This pattern of results could be predicted by accounts that propose that non-arbitrary sounds link directly to semantic representations, while words first connect to lexical representations before reaching semantics (Chen & Spence, 2011). The fact that children have experience interpreting environmental sounds (balls bouncing, things falling) before learning to interpret speech referentially could also predict an advantage for environmental sounds. A third possibility is that children are equally efficient in exploiting these three sound types to guide their attention to a familiar animal. This pattern of results would parallel that of previous studies that failed to find differences in the processing of environmental sounds and words (Cummings et al., 2009). These pattern of findings would be the most difficult to interpret, but would demonstrate children's proficiency in exploiting very different sounds to guide their attention to objects in their surroundings.

Our second question is whether two-year-olds make similar inferences when mapping a novel name and a novel animal vocalization to an unfamiliar animal. In a paradigm similar to the one explained above, children will hear a *novel* animal name or novel animal vocalization (instead of a familiar one), while looking at the picture of a familiar and a *novel* animal (instead of two familiar objects). We compare children's proportion of looking to the novel animal when hearing a novel name and novel vocalization. Considering dozens of studies on children's disambiguation biases, it is safe to assume that children will look at a novel animal when hearing a novel name, thus one of two patterns of results is most likely to emerge: The first is that children look at a novel animal when hearing a novel name, but are at chance or substantially less accurate when hearing a novel animal vocalization. This pattern of result would be compatible with lexical accounts that predict disambiguation only within the domain of word learning, or by pragmatic accounts that predict disambiguation only within communicative contexts. The second pattern of findings is that children look at a novel animal when hearing a novel name and when hearing a novel animal vocalization, with performance indistinguishable or comparable between these two conditions. This pattern of results would be compatible with accounts that propose that disambiguation biases emerge from domain-general learning mechanisms that look for regularities in complex domains.

*Method*

*Participants.* Participants were 21 32-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. Children were from mid/high-SES families, with average maternal education of 16 years (i.e., college degree, range=12-18 years of education).

*Visual stimuli.* The visual stimuli included pictures of four Familiar animals (horse, dog, cow, sheep) and two Novel animals (pangolin, tapir). According to parental report, the familiar animals were known by all children. Parents also reported that the novel animals were completely unfamiliar to the children. Each animal picture was centered on a grey background in a 640 x 480 pixel space.

*Auditory stimuli.* The auditory stimuli consisted of sounds that were either Familiar or Novel to 32-month-olds. Figure 1 serves as a guide to the different sound types. The *Familiar* sounds were used in *Familiar Trials*, and consisted of one of three different *familiar* sounds: *names* (horse, dog, cow and sheep)*,* *onomatopoeic words* (neigh, woof-woof, moo and baa)*,* and *vocalizations* (horse neighing, dog barking, cow mooing and sheep baaing). The *Novel* sounds were used in *Disambiguation Trials*, and consisted of one of two types of *novel* sounds: *names* (capa, nadu)and *vocalizations* (rhino grunting, gorilla snorting).

Trials in which the auditory cue was a familiar or novel animal name (e.g., *Where’s the dog?*)or a familiar or novel lexical sound (*Which one goes woof-woof?)* began with a brief carrier frame. The duration of the target cue was 800 ms for lexical sounds and 750 ms for animal names. The intensity of the phrases was normalized using Praat speech analysis software (Boersma, 2002).

Trials with familiar or novel animal vocalizations began with a single word, used to draw children’s attention (e.g., *Look! “dog barking*”). Familiar animal vocalizations were selected based on prototypicality. After selecting at least three vocalizations for each familiar animal, the authors voted on the one that we thought would be most easily recognized by children. Choosing the novel animal vocalizations was more challenging. A group of research assistants pooled from different websites several vocalizations that they judged as unfamiliar. From these vocalizations, we selected two (i.e., rhino grunting and gorilla snorting) that we judged were equally likely to be produced by the 6 familiar and 2 novel animals based on the their size and vocal tract characteristics. These vocalizations were also maximally distinct from each other and from the familiar animal vocalizations and expected to be unfamiliar to children. We counterbalanced the vocalizations that were paired with the two novel animals, in order to control for the possibility that children judged one of the two novel animals as more likely to produce one of the novel vocalization. All children were reported by parents to have had no exposure to either of the novel animal’s natural vocalization. The duration of all target animal vocalizations was 2000 ms. More details about trial types and conditions in Figure 1 will be given in the Procedure section.



**Figure 1.** Trial types in Experiments 1 organized by type of cue: Familiar vs. Novel. The target animal for each trial type is on the left.

*Familiarization.* There were two main reasons for us to want to make sure that children knew the familiar onomatopoeic words and animal vocalizations before we administered our experiment. First, we wanted to make sure that any differences we observed in children's performance within Familiar-Animal Trials was due to processing speed and could not be explained by the fact that children were not familiar with one of the sound types. The looking-while-listening procedure has been shown to capture differences in processing efficiency even when words are considered "known" by offline reaching tasks or parental report (Fernald et al., 2008). Second, we wanted to make sure that children knew the pairings between familiar vocalizations and animals, a potential prerequisite for success on Disambiguation trials. Since we were working with children from mid/high-SES families growing up in an urban environment, we were particularly worried that they would not be familiar with many animal vocalizations or onomatopoeic words. 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,* a week before their visit*.* Parents were instructed to share each book with their child for 5 to 10 min on at least three days prior to the experiment. The first book 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.* Since we were interested in detecting differences in processing between sounds that we expected to be familiar to children, we choose to access speed and accuracy in identifying the correct target picture with the looking-while-listening (LWL) procedure (see Fernald, et al, 2008). Previous studies have shown that even when objects are reported by parents as familiar to their children, or when children are at ceiling in offline reaching tasks, these real-time processing measures can capture meaningful differences in processing. These differences correlate to properties of the sound stimuli (e.g., word-frequency) and different aspects of the child's experience (e.g., their age, socioeconomic status, amount of parental talk). Looking-time measures have also been used in Disambiguation tasks with children from different ages, capturing differences in accuracy that relate to children's age and vocabulary size (Bion et al., 2013).

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. As seen in Figure 1, we have two main trials types, Familiar Trials and Disambiguation trials, paralleling our original two research questions on children's processing of familiar and novel auditory cues.

Within the *Familiar Trials*, we have three different sub-trials: *name, onomatopoeic word*, and *vocalization.* On 8 *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 *Onomatopoeic-word* trials, each familiar animal served as the target twice. On 16 *Vocalization* trials, each familiar animal served as the target four times. These three familiar sound types show allow us to answer our first research question, asking whether names, onomatopoeic words, or animal vocalizations, are equally effective as acoustic cues in guiding children's attention to animals in a visual scene.

Within the *Disambiguation Trials*, we have two different sub-trials: *name,* and *vocalization.* On 6 *Name* trials, each novel animal was labeled three times with a novel animal name (i.e., *capa, nadu*), always paired with a familiar animal. On 8 *Vocalization* trials, each novel animal vocalization served as the target four times and was paired with each familiar animal once. These two sound types should allow us to answer our second research question, asking whether two-year-olds make similar inferences when mapping a novel name and a novel animal vocalization to an unfamiliar animal.

These different trial types were administered in two different visits. The Familiar and Disambiguation Trials with animal names and onomatopoeic words were administered during their first. The Familiar and Disambiguation Trials with the animal vocalizations were administered during the second visit. We administered the animal vocalizations on the second visit to allow children to become familiar with the procedure and to give parents additional time to use the familiarization books with the vocalizations with their children. During each visit, five 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 animals, were counterbalanced across participants. Caregivers wore darkened sunglasses so that they could not see the pictures and influence infants’ looking throughout the 5-min procedure.

*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. Similar results are found when using a shorter analysis window.

*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).

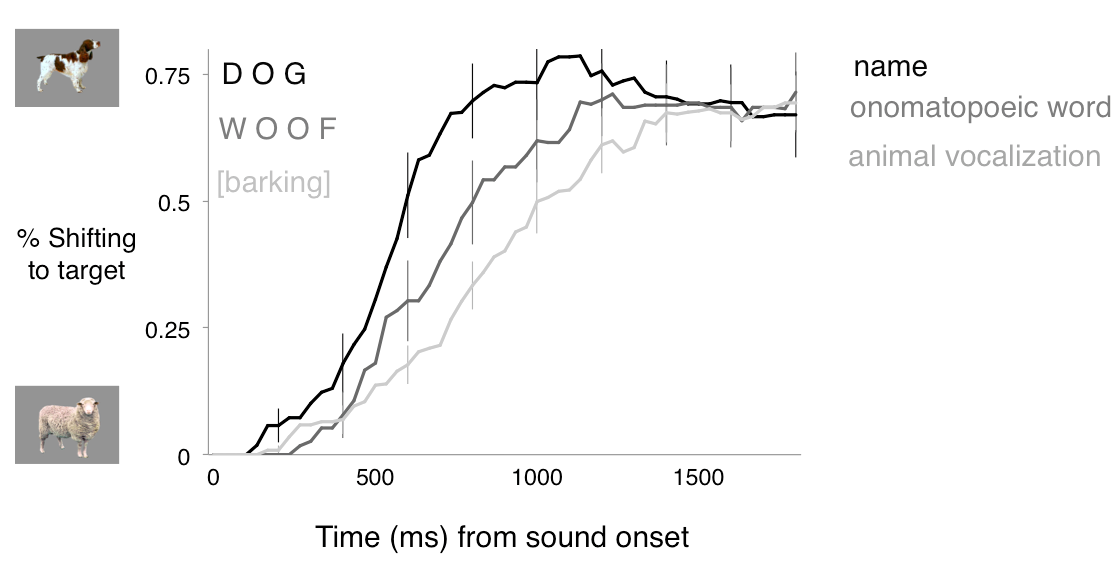
*Results and Discussion*

*Familiar Trials: 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 words (e.g., *bow-wow*), and their natural vocalizations (e.g., *dog barking*).

To begin exploring this question, Figure 2 shows children's looking behavior on the LWL procedure (Fernald et al., 1998). In order to capture children's speed of processing, we only show children's responses on trials in which they start looking at the wrong animal. From sound onset onward, we show the mean proportion of trials in which children were looking at the correct picture, every 33ms, with different lines representing children's responses on *Name* trials (black), *Onomatopoeic-word* trials (dark grey), and *Vocalization* trials (light grey). The y-axis shows the mean proportion of trials on which children are looking at the correct animal. The x-axis represents time from sound onset in milliseconds. As can be seen in Figure 1, at 750ms from sound onset, there is already a substantial difference in the mean proportion of trials in which children are looking at the correct animal depending on whether they heard an animal name or an animal vocalization. Children are looking at the correct animal in a greater proportion of trials when they hear a familiar animal name, as compared to when they hear an animal vocalization, with performance on trials with onomatopoeic words somewhere in between. This difference is robust and continues to increase as the trial unfolds.

To assess these differences statistically, we computed children's reaction time (RT) as the mean time it took them to shift to the correct picture on trials in which they were looking at the wrong picture at sound onset, again separately for the three trial types. Based on this measure, 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). This difference is robust, and not dependent of analysis window or rejection of participants.



**Figure 2.** Time course of children’s looking to the target animal after hearing a familiar animal name, onomatopoeic word, or animal vocalization. The graph shows children's responses for the trials in which they were looking at the distracter animal at sound onset. The x-axis shows time in milliseconds from sound onset. The y-axis shows the mean proportion of trials in which children are looking to the target animal. When hearing a familiar auditory cue, children were faster to orient to the correct animal after hearing its name, than when hearing the lexicalized sound or animal vocalization.

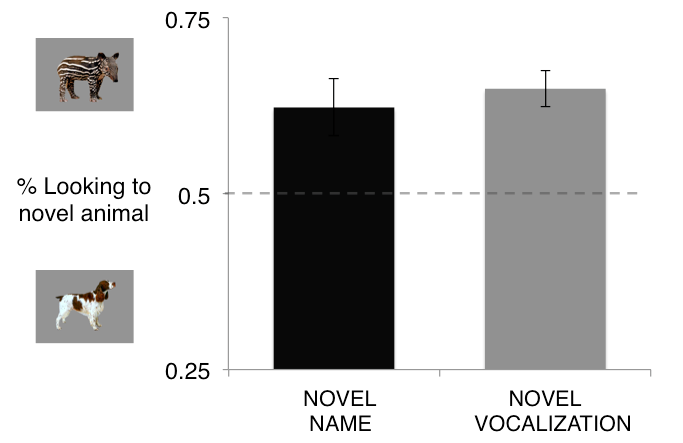
We found a difference in speed of processing for names, onomatopoeic words, and vocalization. Our next step was to confirm that children could use these three trial types to guide their attention to a familiar animal. When accuracy is computed over a window from 300 to 4300 ms after the onset of the 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). Despite differences in RT over a short window in trials in which children start looking at the wrong animal, these results show that the animal vocalization can be as good a cue to identify familiar animals as their names or lexicalized sounds if one gives children enough time to process the stimuli.

Revisiting our first research question and hypotheses, we asked whether 32-month-olds can use familiar animal names (e.g., *dog*), onomatopoeic sounds (e.g. *bow-wow*), and animal vocalizations (e.g., dog barking) to identify familiar animals, and whether these three sounds were equally effective in guiding children's attention to animals in a visual scene. We found that children were faster to orient to a target animal after hearing its name than after hearing its vocalization, with onomatopoeic words somewhere in between. This finding could be explained based on differences in frequency (e.g., our participants probably heard the animal names more often than animal vocalizations), or by developmental accounts that see speech as having a privileged status from early development onwards. These and other possibilities are discussed in more detail in our General Discussion session.

*Disambiguation Trials: 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. 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 3 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. There were no significant differences between children's reaction time for animal names or vocalizations.



**Figure 3.** 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.

Revisiting our second research question and hypotheses, we asked whether two-year-olds make similar inferences when mapping a novel name and a novel animal vocalization to an unfamiliar animal. We found that children looked at a novel animal when hearing a novel animal name, and they also looked at a novel animal when hearing a novel animal vocalization. While not serving as evidence against lexical or pragmatic accounts of disambiguation, this finding is best explained by domain-general accounts that see disambiguation biases as emerging across domains in which consistent cross modal mappings are observed. These findings and how they relate to the different explanation of disambiguation biases will be discussed in mode detail in the General Discussion session.

**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, 2013; Horst, McMurray, & Samuelson, 2006).

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. (2013) replicated these findings in a study with 18-, 24-, and 30-month-old infants using looking time measures of performance.

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. Our prediction is that children will succeed in disambiguation trials, but will show marginal or no retention on subsequent disambiguation trials, paralleling the findings of earlier studies with linguistic stimuli (Bion et al., 2013).

*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. Children were from mid/high SES families, with an average of 17 years of maternal education (range = 14-18 years of education).

*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 (see animals in Figure 4). We decided to change the novel animals in order to confirm that our results were not restricted to the particular stimuli set in Experiment 1. All children were reported by parents to have had no exposure to the novel animals.

*Auditory stimuli.* The auditory stimuli consisted of the same familiar and novel animal vocalizations 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 30 trials, consisting of three trial types (Figure 4). The 16 *Familiar* trials and 8 *Disambiguation* trials were identical in structure to the *Vocalization* trials 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.

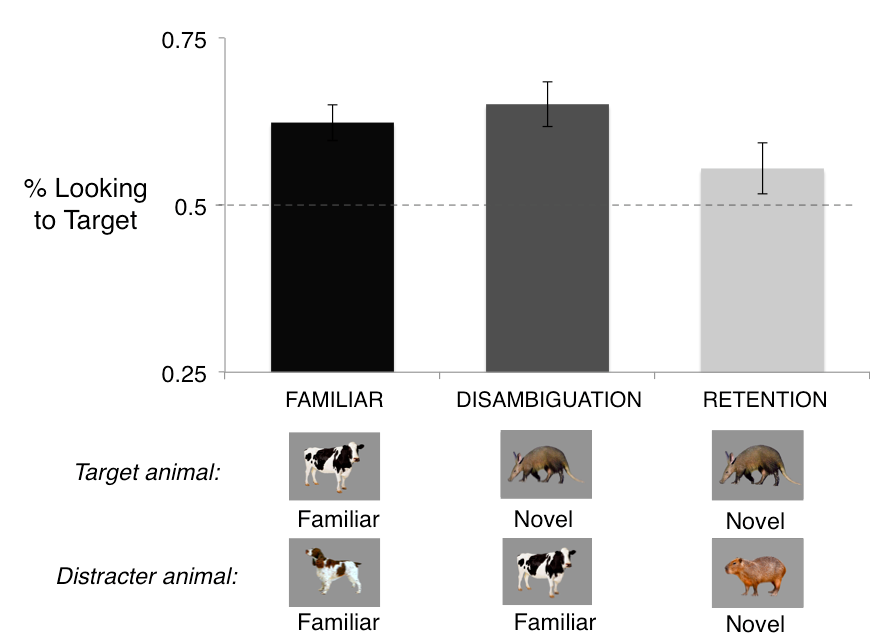


**Figure 4.** Trial types in Experiments 4 organized by type of trial. Children hear familiar and novel vocalizations. The target animal for each trial type is on the left.

*Results and Discussion:*

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

Figure 5 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 5.** 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).

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. These results and their implications will be discussed in more details in the following section.

*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 disambiguation biases for the types of vocalizations that animals produce, similar to their biases in word learning. The third finding was that these biases do not necessarily lead to learning, as children were only marginally successful in retaining the link between novel animals and their vocalizations, paralleling the findings of recent word-learning studies (see McMurray, Horst, et al., 2009).

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), 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 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). 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.

Little is known about children’s and adults’ processing of onomatopoeic 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. 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). Fernald and Morikawa (1993) 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).

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). 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). 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 uniquely 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), 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, 2012). 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; B. McMurray, Horst, et al., 2009; Regier, 2003; Yu & Smith, 2007). 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, 2013). 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. That is, children could use a domain general mechanism to learn about novel animal vocalizations, and they could use a lexical and pragmatic constraint to learn about novel words.

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 (Horst et al., 2006; B. McMurray, Aslin, et al., 2009). 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., 2013).

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; Yurovsky, Bion, Smith, & Fernald, 2012). 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 (Bion, Borovsky, & Fernald, 2013; Weisleder, Hurtado, & Fernald, in preparation; Yurovsky et al., 2012). 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 (Bion, Borovsky, & Fernald, 2013; Horst, McMurray, & Samuelson, 2006; McMurray, Horst, & Samuelson, 2012). And our study adds to studies taking a fresh look at an old question: the scope of disambiguation biases (Suanda & Namy, 2012, 2013). 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). 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), cross-situational statistics (Goodman et al., 1998), syntax (Brown, 1957), and pragmatic and social cues (Baldwin, 1991, 1993), and – why not - disambiguation biases (Markman, 1991). 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*, 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. (2013). Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, *126*, 39-53.

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

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

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

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

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

Cummings, A., & Čeponienė, R. (2010). Verbal and nonverbal semantic processing in children with developmental language impairment. *Neuropsychologia, 48*, 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*, 92-107.

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

Dahan, D., Magnuson, J., & Tanenhaus, M. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology, 42*, 317-367.

de Marchena, A., Eigsti, I., Worek, A., Ono, K., & Snedeker, J. (2011). Mutual Exclusivity in Autism Spectrum Disorders: Testing the Pragmatic Hypothesis. *Cognition*, *119*, 96-113.

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*, 799-816.

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

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development, 64*, 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*, 578.

Fulkerson, Anne L, & Waxman, Sandra R. (2007). Words (but not tones) facilitate object categorization: Evidence from 6-and 12-month-olds. *Cognition, 105*, 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*, 1330-1344.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition, 87*, 23-34.

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*, 1762-1770.

Lewis, M., & Frank, M. (2013). Modeling disambiguation in word learning via multiple probabilistic constraints. *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*

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

Lupyan, G., & Thompson-Schill, S. (2012). The evocative power of words: Activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology-General, 141*, 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*, 121-157.

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

McCleery, J., 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*, 277-286.

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

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

McMurray, B., Horst, J., & Samuelson, L. (2013). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological review*, *119*, 831-877

McMurray, B., Horst, J., Toscano, J., & Samuelson, L. (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*, 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*, 719-751.

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

Namy, L., & Waxman, S. (1998). Words and gestures: Infants' interpretations of different forms of symbolic reference. *Child Development, 69*, 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, 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., & 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.

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

Suanda, S., & Namy, L. (2012). The Organization of Words and Symbolic Gestures in 18‐Month‐Olds’ Lexicons: Evidence From a Disambiguation Task. *Infancy*, *18*, 276-288

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

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

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

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

Vouloumanos, A., Druhen, M., Hauser, M., & Huizink, A. (2009). Five-month-old infants' identification of the sources of vocalizations. *Proceedings of the National Academy of Sciences, 106*, 18867-18872.

Waxman, S., & Gelman, S. (2009). Early word-learning entails reference, not merely associations. *TRENDS in Cognitive Sciences, 13*, 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, A., & Hoyne, K. (1999). Infants' learning about words and sounds in relation to objects. *Child Development, 70*, 65-77.

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

Yoshida, K., Rhemtulla, M., & Vouloumanos, A. (2012). Exclusion constraints facilitate statistical word learnin. *Cognitive Science*. *36*, 933-947

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

Yurovsky, D., Bion, R., Smith, L., & Fernald, A. (2012). *Mutual exclusivity and vocabulary structure.* *Proceedings of the 34th Annual Conference of the Cognitive Science Society.*