

Understanding the Role of Non-Contrastive Variability
in Word Learning and Visual Attention in Infancy

By

HEIDI ANN BAUMGARTNER
B.A. (Stanford University) 2005
M.A. (University of California, Davis) 2010
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Psychology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Lisa M. Oakes, Chair

Katharine Graf Estes

Steven J. Luck

Committee in Charge

2014

Contents

| | |
|--|------------|
| Acknowledgements | iii |
| Abstract | iv |
| Chapter 1: Introduction | 1 |
| Chapter 2: The Role of Variability in Word Learning | 24 |
| <i>General Method</i> | 24 |
| <i>Experiment 1</i> | 43 |
| <i>Experiment 2</i> | 57 |
| <i>Experiment 3</i> | 70 |
| <i>Cross-Experiment Analyses and Summary</i> | 83 |
| Chapter 3: Visual Attention | 89 |
| <i>Experiment 4</i> | 89 |
| Chapter 4: General Discussion | 105 |
| References | 111 |
| Appendix | 122 |

Acknowledgements

This research was supported by funding from a University of California, Davis Humanities Graduate Research Award and a University of California, Davis Provost's Fellowship in the Arts, Humanities, and Social Sciences.

The work presented in this dissertation could not have been completed without the help and support of many people. I would first and foremost like to thank my graduate advisor and mentor, Lisa Oakes, for the unwavering support, guidance, and enthusiasm she provided throughout my graduate career. I am eternally grateful for all she has done for me.

I would also like to thank Steve Luck and Katie Graf Estes for their thoughtful consideration and advice throughout this process. Thanks also to Anne Fernald, Virginia Marchman, Renate Zangl, Kirsten Thorpe Read, and Allan Reiss for igniting and fostering my passion for developmental research.

Thank you to all of the members of the Infant Cognition Lab for their help with nearly every step along the way, especially Karinna Hurley for our office chats, lab managers Mirjam Harrison and Samantha Mitsven for their organizational magic, and coders Tina Vo, Javi Perez, and Sharmi Devarajan for their hours of careful work. Special thanks as well to all of the parents and infants who volunteered their time to make this project possible.

Thanks to my fantastic network of friends who have supported me via phone calls, gchats, and wine-fueled dinners, without whom I could not have made it through this process.

Lastly, to my amazing parents Bobbi and Lee Baumgartner and my brother Craig: thank you for the constant love and support, and for knowing I'd get through it even when I wasn't so sure. I know how lucky I am, and frankly, I couldn't have done it without you.

Abstract

Studies of word learning and comprehension in infancy often feature labeling events that feature a single exemplar of an object paired with a single label token, but the labeling events infants encounter in the real world are full of both acoustic and visual variability. Recent studies suggest that variability in phonemically non-contrastive features of speech (e.g., speaker voice, intonation) improves minimal pair learning at 14 months (Galle, Apfelbaum, & McMurray, 2014; Rost & McMurray, 2010). The present experiments investigate if the facilitative effect of acoustic variability in non-contrastive features extends to younger infants' learning of dissimilar words, and if the effect is present for variability in non-contrastive visual features of objects.

Twelve-month-old infants were habituated to novel object-label pairs featuring non-variable labels and objects (Experiment 1), variable labels paired with non-variable objects (Experiment 2), or non-variable labels paired with variable objects (Experiment 3). Following habituation, infants' learning of the novel associations, as well as their ability to recognize familiar words, were then tested in a *looking-while-listening* task. Only infants who habituated to variable object exemplars in Experiment 3 demonstrated word learning at test. Recognition of familiar words was enhanced following habituation to variable stimuli in Experiments 2 and 3, consistent with other findings that infants' ability to demonstrate linguistic knowledge is influenced by the broader context of the testing environment (e.g., Bergelson & Swingley, 2012; Fennell & Waxman, 2010; Fennell, 2012). Infants in Experiment 1 were also tested in a visual search task in which search items either did or did not vary in a task-irrelevant feature (Experiment 4). Infants who were better able to direct their attention to a shape-defined target in the more difficult, non-variable search condition showed more evidence of having learned the novel words in Experiment 1, providing evidence that visual attention and word learning are closely related.

Chapter 1: Introduction

Much research in the field of cognitive development attempts to answer the question of how infants make sense of the world around them. Infants must learn how to organize a large amount of information in their environment into coherent units in order to recognize objects, form categories, and learn words. Because there are a nearly infinite number of ways that input can be organized and the processing capabilities of infants are limited, infants must selectively attend to some pieces of information at the cost of neglecting other pieces of information. How do infants come to discriminate between meaningful information that is worthy of their attention and “noise” that should be ignored?

Many theories of development address how infants deal with the task of selecting and attending to meaningful information. For example, according to some theories, infants are endowed with innate, domain-specific learning mechanisms (“core cognition”) which provide the foundation for concept formation and perceptual input analysis (e.g., Carey, 2009). Alternatively, more empiricist theories of development argue that while there might be initial, domain-general learning mechanisms and constraints on attention (e.g., a bias to attend to certain aspects of input such as motion; Rakison, 2005), infants learn to recognize meaningful information through experience interacting with their environment. For example, statistical learning and connectionist accounts of associative learning provide evidence that domain-general learning mechanisms can account for the way that infants are able to successfully parse information into units and group those units into categories by attending to and processing regularities in their environment (Rogers & McClelland, 2004; Saffran, Aslin, & Newport, 1996).

The present experiments investigate how infants make use of regularities and variability in the environment in order to guide their attention and facilitate learning. The effect of variability on attention and learning in infancy is examined in the contexts of word learning (Chapter 2) and visual attention (Chapter 3).

Learning of Associations in Infancy

To successfully learn about the world around them, infants must not only parse the input in their environment into meaningful pieces, they must also be sensitive to associations between features. Attending to regularities regarding the co-occurrence of features is important for grouping things like objects and sounds into categories (e.g., Rogers & McClelland, 2004; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Younger, 1985), learning about object function (e.g., Barsalou, Sloman, & Chaigneau, 2005; Oakes & Madole, 2008), and learning object labels (e.g., Apfelbaum & McMurray, 2011).

Although infants have been shown to recognize such co-occurrences or correlations in many domains (e.g., Madole & Cohen, 1995; Saffran et al., 1996; Werker, Cohen, Lloyd, Casasola, & Stager, 1998), the evidence regarding the developmental trajectory of attention to feature correlations is mixed. Generally, infants initially attend to individual features of objects and gradually develop the ability to process relations among features (e.g., Bahrick, Hernandez-Reif, & Flom, 2005; Baumgartner & Oakes, 2011; Werker et al., 1998; Younger & Cohen, 1983, 1986). The question is how infants' sensitivity to correlations changes, once they begin to recognize those relations. Some studies indicate that infants gradually attend to an increasing number of feature correlations as they develop. In a series of studies, Rakison (2004; Rakison & Poulin-Dubois, 2002) found that 14-month-old infants attended to relations between two dynamic features (global motion of an object and local motion of object parts) before they

demonstrated learning of relations between one static and one dynamic feature (object appearance and motion). Similarly, Hernandez-Reif and Bahrick (2001) showed that 4-month-old infants associated tactile information with object shape but not color, while 6-month-old infants associated tactile information with both shape and color.

Research from our own lab has shown that infants gradually attend to more relations between features in complex, dynamic events as they develop (Baumgartner & Oakes, 2011; Perone & Oakes, 2006). When habituated to events involving an object being acted upon by a hand, accompanied by a sound, eight-month-old infants learned the individual features of the events (object appearance, action, and sound), but did not attend to relations between the features (Baumgartner & Oakes, 2011). Ten-month-old infants were sensitive to the relation between the appearance of an object and the action that was performed, but did not attend to relations involving the sound (Perone & Oakes, 2006). Twelve-month-old infants showed evidence of learning the relation between the object and sound in addition to the relation between the object and action (Baumgartner & Oakes, 2011). This increase in relations that are attended to over developmental time is thought to be related to increased information processing capabilities.

Other studies suggest that infants initially are unselective in the features and relations they attend to and they gradually become more selective, perhaps as a result of a period of “perceptual narrowing” (Lewkowicz & Ghazanfar, 2009; Scott, Pascalis, & Nelson, 2007). According to this theory, infants’ experience tunes their perception to familiar features and relations, eventually causing them to be relatively insensitive to unfamiliar features and relations. For example, Madole and Cohen (1995) found that 14-month-old infants attended to a relation between the appearance of one feature of an object and the function of another feature of the same object, while 18-month-old infants only attended to the relation between the appearance

and function of the same feature. Madole and Cohen interpreted this finding as an indication that older infants were selectively attending only to the relation that is most likely to occur and be informative in the real world, whereas younger infants' attention was not constrained by such knowledge. Namy and Waxman (1998) observed a similar narrowing of infants' attention to associations between gestures and objects. Eighteen-month-old infants associated both gestures and words with objects, but 26-month-old infants attended only to relations between words and objects. In a similar study, Woodward and Hoyne (1999) found that 13-month-old infants learned both word-object and sound-object associations, while 20-month-old infants attended only to associations between words and objects. The authors of these studies interpreted their findings as an indication that a general ability to associate symbols with object categories narrows as a result of experience.

Recent word recognition and word learning studies have indicated that younger infants appear to also be initially unselective in the feature associations they attend to in linguistic tasks. Younger infants link acoustic features that are phonemically non-contrastive (e.g., speaker, pitch) to early word representations, while older infants are more "adult-like" in their ability to prioritize contrastive phonemic information (e.g. voice onset time) in their word representations (e.g., Rost & McMurray, 2010; Singh, 2008). This pattern of perceptual narrowing across infancy has been found across multiple domains including discrimination of phonemes (Werker & Tees, 1984), linguistic signs (Palmer, Fais, Golinkoff, & Werker, 2012), and faces (Pascalis, de Haan, & Nelson, 2002).

Taken as a whole, this work demonstrates developmental changes in infants' selective attention to relations between features across domains. Infants' attention to some kinds of relations appears to develop gradually, while attention to other features undergoes a period of

perceptual narrowing. Regardless of the trajectory, the research reviewed thus far does not answer the question of what factors allow infants to increasingly attend to relations that are meaningful for learning and ignore irrelevant correlations. The following sections will explore possible mechanisms behind infants' ability to selectively attend to some features and relations and ignore others.

Statistical Learning

There are practically limitless possibilities for feature associations that *could be* meaningful in the environment. Regardless of their initial level of selectivity, infants have the task of determining which associations *are* meaningful. Infants are likely aided in this task by the way that features are associated in the real world. Some associations are much more likely to occur in the real world than others, and many categories—especially at the basic level—are formed around clusters of frequently correlated attributes (Rosch et al., 1976). Therefore, if infants are sensitive to the frequency with which various attributes co-occur, they can make use of this information to group objects into categories.

A growing body of research shows that infants do, in fact, attend to such statistical regularities. In a groundbreaking study of statistical learning in infancy, Saffran, Aslin, and Newport (1996) demonstrated that by the time infants are 8 months old, they can segment words from a fluent stream of artificial speech by tracking transitional probabilities between syllables. A follow-up study showed that statistical learning is not limited to linguistic stimuli; infants used statistics to segment sequences of non-linguistic tones as well (Saffran, Johnson, Aslin, & Newport, 1999). Moreover, using the statistical features of a speech stream to divide it into units is meaningful for word learning. Infants are more likely to learn object labels after having had the opportunity to segment those labels from fluent speech based on statistical probabilities (Graf

Estes, Evans, Alibali, & Saffran, 2007). Subsequent studies have provided support for the notion that not only is statistical learning relevant to learning linguistic and other types of auditory information, it is a truly domain-general mechanism. For example, Kirkham, Slemmer, and Johnson (2002) modified Saffran et al.'s (1996) design and found that infants as young as 2 months of age made use of statistical information to segment sequences of visual stimuli.

In the domain of categorization, Younger (1985) found that 10-month-old infants were sensitive to the statistical distribution of correlations between features in a set of exemplars, and could segregate items into categories based on the nature of this distribution. Infants familiarized to a broad set of exemplars in which features varied independently showed evidence of having formed a single, broad category. In contrast, infants familiarized with a similar set of exemplars with two subgroups of objects defined by stable correlations between values of attributes (e.g., short legs correlated with long neck, thick tail, and close-together ears; long legs correlated with short neck, thin tail, and wide-set ears) showed evidence of forming two categories on the basis of the distribution of correlated features.

Maye, Werker, and Gerken (2002) found that 6- and 8-month-old infants' discrimination of phonemes is also influenced by distributional cues. Infants familiarized with speech sounds along a continuum of voice onset time (VOT) exhibiting a bimodal frequency distribution were able to discriminate between tokens from the endpoints of the continuum ([da] and [ta]), whereas infants familiarized to sounds from the same continuum but with a unimodal frequency distribution did not discriminate between these same tokens. It is important to note that in this experiment exposure to a unimodal frequency distribution of phonemes actually *reduced* infants' phonemic discrimination, as other studies have shown that infants are able to discriminate between these same phonemes when given no pre-exposure (Pegg & Werker, 1997).

Connectionist models of development like the one proposed by Rogers and McClelland (2004) operate under the principle that learning and category formation take place as a result of attending to the coherent co-variation of attributes across items. These models assume that infants' early learning strategies are naïve with regards to the relative weight given to various features and associations between features that are present in the environment. As infants gain experience with tracking statistics and regularities in the world around them, however, more weight is given to features that consistently co-occur (e.g., wings and feathers), and those features become prioritized for determining category membership. As a result, with experience infants become increasingly sophisticated in their ability to discover which features (and associations between features) are important in a given context.

The role of variability

One factor that appears to play an important role in how infants learn to distinguish between contrastive and non-contrastive cues as they track regularities in the environment is the *variability* of the input. The effect of variability on learning has been studied in many domains in the field of infant cognition. The following section will review the relevant literature in this area.

Infants are sensitive to many sources of variability in the input they are exposed to, and this attention to variability has the potential to both hinder and facilitate learning. Increased variability may interfere with learning, as it might overload infants' finite processing abilities. Exposure to variable input (rather than repeated exposure to a single stimulus) means more information for the infant to attend to, remember, and compare across exemplars. Increased variability could also mean more "noise" that the infant must successfully ignore in order to organize the input into meaningful units. Variability in features that are particularly salient for young infants (e.g., sound or action) could make processing features that are potentially

informative but are also less salient even more difficult. Decreasing the amount of variability in the input might therefore help infants by reducing the “task demands” that a noisy environment places on the infant. For example, Cohen and Oakes (1993) found that 10-month-old infants had a harder time perceiving causality when the particular objects used in the events varied from trial to trial than when the objects remained constant. Apparently at this age infants initially perceive causality in terms of a relation between specific objects, and have difficulty generalizing a causal relationship when presented with events involving a variety of exemplars. Reducing the variability of exemplars has also been shown to facilitate infants’ early learning of abstract spatial relations between objects such as *above* and *below* (Quinn, Cummins, Kase, Martin, & Weissman, 1996), *between* (Quinn, Adams, Kennedy, Shettler, & Wasnik, 2003), and *on* (Casasola, 2005).

Alternatively, there is reason to think variability may actually facilitate learning in many circumstances. In the absence of variability, there are few cues for the infant to use to distinguish between features that are informative (contrastive) or irrelevant (non-contrastive) for categorization. For example, if an infant is exposed to only one exemplar from the category *ball*, there is no way for her to determine which features are relevant for category membership (e.g., round shape, able to be rolled or bounced) and which are irrelevant (e.g., color, size). When some aspects of the input vary and other aspects do not, the variable input may serve as a cue to help direct the infant’s attention to features that are consistent across exemplars and are thus important for determining category membership, or away from features that are irrelevant for categorization. Continuing the example from above, exposure to many exemplars from the category *ball* of many sizes and colors might help the infant learn that color and size are features are not relevant to category membership. In support of this, Mather and Plunkett (2011) refer to

variability as “a central feature to defining category membership and structure” (p. 445), and Bhatt and Quinn (2011) identified exposure to variability as one of five kinds of experiences that induce perceptual learning by allowing infants to “hone in on functionally critical invariants in the environment” (p. 13).

Research on variability and categorization has revealed that the issue is more complex than simply learning in the presence or absence of variability; instead, the *degree* of variability influences categorization by infants. Specifically, the degree to which exemplars are perceptually similar to one another affects the inclusivity or exclusivity of the categories formed by infants (Oakes, Coppage, & Dingel, 1997; Quinn, Eimas, & Rosenkrantz, 1993). These studies have shown that infants form different categories when they are shown exemplars that are very perceptually similar than they do when they are shown a more perceptually variable set of exemplars. Using a visual preference task, Quinn et al. (1993) found that 3- and 4-month-old infants formed more exclusive categories when shown perceptually similar exemplars than they did when they were exposed to exemplars that were more perceptually variable. Oakes and colleagues (1997) exposed 10- and 13-month-old infants to perceptually similar or variable exemplars in an object-examining task. Like Quinn et al. (1993), they found that infants formed more inclusive categories when familiarized with variable exemplars and less inclusive categories when familiarized with perceptually similar exemplars.

There is also evidence to suggest that the amount of perceptual variability between exemplars influences infants’ categorization on a moment-to-moment basis. Using stimuli based on those used by Younger (1985), Mather and Plunkett (2011) found that infants’ categorization is influenced by the order in which exemplars are presented. In this study, 10-month-old infants categorized the same set of exemplars differently based on the perceptual distance between

consecutive exemplars. When familiarized with a set of figures with a large perceptual distance between consecutive exemplars, infants looked longer to an atypical figure than to the category prototype, providing evidence of having organized the exemplars into a single category. Infants familiarized to a stimulus set which minimized the perceptual distance between consecutive exemplars, however, did not show a preference for either test stimulus. The authors interpreted their findings as evidence that sequences with more moment-to-moment variation between exemplars facilitated categorization more than sequences with small differences between consecutive items. However, the authors also acknowledge that there is likely a limit to the facilitative effects of variability. According to Plunkett's (2011) 'perceptual load hypothesis,' the relation between infants' categorization and perceptual load is characterized by an inverted U-shaped function in which categorization is optimal at moderate levels of perceptual load (also referred to as the 'Goldilocks Principle'; Kidd, Piantadosi, & Aslin, 2012). Therefore, some amount of variability between exemplars will facilitate categorization, but too much variability will be disruptive.

Importantly, in these two examples, the level of variability in the features that were *informative* for category membership was manipulated. Recall that Bhatt and Quinn (2011) suggested that variability in irrelevant features would highlight critical invariants and facilitate perceptual learning. Thus, variability in contrastive or relevant aspects of the input should have a different effect on learning than variability in non-contrastive or irrelevant aspects of the input.

This distinction between variability in contrastive and non-contrastive features has been studied in the domain of word learning. Specifically, variability in linguistically *non-contrastive* cues appears to aid infants' learning of linguistic input. In a statistical word segmentation task similar to Saffran et al. (1996), variability in irrelevant features helped infants learn relations

between non-adjacent elements in a speech stream (Gomez, 2002). In this study, infants were familiarized with a speech stream composed of three-element strings. The only cue for segmenting the stream into strings was a stable relation between the first and third elements within each string. Eighteen-month-old infants showed evidence of learning these non-adjacent dependencies only when there was increased variability in the middle (irrelevant) element of each string. Variability in another linguistically non-contrastive feature, affect, also appears to help infants recognize words in new contexts (Singh, 2008). In this study, 7.5-month-old infants were better able to recognize words in passages when they were familiarized to exemplars of the word that varied in affect (e.g., happy and neutral) than when the familiarization exemplars did not vary. The results from a series of experiments support the conclusion that affective variability reduced the perceived importance of surface features and allowed infants to prioritize phonological invariants in their word representations.

In addition to word segmentation and recognition, there is evidence that variability in non-contrastive features helps infants form associations between words and objects. Many studies of early word learning have utilized the “Switch” design (Werker et al., 1998), in which infants are habituated to two word-object pairings and then tested with two test events: a ‘same’ event containing one of the word-object pairings from habituation, and a ‘switch’ event containing the object from one pairing and the word from the other pairing (e.g., object A with word B). In a pair of influential studies, Werker and colleagues found that 14-month-old infants were able to learn associations between two novel words and their referents when the novel words were easily discriminable from one another (e.g., *lif* and *neem*; Werker et al., 1998), but did not dishabituate to the switch test event when the two labels were phonemically similar (e.g.,

bih and *dih*), despite the fact that infants succeeded on a pure discrimination task with this minimal pair of words (Stager & Werker, 1997).

In an effort to examine why infants had difficulty learning phonologically similar words in this task despite their well-developed ability to discriminate phonemes, Rost and McMurray (2009, 2010) conducted a set of follow-up studies. They hypothesized that part of infants' difficulty in Stager and Werker's (1997) task came from a lack of variability in the word tokens infants were exposed to. Because speaker and context did not vary, infants might have given these features as much weight as the phonemically relevant feature of VOT in their newly formed word representations. As a result, the representations for the two newly learned words would be highly overlapping, causing infants to view the switch test event as an acceptable pairing. To test this hypothesis, Rost and McMurray (2010) exposed infants to multiple exemplars of each label that varied in either lexically contrastive (VOT, fundamental frequency) or non-contrastive (speaker, prosody) features. They found that variability in non-contrastive cues helped 14-month-old infants succeed in the minimal pair switch task, but variability in contrastive voicing cues did not facilitate learning. These findings were interpreted as evidence that increased indexical variability within each label category helped infants learn that those cues were not informative and allowed infants to focus instead on the relatively invariant (and contrastive) cues such as VOT. Subsequent modeling work by McMurray and colleagues (Apfelbaum & McMurray, 2011; Toscano & McMurray, 2010) indicates that speaker variability is helpful because it allows associative learning mechanisms to place more weight on features that are consistent across exemplars and reduces the chance of erroneous links forming between indexical cues and words. In a recent follow-up, McMurray and colleagues found that 14-month-old infants also succeeded in the minimal pair switch task when they heard variable tokens

produced by a single speaker, supporting their claim that the benefits of non-contrastive feature variability are not specific to features that vary across speakers (Galle et al., 2014).

It is also possible that hearing labels produced by multiple speakers facilitates learning because it guides infants' attention to the referential nature of the task. This interpretation is supported by a similar series of experiments conducted by Fennell and Waxman (2010). In one condition, infants were shown training trials with familiar object-label pairs prior to the habituation phase, and in another condition the novel labels were embedded in familiar sentence frames. In both cases, 14-month-old infants learned the associations between minimal word pairs and objects that were not learned in the absence of such referential cues (Stager & Werker, 1997).

Less is known about the distinction between variability in contrastive (relevant for category membership) and non-contrastive (irrelevant for category membership) features in object categorization. Kloos and Sloutsky (2008) investigated how differences in category density—defined as the ratio of category-relevant variance to total variance—affected category learning. They tested both adults and preschool-aged children in one of two conditions: a *supervised* condition in which they were given explicit information about category-inclusion rules (e.g., “Ziblets have a short tail”), and an *unsupervised* condition in which no information was given. In the unsupervised condition, which is especially relevant to learning in infancy, dense categories with a high ratio of relevant to irrelevant variation were easier to learn than sparse categories in which only some sources of variation were relevant to category membership. The authors proposed that better learning in the dense category condition was related to the increased demands on selective attention in the sparse category condition. In the sparse condition, the learner had to determine which features to attend to and selectively ignore the

large number of other features. It is important to note, however, that this study did not include a condition comparable to the non-contrastive variability condition of Rost and McMurray (2010) in which *only* category-irrelevant features varied while category-relevant features remained constant, so it is not possible to draw conclusions about category learning in such conditions. This question also has yet to be addressed with infants or in the domain of word learning, and will be the focus of one of the experiments proposed here.

The Role of Selective Visual Attention

The discussion to this point has been about how infants develop the ability to attend to and learn relevant or contrastive aspects of the input they are exposed to, and inhibit their attention to or learning of irrelevant or non-contrastive aspects of this input. In the previous discussion, the focus was on how variation in the input itself contributes to this process. However, infants' basic ability to selectively attend to some information—and therefore inhibit responding to other information—must also be a factor in this process. Because infants can process only a limited amount of information at once, the ability to selectively control their visual attention is crucial for learning. Selectivity has been defined as “the ability to filter out unwanted information” (Desimone & Duncan, 1995, p.193), and selective attention is “how action is directed only to stimuli relevant to the current behavior and not to other irrelevant stimuli” (Tipper, Bourque, Anderson, & Brehaut, 1989, p. 353).

Various theories have been proposed to account for how selection of relevant stimuli and inhibition of distractors takes place. Inhibitory accounts of visual selective attention highlight the importance of inhibitory processes (e.g., Tipper, 1985). Stimuli and features of stimuli in the environment (some relevant and others irrelevant) compete for attention, and this competition is resolved not only by the selection of relevant stimuli to attend to, but also by the inhibition of

irrelevant distractors. According to Tipper and colleagues' (1989) model of selective attention, excitatory and inhibitory mechanisms operate together to direct attention toward a relevant target and actively inhibit competition from distracting stimuli. According to the biased-competition model of attention (Desimone & Duncan, 1995), multiple objects in the visual input compete for processing resources, and the visual system uses any relevant information it has (either bottom-up or top-down) to bias the competition toward some objects over others. Treisman's (Treisman & Gelade, 1980) Feature-Integration Theory (FIT) posits that some features such as color and shape are processed automatically and in parallel across the visual field without the need for selective attention, while conjunctions of multiple features within an object (e.g., blue X, red O) require focused attention and are processed serially. And according to Wolfe's (2007) Guided Search Model, attention is mediated by a "guiding representation" that is derived from early visual processes. For example, if a target differs from distractors in terms of orientation, the bottom-up salience of orientation will guide attention to the target and inhibit attention to the distractors. Despite their differences, these theories all predict that targets that differ from distractors on a single, salient dimension (e.g., color, orientation) will "pop out" from the display, while targets that differ from distractors based on the conjunction of features will require focused attention in the form of a serial search in order to locate. While successful searches are often thought of in terms of directing attention to a target item, an equally important component of such focused attention is the use of inhibitory processes to ignore other elements in such displays.

In studies of visual search with 3-month-old infants, Adler and colleagues demonstrated that this form of bottom-up pop-out occurs under some conditions in infancy as well as adulthood. In one study, they found that the average latency of infants' first saccade to a feature-

present target (e.g., an R among P distractors) did not vary with the number of distractors, while the average latency of their first saccade to a feature-absent target (e.g., P among Rs) did (Adler & Gallego, 2014). This asymmetry in search performance was interpreted as evidence that visual search in infancy might be guided by adult-like mechanisms of selective attention.

The kinds of selective attention measured in these studies seem to play a role in infants' visual perception. In a study of unity perception in which 3-month-old infants viewed a rod moving behind an occluder, perception of the visible rod pieces as a unitary whole was related to infants' visual scanning patterns of the scene (Johnson, Slemmer, & Amslo, 2004). Infants who were classified as "perceivers" selectively attended to some aspects of the scene (the moving pieces of the rod) and inhibited looks to the irrelevant area (the occluder). A second set of studies showed that unity perception in the occluded rod task was related not only to scanning patterns during that task, but also to performance on a visual search task (Amslo & Johnson, 2006). In the visual search task, infants were shown an array of 27 vertical distractor rods with a target rod that differed from the distractors in one of two ways: in the competition condition, the target differed from the distractors by vertical angle (30° , 60° , or 90° from vertical); in a control condition, the target rod moved laterally at one of three speeds (1 Hz, 1.5 Hz, or 2 Hz). Unity perception in the occluded rod task was associated with higher accuracy (proportion of trials with looks to the target) and longer saccade latencies to the target in the competition condition relative to the control condition. The perceivers' longer saccade latencies to the target in the competition condition was interpreted by the authors as suggesting that those infants were using a visual search mechanism more effectively than were non-perceivers in order to overcome competition between the target and distractors.

There is additional evidence that selective attention influences how infants learn about novel stimuli. Jankowski, Rose, and Feldman (2001) found that infants who examined targets with shorter, more distributed looks (“short lookers”) were more likely to exhibit a novelty preference in a subsequent discrimination task than were infants who were classified as “long lookers.” The results are consistent with the idea that the ability to disengage one’s attention from a look and selectively attend elsewhere is related to efficient stimulus processing. In a second study in which long lookers were trained to scan the targets with shorter, more distributed looks (a red light highlighted one quadrant of the target at a time), those infants also showed a novelty preference in the subsequent discrimination task. The results of this training study suggest a causal relation between the distribution of visual attention and efficient processing of novel stimuli. Inducing long lookers to disengage attention from one region and scan stimuli more broadly appeared to improve processing speed and/or depth of encoding.

Other work shows developmental changes in infants’ ability to disengage attention from distractors. Studies using variations of the gap overlap procedure have shown that infants are faster to orient to a peripheral stimulus when there is a temporal gap between its presentation and the presentation of a preceding center stimulus than they are in an overlap condition in which the peripheral target is presented before the offset of the central stimulus (Hood & Atkinson, 1993; M. H. Johnson, Posner, & Rothbart, 1991). The slower saccade latencies to the peripheral stimulus in the overlap condition are thought to be the result of the need for the infant to disengage attention from the center stimulus. These studies have shown that with age, infants become increasingly adept at disengaging from a center stimulus to attend to a peripheral target.

The kinds of inhibition and disengagement processes that develop during infancy are important in adults’ cognitive processes. For example, using a visual search task, Fukuda and

Vogel (2011) tested the relation between working memory (WM) capacity and inhibition. Previous research indicated that adult WM capacity is related to attentional capture by distractors, as high-capacity individuals appeared better able to resist orienting to irrelevant items in a visual choice task (Fukuda & Vogel, 2009). Here the authors tested an alternative hypothesis: differences in WM capacity are related not to an individual's susceptibility to attentional capture, but rather to that individual's ability to disengage from distractors and redirect his or her attention to a target. In this study, adults completed a visual search task in which a flanker prime preceded a test array with a target stimulus. The flanker was either the same color (relevant) or a different color (irrelevant) than the target. Individuals with high and low WM capacity showed no differences in their susceptibility to attentional capture (as measured by target accuracy following an irrelevant flanker), but individuals with high WM capacity showed faster recovery times from this attentional capture following a relevant flanker. In other words, the ability to quickly inhibit attention to and disengage from distractors and redirect attention to a target was correlated with a higher WM capacity. This relation between inhibition and WM capacity has implications for the relationship between inhibitory processes and general cognition, as WM capacity has also been found to be correlated with cognitive measures such as fluid intelligence (Fukuda, Vogel, Mayr, & Awh, 2010).

These studies confirm that inhibitory processes are particularly important for selective attention. These inhibitory processes develop over the period of infancy (M. H. Johnson, 1990; Ruff & Rothbart, 1996). For example, Amso and Johnson (2005) investigated inhibition of visual attention in infancy using a variation of the spatial negative priming (SNP) paradigm (e.g., Tipper, 1985). In this study, 9-month-old infants were shown a prime array with a target object (e.g., a toy bear) and a distractor (e.g., a gray diamond) in 2 locations. After a brief delay (67,

200, or 550 ms), infants were probed with a target object (e.g., a duck) that appeared in either the same location as the distractor (*ignored repetition* condition) or a new location (*control* condition). The logic of the task is that during the prime, infants will attend to the more salient target and inhibit looks to the distractor. If infants do indeed inhibit responding to the location of the distractor during the prime, then it should be difficult for them to then shift attention to that location during the probe when it contains a target. Because in the control condition the target is in a location that has neither been inhibited nor primed, infants' responding to these targets should provide a good comparison for their responding to targets in the ignored repetition condition. The 9-month-old infants showed evidence of spatial negative priming through impaired performance on ignored repetition trials compared with control trials, indicating that by this age infants are capable of inhibiting attention to distractors in a visual selection task (Amso & Johnson, 2005). The study also revealed differences in inhibition between infants and adults as a function of the interstimulus interval (ISI) between the prime and probe displays, suggesting that the inhibitory mechanism measured by this task is still developing at 9 months of age.

There is also a growing body of evidence that inhibitory selective attention is closely related to learning and memory. For example, Markant and Amso (2013) tested 9-month-old infants in a version of Posner's (1980) classic spatial cueing task in which target items were presented in either cued or non-cued locations. They found that infants showed enhanced memory for specific items as well as abstraction of category information when items were presented in the non-cued location following a delay long enough to allow for suppression of attention to the cued location (700 ms). In contrast, infants showed no evidence of having learned item or category information in a facilitation condition in which target items appeared in cued locations following a short delay (< 100 ms) and. These results demonstrate that the

attentional mechanisms activated by a learning context are closely related to (and perhaps determine) the amount of learning that takes place.

Additional research suggests significant developmental change during infancy in the ability to focus attention on relevant features/stimuli and resist attending to irrelevant features/stimuli. Oakes and Tellinghuisen (1994) found developmental changes between 6 and 10 months in infants' ability to inhibit responding to a distractor as they actively attended to a toy; older infants exhibited longer latencies to orient to the distractor than younger infants. Studies of the physiological correlates of sustained attention in infancy (e.g., heart rate deceleration) have shown greater heart rate deceleration in older infants than younger infants in both cross-sectional and longitudinal designs, suggesting a developmental increase in sustained attention between 14 and 26 weeks (Casey & Richards, 1988; Richards, 1989). Although measures of distractibility vary as a function of the relative complexities of the target and distractor objects as well as developmental differences in information-processing abilities, studies have consistently shown that infants are less distractible during these periods of sustained attention (Casey & Richards, 1988; Oakes & Tellinghuisen, 1994). Importantly, these measures of sustained attention in infancy seem to reflect stable individual differences in information processing abilities. Kannass, Oakes, and Shaddy (2006) found that measures of distractibility in infancy were related to measures of focused attention in toddlers engaged in an object free play task, showing relative stability across age and contexts in children's selective attention abilities.

These developmental changes in selective attention may be related to other developmental changes observed in studies of attention to relations between features in dynamic events, in which it appears infants initially have difficulty inhibiting attention to dynamic features such as action in order to attend to relatively static appearance features of objects or

faces (e.g., Bahrick, Gogate, & Ruiz, 2002; Baumgartner & Oakes, 2011). These changes in the information infants attend to and how infants prioritize some features over others may reflect maturation of the same inhibitory processes that underlie infants' developing ability to resist distracting stimuli in other contexts.

The evidence reviewed above suggests a close link between infants' ability to selectively control their attention and their ability to perceive and learn about relevant aspects of input in their environment. It is possible that the mechanism that allows infants to attend to a target stimulus while inhibiting attention to distracting stimuli in a visual search context is also involved in directing attention toward features of stimuli that are relevant for learning and away from features that are irrelevant. This possibility will be tested by the experiments proposed in this dissertation.

The Present Experiments

The literature reviewed above shows that the variability of object exemplars affects the nature of the categories infants form (Mather & Plunkett, 2011; Oakes et al., 1997; Quinn et al., 1993), and the variability of speech features affects how infants recognize words and attend to object-label associations (Rost & McMurray, 2009; Singh, 2008). In addition, recent research suggests that word learning is particularly facilitated by variation in acoustic features that are not informative for word meaning, such as speaker or intonation (Galle et al., 2014; Rost & McMurray, 2010). However, those studies have not addressed if indexical label variability of this kind is helpful only in the context of minimal pair word learning (e.g., *buk* and *puk*), or if it also facilitates learning when words are dissimilar (e.g., *lif* and *neem*). In her original examination of infants' learning of novel object-label associations, Werker found that 14-month-old infants successfully matched easily discriminable words with novel referents, but 12-month-old infants

did not (Werker et al., 1998). If indexical variability is generally beneficial for word learning, we might expect it to help 12-month-old infants succeed on a word-learning task of this sort when the words are dissimilar. If, however, the effect of indexical variability observed by McMurray and colleagues is limited to facilitating the disambiguation of minimal pairs, then we would not expect speaker variability to help 12-month-old infants' learning of dissimilar words.

The studies previously presented also do not address the question of whether the facilitative effect of non-contrastive variability is the result of a domain-general learning mechanism that allows infants to attend to statistical regularities in the world around them, or if it is specific to speech. If it is domain-general, then it is possible that infants' learning of associations between objects and labels should be facilitated not only by variability in non-contrastive acoustic features of the labels, but by variability in perceptual features of the objects themselves that are not informative for category membership (e.g., color).

Few studies have demonstrated consistent word learning or recognition in infants younger than 14 months of age. We therefore do not expect 12-month-old infants to demonstrate learning of novel object-label associations in the absence of stimulus variability, a learning context akin to Werker et al.'s (1998) original study, on a group level. However, we anticipate that there will be variation in individual infants' responding to these novel labeling events. As discussed above, there is evidence that infants' selective attention is linked to learning and memory across domains, and we are interested in the role that selective attention might play in individual differences in novel word learning.

These questions form the basis for the present experiments, which test the following hypotheses:

Hypothesis 1: Infants will not show robust evidence of learning object-label associations at 12 months in the absence of variability (Experiment 1).

Hypothesis 2: Variability affects learning via a domain-general learning mechanism, and infants' learning of object-label associations will be facilitated by variability in non-contrastive features of the labels, speaker and prosody (Experiment 2), as well as variability in a non-contrastive feature of the objects, color (Experiment 3).

Hypothesis 3: Measures of selective attention, as assessed through a visual search task (Experiment 4) will be related to infants' learning of object-label associations in the absence of variability (Experiment 1).

Chapter 2: The Role of Variability in Word Learning

We tested 12-month-old infants in three experiments designed to investigate their ability to recognize and learn words in the presence and absence of task-irrelevant variability. In Experiments 1-3, infants were exposed to two object-label associations using a standard habituation procedure (Oakes, 2010), and their learning was assessed using the *looking-while-listening* procedure (Fernald, Zangl, Portillo, & Marchman, 2008). Experiments differed in the amount and kind of stimulus variability to which infants were exposed. In Experiment 1, infants saw a single exemplar from each of two novel object categories and heard each object paired with a single token of a novel label. In Experiment 2, we introduced acoustic variability in the form of multiple speakers providing multiple label tokens for each object. In Experiment 3, we varied a visual feature of the objects (color). Methods that are shared by Experiments 1-3 are discussed in the following section, and methods that differ by experiment are discussed in the context of each experiment.

General Method

Participants

Participants in all experiments were 12-month-old English-learning¹ infants. All infants were born full-term (within 21 days of due date) and had no reported health, vision, or hearing problems. Infant names were obtained from the California State Vital Records office. Parents were sent information regarding the studies and an invitation to participate, and parents who responded to this mailing were contacted when their infants reached the appropriate age to schedule an appointment. Parents were not compensated, but all infants were given a toy, book, t-shirt, or bib in appreciation for their participation.

¹ Greater than 70% of language exposure is to English, based on parental report.

Stimuli

Stimuli for Experiments 1-3 were color pictures of familiar and novel objects paired with audio recordings of familiar and novel labels. Familiar objects and labels were used for pretest and familiar test trials, and novel objects and labels were used for habituation trials and novel test trials (see Design and procedure, below). The entire set of visual and auditory stimuli will be described here. To examine the effects of different types of variability on learning, different subsets of these items were used as pretest and habituation stimuli in each of the three experiments. These subsets will be discussed in the context of each experiment.

Visual stimuli. Visual stimuli were color photographs of six everyday object categories likely to be familiar to 12-month-old infants (baby, ball, car, dog, kitty, shoe), and three-dimensional computer-generated images of novel objects selected from two categories (to be referred to as Category A and Category B) of the “Fribbles” stimulus set (Behrmann & Williams, 2007) (see Figure 1). The novel objects were monochromatic and uniform in texture, but had multiple parts that differed in value and configuration between categories.



Figure 1. Visual stimuli used during test trials for all experiments. (Note: Experiment 1 used a multi-colored beach ball instead of an orange basketball.)

Additional exemplars of each familiar and novel object category varying in color were created using Adobe Photoshop. Exemplars were edited from the original items to vary both within and across color categories along multiple color dimensions including hue, brightness, and saturation. Five exemplars from each of seven color categories (blue, brown, green, orange, purple, red, yellow) were created for each novel object category (see Figure 2). These exemplars were added to the original blue exemplar from Category A and the original green exemplar from Category B (see Figure 1) for a total stimulus set of 36 color exemplars from each novel object category. Two additional exemplars from each of five color categories were created for the ball (blue, green, purple, red, yellow), car (green, orange, purple, red, yellow), and shoe (blue, green, orange, purple, yellow) familiar object categories, for a total of 11 color exemplars from these familiar object categories (additional color exemplars were not created for the baby, dog, and kitty pictures for practical reasons that will be discussed in the context of Experiment 3). CIE coordinate ranges for each color category within each object category are found in Table 1.

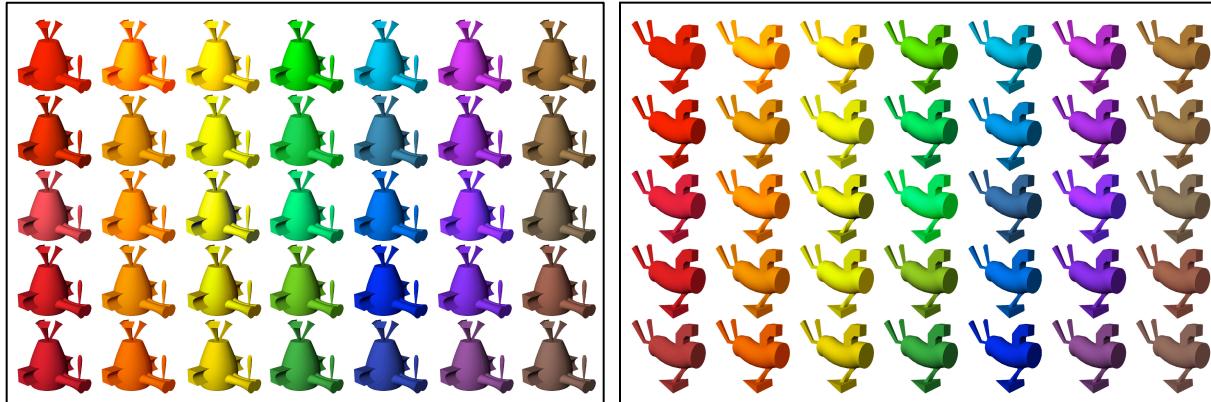


Figure 2. Color exemplars of novel objects from Category A (left panel) and Category B (right panel).

Table 1. CIE [x, y, Luminance] Coordinates for Experiment 3 Pretest and Habituation Visual Stimuli

| Color Category | Category A | | | Category B | | | Ball | | | Familiar Objects | | | Shoe | | |
|----------------|--------------|--------------|---------------|--------------|--------------|---------------|---------------|---------------|---------------|------------------|---------------|---------------|---------------|--------------|---------------|
| | x | y | Luminance | x | y | Luminance | x | y | Luminance | x | y | Luminance | x | y | Luminance |
| Blue | 0.17 0.02 | 0.19 0.10 | 47.4 39.2 | 0.18 0.02 | 0.20 0.09 | 46.5 29.8 | 0.18 0.02 | 0.21 0.08 | 45.8 28.8 | | | | 0.19 0.02 | 0.25 0.09 | 55.7 30.1 |
| Brown | 0.41 0.04 | 0.40 0.02 | 42.5 9.6 | 0.41 0.04 | 0.40 0.03 | 47.5 11.9 | | | | | | | | | |
| Green | 0.27 0.02 | 0.54 0.02 | 125.9 38.4 | 0.28 0.03 | 0.54 0.01 | 118.0 42 | 0.26 0.01 | 0.53 0.03 | 100.5 41.8 | 0.29 0.04 | 0.52 0.02 | 115.4 37.7 | 0.27 0.003 | 0.52 0.04 | 87.6 15.3 |
| Orange | 0.51 0.03 | 0.42 0.02 | 112.2 22.8 | 0.52 0.04 | 0.41 0.02 | 99.5 29.9 | | | | 0.53 0.02 | 0.40 0.02 | 83.7 5.3 | 0.57 0.04 | 0.37 0.02 | 53.1 0.7 |
| Purple | 0.25 0.05 | 0.15 0.05 | 38.0 10 | 0.26 0.05 | 0.15 0.05 | 32.8 8 | 0.25 0.07 | 0.16 0.07 | 25.3 6.9 | 0.3 0.03 | 0.22 0.06 | 35.7 8.6 | 0.30 0.07 | 0.19 0.06 | 28.8 2.3 |
| Red | 0.61 0.03 | 0.34 0.01 | 51.9 10.6 | 0.61 0.04 | 0.34 0.01 | 45.8 7.4 | 0.62 0.02 | 0.34 0.003 | 26.6 8.7 | 0.59 0.01 | 0.33 0.001 | 37.6 2 | | | |
| Yellow | 0.39 0.01 | 0.49 0.01 | 216.4 28.6 | 0.39 0.02 | 0.49 0.01 | 193.2 39.1 | 0.44 0.003 | 0.46 0.001 | 147 36.8 | 0.40 0.002 | 0.48 0.01 | 234.5 30.4 | 0.40 0.02 | 0.49 0.01 | 159.5 17.7 |
| Test | 0.18 | 0.23 | 63.6 | 0.26 | 0.57 | 149.0 | 0.59 | 0.36 | 33.5 | 0.16 | 0.14 | 17.7 | 0.59 | 0.33 | 29.8 |

Note. Averages were calculated across five exemplars per color category for novel object categories and two exemplars per color category for familiar object categories. Values in italics are standard deviations.

Visual stimulus discrimination testing. To ensure that 12-month-old infants could discriminate between the novel object categories (i.e., "Fribbles") and between color exemplars within each novel object category, a separate group of 28 infants ($M = 383.3$ days, 17 female) completed a preliminary discrimination task. This task had two conditions: *between-category discrimination* and *within-category discrimination*. Testing took place using the same apparatus as Experiments 1-3 (see Apparatus, below).

Between-category discrimination. Infants were tested for their ability to discriminate between the novel object categories using a familiarization-test procedure. Infants were first familiarized with a single blue exemplar from one of the two novel object categories (randomly selected for each participant), presented on both the left and right sides of the screen (Figure 3, left). A trained observer initiated each trial when the infant fixated a central attention-getter, and recorded infants' looking to the left and right sides of the screen. The objects remained visible on the screen until the infant accumulated 10 s of total looking to the left and/or right object(s). This familiarization phase was followed by two test trials designed to assess infants' ability to discriminate between the two novel object categories. On each test trial, the object seen during familiarization was presented on one side of the screen and an exemplar of the same color from the other novel object category was simultaneously presented on the other side of the screen. Each pair of images remained visible on the screen for 5 s after an infant's first look to the left or right object. The location (left or right) of the familiarized object was counterbalanced by trial.

A trained observer pressed one key to record the duration of infants' looking to the left object and another key to record the duration of looking to the right object. The measure of interest was the proportion of total looking to both objects that was spent looking at the novel (i.e., not seen during familiarization) object. While a preference for the novel was expected, a

significant preference for either object would indicate successful discrimination between object categories. A two-tailed, paired samples t test indicated that infants looked more to the novel object than to the familiarized object than would be expected by chance, $t(27) = 2.77, p = .01, d = .52$. Thus, 12-month-old infants were able to discriminate between the two novel “Fribble” object categories after a short familiarization period.

Within-category discrimination. To test infants’ ability to discriminate between exemplars within a novel object category, we used a paired preference task. Each infant was presented with four 10-s trials in which two simultaneous streams of visual stimuli were presented side-by-side on the stimulus monitor (similar to the procedure used by Ross-Sheehy, Oakes, & Luck, 2003) (Figure 3, right). Within each stimulus stream, a single item was presented for 1 s, disappeared for 250 ms, and then reappeared for 1 s; this cycle of repeatedly appearing and disappearing continued for the duration of the trial (for a total of eight presentations of an object over each 10-s trial). On each trial, one stream was *unchanging*, and consisted of a single novel object exemplar presented on each cycle (e.g., an identical orange exemplar from Category B reappeared after each 250 ms blank period). The other stream was *changing*, and consisted of a different color exemplar from the same novel object category presented on each cycle (e.g., a green exemplar from Category B followed by a purple exemplar from Category B followed by a yellow exemplar from Category B, etc.). Given that infants generally prefer novel, complex stimuli to simple, familiar stimuli (e.g., Brennan, Ames, & Moore, 1966), infants are expected to preferentially look at the changing stream than at a stream in which the same item is presented on each cycle. Indeed, using this task, infants from 4 to 12 months of age have been shown to robustly prefer changing streams when the items are simple squares or complex objects that change from cycle to cycle (Kwon, Luck, & Oakes, 2014; Ross-Sheehy et al., 2003). The

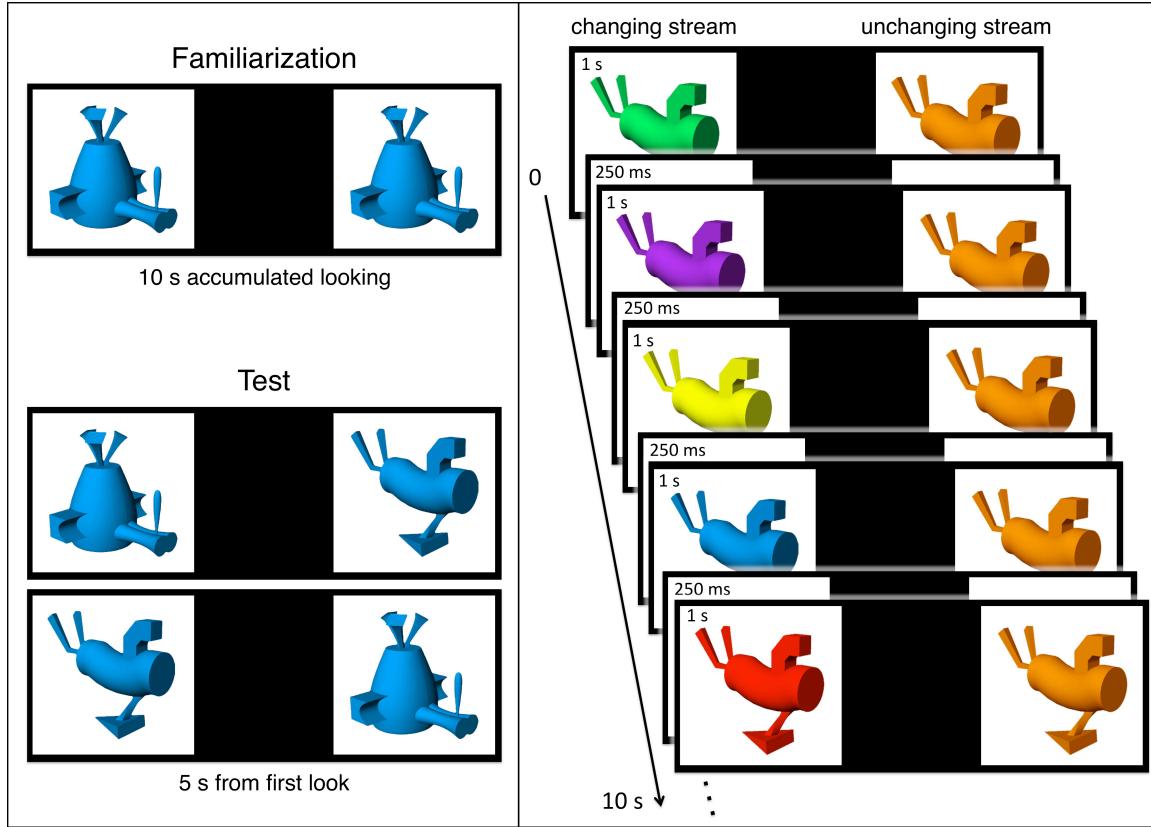


Figure 3. Examples of *between-category discrimination* (left panel) and *within-category discrimination* (right panel) conditions in the preliminary stimulus discrimination task.

novel object category (Category A or Category B) and side of the changing stream was counterbalanced across trials. Note for the present purposes the question was whether infants would discriminate between exemplars within each novel object category and therefore see the changes in the changing stream. Thus, a preference for the changing stream in this context indicates that infants can discriminate the different items comprising those changing streams.

A trained observer initiated each trial when the infant fixated an attention-getting stimulus in the center of the monitor. The observer pressed one computer key to record the duration of the infant's looking to the left stream, and another key to record duration of looking to the right stream. The main measure of discrimination was infants' preference for the changing stream over the unchanging stream, and this measure was calculated for each trial by dividing the

duration of time spent looking to the changing stream by the duration of time spent looking to both the changing and unchanging streams. A two-tailed t test indicated that infants preferred the changing stream over the unchanging stream more than would be expected by chance, $t(27) = 2.83, p < .01, d = .54$, showing that they were able to discriminate between color exemplars within each novel object category. Therefore, these visual stimuli are discriminable by 12-month-old infants.

Auditory stimuli were recorded using Praat software in a sound-attenuated room. Seven female native English speakers produced five tokens of each novel label (*lif, neem*) in infant directed speech that varied along many acoustic dimensions including pitch, intonation, and vowel length. These same speakers also produced two tokens of each familiar label (*baby, ball, car, dog, kitty, shoe*) that varied along the same dimensions. Table 2 shows the mean duration and fundamental frequency (F0) for each speaker's novel and familiar label tokens. One of the speakers (Speaker 1) produced an additional token of each label (Table 3), for a total stimulus set of 36 tokens of each novel label and 15 tokens of each familiar label.

Test stimuli. Test stimuli were identical across Experiments 1-3. A single exemplar from each object category was used during all test trials (novel objects: blue exemplar from Category A, green exemplar from Category B; familiar objects: orange² ball, baby, blue car, dog, kitty, red shoe; see Figure 1). The auditory stimuli for the test phase were a single token of each novel and familiar label, all produced by the same female, adult, native English speaker (Speaker 1). This speaker also produced two familiar sentence frames ("Where's the ___?" and "Can you find the ___?") in which the labels were embedded during test trials (see Table 3).

² Note: A different, multi-colored ball exemplar was used for pretest and familiar test trials in Experiment 1, but was replaced with the exemplar shown in Figure 1 for Experiments 2 and 3 for practical purposes regarding the creation of additional exemplars varying in color.

Table 2. *Mean Fundamental Frequency (F0) and Duration, By Speaker, of Novel and Familiar Label Tokens Heard in Experiment 2*

| Speaker | Novel Labels | | | | Familiar Labels | | | |
|------------------|--------------|------------------|-------------|------------------|-----------------|------------------|------------|------------------|
| | <i>Lif</i> | <i>Neem</i> | <i>Ball</i> | <i>Car</i> | <i>Shoe</i> | <i>Ball</i> | <i>Car</i> | <i>Shoe</i> |
| | F0 (Hz) | Duration (ms) | F0 (Hz) | Duration (ms) | F0 (Hz) | Duration (ms) | F0 (Hz) | Duration (ms) |
| <i>Speaker 1</i> | 317 (60) | 951 (715) | 284 (51) | 1122 (93) | 251 (49) | 958 (753) | 245 (46) | 970 (781) |
| <i>Speaker 2</i> | 241 (48) | 890 (95) | 250 (39) | 822 (80) | 267 (73) | 770 (104) | 244 (74) | 857 (40) |
| <i>Speaker 3</i> | 328 (25) | 827 (69) | 326 (38) | 943 (112) | 270 (10) | 685 (216) | 315 (8) | 775 (32) |
| <i>Speaker 4</i> | 285 (81) | 789 (93) | 302 (66) | 906 (76) | 273 (56) | 953 (4) | 276 (35) | 1006 (13) |
| <i>Speaker 5</i> | 280 (71) | 999 (174) | 232 (39) | 1032 (52) | 239 (70) | 851 (69) | 258 (45) | 993 (95) |
| <i>Speaker 6</i> | 345 (58) | 874 (114) | 330 (39) | 883 (170) | 258 (52) | 808 (6) | 261 (75) | 966 (60) |
| <i>Speaker 7</i> | 311 (39) | 1043 (105) | 207 (39) | 1143 (87) | 207 (52) | 1018 (94) | 215 (46) | 902 (4) |
| | | | | | | | 285 (49) | 1255 (23) |

Note. Averages were calculated across five tokens per speaker for *novel* labels and two tokens per speaker for *familiar* labels. Values in parentheses are standard deviations. Frequency and duration were calculated for each token using Praat software.

Table 3. *Fundamental Frequency (F0) and Duration of Label Tokens Heard On Habituation and Pretest Trials in Experiments 1 and 3, and Test Trials in Experiments 1-3*

| Stimulus Type | Label | F0 (Hz) | Duration (ms) |
|-----------------|-------------------------------|------------|------------------|
| Novel labels | <i>Lif</i> | 317 | 849 |
| | <i>Neem</i> | 272 | 1026 |
| Familiar labels | <i>Baby</i> | 296 | 1328 |
| | <i>Ball</i> | 289 | 855 |
| | <i>Car</i> | 285 | 920 |
| | <i>Dog</i> | 307 | 860 |
| | <i>Kitty</i> | 246 | 1154 |
| | <i>Shoe</i> | 302 | 1090 |
| Sentence frames | <i>Can you find the ... ?</i> | 219 | 1591 |
| | <i>Where's the ... ?</i> | 244 | 1320 |

Note. Speaker 1 was the voice used for habituation and pretest trials in Experiments 1 and 3, and for all test trials across experiments.

Apparatus

Infants were tested in a dimly lit experimental room. Infants sat on a parent's lap approximately 100 cm from a 93.98-cm LCD monitor. To minimize bias and infants' ability to reference where parents were looking, parents wore felt-lined glasses and headphones playing classical music for the duration of the experiment. A black curtain hanging from the ceiling with openings for the monitor and a video camera below the monitor divided the room. A trained observer sat behind the curtain and observed the infant on a monitor connected to the video camera. The observer used an Apple computer and a custom Matlab script to present the stimulus events on the LCD monitor and record the duration of infants' looking on each trial. A digital video recording of the session was stored for use in offline coding of test trials and reliability testing of online habituation coding.

Design and Procedure

The experiments used a habituation-of-looking-time procedure, one of the primary methods for investigating cognitive processes in infancy (Oakes, 2010). It is based on the basic biological finding that animals' responses to stimuli decrease (habituate) with repeated exposure (e.g., Berkson & Fitz-Gerald, 1963), and Fantz's (1964) classic finding that infants prefer to look at novel stimuli over familiar stimuli. These principles form the basis for studies utilizing an habituation-of-looking procedure, in which infants are repeatedly presented with a stimulus or stimuli until the duration of their looking decreases to a certain criterion relative to their initial looking.

The experimental design for the experiments was a variation of the *switch task*, in which infants are typically habituated two stimulus pictures or events, each characterized by a stable relation between two features (e.g., object A with label A, object B with label B). The logic is that infants may learn about individual features of the events (e.g., object A, label B), or they may learn the stable relation between features that exists in both events. In the traditional switch task, following habituation infants are presented with two test events one at a time, characterized by Werker et al. (1998) as *switch* and *same* events. These labels refer to the pairing of features. In the *switch* event, the pairing established during habituation is switched (e.g., object A with label B), and in the *same* event the pairing is preserved (e.g., object B with label B). Because all of the individual features are equally familiar to the infant in both events, the *switch* test event will be more interesting than the *same* event only if the infant has learned the association between those features. Variants of the switch task have been successfully used to uncover the development of attention to many different kinds of feature associations in infancy, including correlations between features of line drawings of animals (Younger & Cohen, 1986), relations

between object part appearance and object function (Madole & Cohen, 1995), correlations between object appearance, actions, and sounds (Baumgartner & Oakes, 2011; Perone & Oakes, 2006), and associations between words and objects (Rost & McMurray, 2010; Werker et al., 1998).

Although assessing infants' looking to individually presented switch and same stimuli on successive trials has proven to be a useful tool for studying infants' representations of feature associations, this procedure may not be the most appropriate way to assess learning in every circumstance and may underestimate infants' emerging attention to object-label associations. Specifically, infants might recognize that the *switch* event is not the same as what they learned, but the novel pairing might not be new or wrong enough (or the newly formed object-label representations aren't strong enough) to cause infants to reject the pair and dishabituate (Apfelbaum & McMurray, 2011; Yoshida, Fennell, Swingley, & Werker, 2009).

An alternative procedure for testing what infants have learned during an experimental familiarization phase that may be more sensitive to infants' learning of label-object (or auditory-visual feature) associations is the *looking-while-listening* procedure (Fernald et al., 2008). In this procedure, infants are simultaneously presented with two (or more) visual stimuli and prompted by an auditory stimulus to look at a target stimulus. In several domains, it has been shown that side-by-side presentations during test reduce memory demands and allow for active comparison between stimuli, and might therefore result in a more sensitive measure of infant learning (Oakes & Ribar, 2005; Reznick & Kagan, 1983; Younger & Furrer, 2003). Studies using looking-while-listening indicate that this procedure may allow for a more graded measure of infants' early learning of word-object associations than the switch task (Yoshida et al., 2009; Zangl, Klarman, Thal, Fernald, & Bates, 2005), a conclusion that has been supported by Apfelbaum and

McMurray (2011) in simulations of such learning. Because we were testing infants on the cusp of word learning, we used a hybrid habituation/looking-while-listening task modeled after Yoshida et al. (2009) to probe infants' learning after repeated exposure to a set of stimuli. In addition to potentially being a more sensitive test of infants' learning, the task allowed us to examine the time course of infants' looking (e.g., latency to shift to the target after the onset of a label) in addition to overall accuracy.

Prior to each trial, an attention-getting stimulus (a looming series of colorful geometric shapes accompanied by a series of attention-getting sounds) was presented in the center of the monitor. The observer initiated each trial when she determined that the infant visually fixated the attention-getting stimulus. After the observer initiated a trial, one object (approximately 19 cm² in size; 11°² visual angle) was presented moving slowly up and down against a white background in a 36-cm by 30-cm (20.5° x 17°) region in the center of the monitor. A single cycle of moving up from and back down to the starting position took 1700 ms. Each motion cycle was accompanied by a single auditory presentation of a label. Between each cycle the object disappeared for 300 ms and then reappeared. The observer pressed a key to record the duration of the infant's looking to the stimulus event. If the infant did not look to the monitor within the first 4 s of the trial, the trial terminated and was repeated.

The experiment began with two 10-s pretest trials in which a single familiar object (e.g., ball, shoe, etc.) was presented for five motion cycles, paired with five repetitions of its label at 2-s intervals. These trials were intended to highlight the referential relationship between the auditory and visual stimuli (Fennell & Waxman, 2010). In addition, infants' interest during the pretest (as measured by duration of looking) may reflect general individual differences in interest

in the task in general, or in object labeling events in particular. The familiar objects used on pretest trials were randomly selected for each participant from the set of available stimuli.

Following pretest trials, the habituation phase began. During the habituation phase, infants were familiarized to two novel object-label pairs. The specific pairing between novel object categories and labels was counterbalanced across infants (i.e., for half of the participants, Category A was labeled *lif* and Category B was labeled *neem*; for the other half of participants, Category A was labeled *neem* and Category B was labeled *lif*). On each fixed-length 14-s habituation trial, one novel object was presented for seven motion cycles, accompanied by seven repetitions of a novel label at 2-s intervals. The habituation phase continued until looking time over a non-overlapping block of four trials decreased to 50% or less of looking time over the first block of four trials, or until a maximum of 24 trials had been presented (trial length, frequency of label tokens, and maximum number of trials modeled after previous studies using a similar design, including Rost & McMurray, 2009, 2010; Werker et al., 1998; Yoshida et al., 2009). The order of habituation trials was pseudo-randomized, with each novel object-label pair presented twice in each block of four trials and no pair occurring on more than two consecutive trials. This constraint ensured that the infant was exposed to each object-label pair an equal number of times regardless of the number of trials she saw before reaching the habituation criteria.

Immediately following the habituation phase, the looking-while-listening test phase began. The test phase, modeled after Yoshida et al. (2009),³ consisted of four *novel* test trials and four *familiar* test trials, both counterbalanced for target side (the target picture appeared an equal number of times on each side and on the same side for no more than two consecutive trials). On

³ Yoshida et al. (2009) included 16 test trials (8 novel and 8 familiar), but significant results in that study were limited to first block of 4 test trials. In the interest of keeping the task as short as possible, only 8 test trials were included here.

each 7900-ms test trial, two objects were simultaneously presented in white boxes side-by-side on the monitor, with a center-to-center distance between objects of approximately 51 cm (26.5°) (Figure 4). Test trials had the following sequence: 1) two objects were presented on the screen for 1900 ms in silence (this period provided infants the opportunity to inspect both objects before hearing the label); 2) a label for one of the objects (referred to hereafter as the *target*) was presented in isolation; 3) the target label was presented again, this time embedded in a familiar sentence frame (the sentence frame began 2500 ms after the first onset of the target label, and the second target onset occurred 1615 ms after the onset of the sentence frame). For example, the entire utterance heard on a test trial would have the following format: “*Dog! Can you find the dog?*”

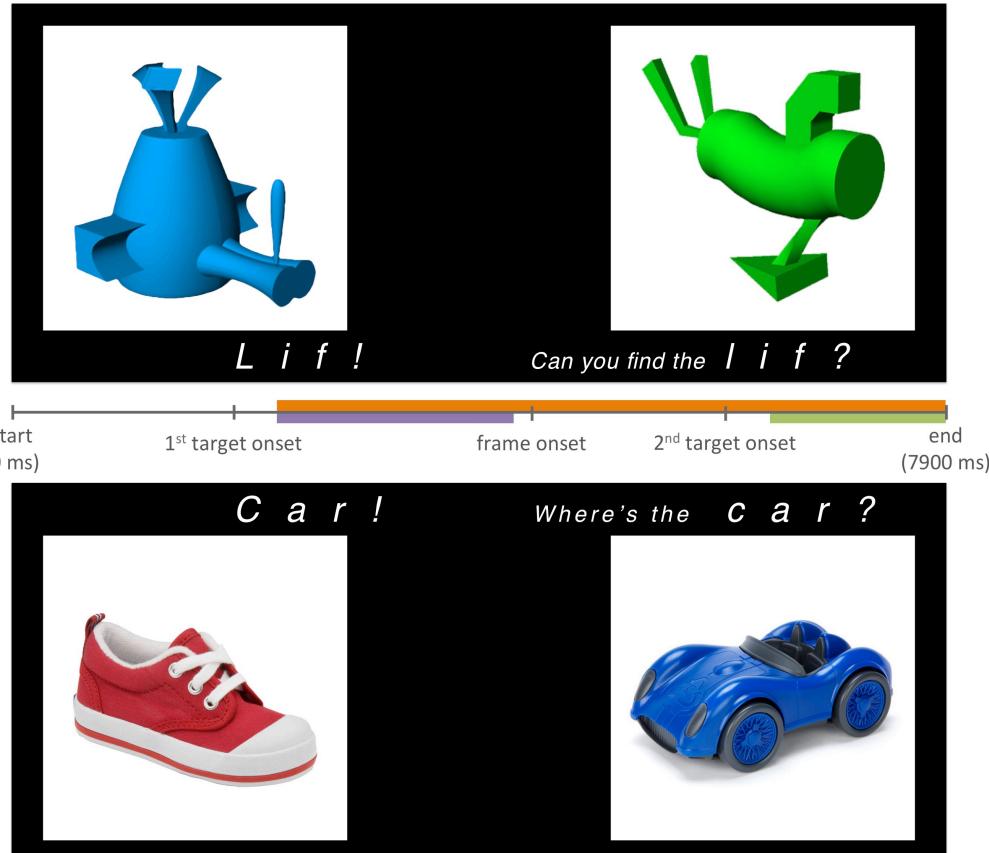


Figure 4. Examples of *novel* (upper) and *familiar* (lower) test trials. The colored bars on the timeline indicate the *whole trial* (orange), *first presentation* (purple), and *second presentation* (green) analysis windows.

The *familiar* test trials each featured two familiar objects (e.g., ball and dog) along with the label for one of the objects. All four *familiar* test trials featured a different target randomly selected from the set of six familiar objects. Distractor objects for each *familiar* test trial were randomly selected from the remaining set of stimuli, with the constraint that the label for the distractor object did not start with the same initial consonant sound as the target (e.g., trials with the ball as the target object featured a distractor that was any familiar item other than the baby). This constraint allowed for correct identification of the target object as early as the initial consonant sound for all test trials. The *novel* test trials featured the two novel objects (Category A and Category B) from the habituation phase, along with one of the novel labels. Each novel object was the target on two of the four *novel* test trials, with target side counterbalanced for each object.

The first test trial was always a *familiar* trial in order to ease the transition from the single-stimulus habituation phase to the two-stimuli looking-while-listening test phase, and to familiarize the infant with the task. The second test trial was always a *novel* trial, and the order of the remaining test trials was pseudo-randomized with the constraint that no more than two consecutive test trials were *familiar* or *novel* trials. The observer pressed a computer key to record the duration of infants' looking to the monitor during each test trial, and trials with no looking in the first 4 s were terminated and repeated (a situation which rarely occurred).

A second observer coded the habituation phase from 25% of the sessions from the digital video recording of the infants' behavior for reliability analysis. Agreement between coders for the duration of looking on each pretest and habituation trial was high, average $r > .98$, and the mean difference in the duration of looking on each trial was low, $M = .45$ s. Infants' looking to the left and right sides of the screen during the test trials was coded offline (see *Coding*, below).

Coding

Test trials were coded from the video recording by trained coders who were unaware of trial type or target location (e.g., Fernald, Perfors, & Marchman, 2006). Videos of infants gaze were recorded at a rate of 30 Hz (i.e., one sample every 33.33 ms). On each sample, the coder recorded whether the infant was looking to the left or right stimulus, shifting between stimuli, or looking away from the stimuli. A second coder coded the test trials from 22% of the sessions for reliability analysis. Two measures of inter-coder reliability were obtained. The first measure, the proportion of all samples on which the coders agreed, averaged 98.7%. A more conservative measure of reliability for the mean proportion of shifts in gaze on which coders agreed within one sample was also high, average 96.8%.

Vocabulary Measures

Prior to the study, parents were asked to complete a short vocabulary checklist in which they indicated their infants' comprehension and production of a short list of nine words that included all of the familiar words used in the experiments (see Appendix). For each word, parents indicated whether their infant 1) "does not understand, does not say," 2) "maybe understands, does not say," 3) "understands, does not say," 4) "understands, maybe says," or 5) "understands and says" the word. This checklist was designed to allow the parent to provide a graded measure of each infant's knowledge of the specific words used in the experiments.

Following the study, parents were asked to complete the MacArthur Short Form Vocabulary Checklist: Level I (Fenson et al., 2000). This 90-word checklist,⁴ designed and normed for children between 8 and 18 months, asks parents to check a box if their infant either "understands" or "understands and says" each word. While not meant to be a complete inventory

⁴ The original questionnaire contains 89 items, but the word *baby* was added so that all familiar words used in the experiments appeared on the questionnaire.

of receptive or productive vocabulary, this checklist provides a more general and standardized measure of each infant's relative level of word comprehension and production.

Data Processing and Analysis

To maximize the likelihood that results reflect infants' attention to visual stimuli in relation to labels, we used the following inclusion criteria. Test trials were included if: 1) the infant looked to the stimuli for at least 50% of the total test period beginning 367 ms after the first onset of the target word, 2) the infant looked to at least one of the stimuli during the periods following both presentations of the target word, and 3) the infant looked to at least one of the stimuli prior to the first presentation of the target word. These criteria are consistent with other studies using the same procedure (e.g., Fernald et al., 2006), and the application of these criteria resulted in the exclusion of 124 trials (out of a total 770 trials). In addition, we excluded any infants who contributed fewer than two trials from each test condition (*familiar* and *novel*) from the final analysis.

Because our primary measure of interest was infants' preference for the labeled target relative to a distractor rather than a preference for novelty, our main analyses included infants who did and did not reach the habituation criterion. However, because it is possible that infants who do and do not reach the habituation criteria have acquired different levels of learning about the stimuli (see Oakes, 2010 for a discussion; see Werker et al., 1998 for an example), we also conducted supplemental analyses on just those infants who met the habituation criteria.

Target preference analysis. One set of analyses evaluated infants' looking time for overall target preference (i.e., their tendency to look at the object that was the referent of the target label). Target preference was computed from the frame-by-frame offline coding by dividing the time spent looking to the target object by the time spent looking to both the target

and distractor objects combined. To allow infants time to process the initial consonant sound of the label and execute an eye movement, analysis windows began 367 ms after the onset of target words, consistent with other studies of target recognition in infancy (e.g., Fernald et al., 2008).

Target preference was computed over three different time windows: 1) *whole trial*: 367 ms after the first onset of the target label to the end of the trial; 2) in response to the *first presentation* of the label: starting 367 ms after the first onset of the target label (presented in isolation) and ending 2000 ms later; and 3) in response to the *second presentation* of the label: starting 367 ms after the second onset of the target label (embedded in a sentence frame) and ending 1500 ms later at the end of the trial.

Time course analysis. One major advantage of the looking-while-listening procedure is the high resolution, time-locked nature of the data. In addition to calculating infants' average preferences over windows of time, as is traditional in many studies using infants' looking behavior as a dependent measure, we can examine the temporal dynamics of infants' preferences over the course of the trial.

To analyze infants' looking over the course of a trial, the time course of infants' looking during test trials was analyzed with a Monte Carlo simulation analysis using customized Matlab scripts to test the null hypothesis that infants are equally likely to look at both target and distractor stimuli after a labeling event (Oakes, Baumgartner, Barrett, Messenger, & Luck, 2013). Under this null hypothesis, looks to the target and distractor are interchangeable within a trial (e.g., the target stimulus could be on the left or the right). To test the null hypothesis, looks to the target or distractor were randomly permuted within a trial and the proportion of looks to either side at each sample was calculated for each subject and group of subjects on each of 1000 permutations. Null distributions of the following statistics were estimated: proportion of looking,

F-statistic of the difference of proportion of looking from chance (0.5), number of consecutive samples with a proportion of looking significantly different from chance (run length), and mass of a significant run of samples (sum of squared t values of samples in a run compared to chance). The 95% cutoff of each of these statistics was calculated and applied to the observed time course of looking to identify periods of looking to the target that were significantly different from chance. Using t tests to compare looking to the target at each sample would require controls for multiple comparisons that would likely result in too conservative a test. The use of the Monte Carlo simulation analysis allowed us to control for multiple comparisons without losing the power to detect significant differences from chance.

Experiment 1

Experiment 1 was designed to assess infants' learning and recognition of object-label associations in the absence of stimulus variability. It is predicted that 12-month-old infants will have difficulty both identifying referents of familiar words and forming associations between novel words and objects in this task, consistent with previous studies of word learning at this age (e.g., Fernald et al., 2008; Werker et al., 1998). Thus, Experiment 1 will provide a baseline measure of 12-month-old infants' general ability to match labels to visual targets, as well as provide an opportunity to look for relations between individual differences in responding and other indices of development such as age and vocabulary size. We will then be able to investigate how changes to the learning context in the form of increased acoustic (Experiment 2) and visual (Experiment 3) stimulus variability affect infants' behavior. The results from this experiment will also be used to examine how individual differences in attentional abilities might be related to infants' ability to respond to novel object-label associations (Experiment 4).

Method

Participants. Participants were 34 12-month-old infants ($M = 378.2$ days, $SD = 6.8$ days, range = 367-392 days, 18 females). Eighteen infants were white, 2 were Asian, 2 were black/African American, 11 were mixed race, and 1 infant's race was not reported. Across these groups, 5 infants were Hispanic. All mothers had graduated from high school, and 27 had at least a Bachelor's degree. An additional 19 infants were tested but excluded from the final analysis for the following reasons: fussiness or lack of interest ($n = 12$), parental interference ($n = 2$), equipment error ($n = 1$), or fewer than two test trials from each condition that met the inclusion criteria ($n = 4$).

Stimuli. Visual stimuli were a single exemplar from each object category described in the General Method section above (Figure 1). Pretest stimuli included a single exemplar from each of 6 familiar object categories (baby, ball, car, dog, kitty, shoe). As noted in the General Method section, a different ball exemplar was used for pretest trials and familiar test trials in this experiment than the ball pictured in Figure 1. The ball used in this experiment was a multi-colored beach ball, but was replaced for Experiments 2 and 3 with a monochromatic ball more suitable for color editing. Habituation stimuli were a blue exemplar from novel Category A and a green exemplar from novel Category B. Auditory stimuli were a single token of each familiar and novel label, produced by a single female speaker (Speaker 1; see Table 3). Pretest and habituation stimuli were identical to stimuli used for test trials.

Procedure and design. The procedure and design of this study are as described in the General Method section. Importantly, infants in this experiment saw only a single exemplar from each object category and heard only a single token of each label throughout the duration of the study, with all labels produced by a single speaker (Table 4). On each 10-s pretest trial, infants

Table 4. Number of Object Exemplars and Label Tokens Used on Each Habituation Trial and Over the Course of the Habituation Phase for Experiments 1-3

| Experiment | Object exemplars (per object) | | | Label tokens (per label) | | |
|------------|-------------------------------|--------------------|-------|--------------------------|-------------|-------|
| | Per trial | Per color category | Total | Per trial | Per speaker | Total |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 1 | 1 | 7 | 4 | 28 |
| 3 | 7 | 4 | 28 | 1 | 1 | 1 |

saw a single exemplar from a familiar object category paired with five repetitions of a single token of its label produced by Speaker 1. On each 14-s habituation trial, infants saw a single novel object exemplar presented seven times, paired with seven repetitions of a single novel label token, also produced by Speaker 1. On half of the habituation trials an infant would see the blue exemplar from Category A paired with a single novel label token (e.g., *lif*), and on the other half of habituation trials the infant would see the green exemplar from Category B paired with a single token of the other novel label (e.g., *neem*). Thus, in this experiment, infants could learn that a specific object was associated with a specific label as spoken by a specific speaker. The test trials involved the same object exemplars and label tokens seen and heard during pretest (the familiar items) and habituation (the novel object-label pairings) trials.

Data analysis. Across all included participants three *familiar* test trials and 24 *novel* test trials were excluded using the criteria described earlier. The higher rate of excluded *novel test* trials relative to *familiar* trials is not surprising, given that the stimuli seen on these trials were identical to those seen throughout the habituation phase. On average, infants contributed 3.91 *familiar* test trials and 3.29 *novel* test trials to the analysis.

Target preference scores for *familiar* and *novel* test trials were calculated for each infant by averaging across trials within each condition. Group averages were then calculated for each test condition by averaging across subject averages.

Results

Nineteen of the 34 infants met the habituation criteria, requiring an average of 17.5 trials to do so. The remaining 15 infants did not meet the habituation criteria and saw the maximum number of habituation trials (24). Although this group did not meet the habituation criteria, they did show a significant decrease in looking from the first habituation block ($M = 43.1$ s) to the last habituation block ($M = 32.1$ s), $t(14) = 6.29, p < .001, d = 1.62$. Initial analyses revealed no significant differences between infants who did and did not meet the habituation criteria on any of our outcome measures of interest, so our main analyses are collapsed across all infants, regardless of habituation status.

Our primary interest was if infants looked more to the labeled target object than to the distractor object on either *familiar* or *novel* trials. Looking more at the target on the *familiar* test trials would indicate that the task is sensitive to infants' word comprehension, even at 12 months (assuming that infants actually know the words, an issue that will be addressed in subsequent analyses). Looking more at the target on *novel* test trials would indicate that infants learned the new object-label associations that were taught during habituation.

First, we compared infants' preference for the target on each type of trial to determine if they preferred the target more than would be expected by chance (50%). Recall that a greater than chance preference for the labeled object is taken as evidence that infants recognized the target-label correspondence. Preference scores for both test conditions in all three analysis windows (whole trial, first presentation, second presentation) were compared to chance using

one-sample t tests (all tests two-tailed). Infants' preference for the target object did not differ from chance on either *familiar* or *novel* test trials in any of the three analysis windows, $ts < 0.63$, $ps > .53$ (Figure 5). Thus, infants did not appear to recognize the correspondence between the target object and the label for either type of trial. Because it is possible that infants who did and did not reach the habituation criteria performed differently on test trials, we also conducted supplemental analyses on just those infants who met the habituation criteria. Comparison of the target preference scores for this subset of infants revealed identical results; the infants who habituated also showed no evidence of preferring the target on either *familiar* or *novel* test trials, $ts < 0.89$, $ps > .38$.

Next, to gain deeper understanding into infants' moment-to-moment changes in target preference we conducted an analysis of the time course of infants' looking on *familiar* and *novel*

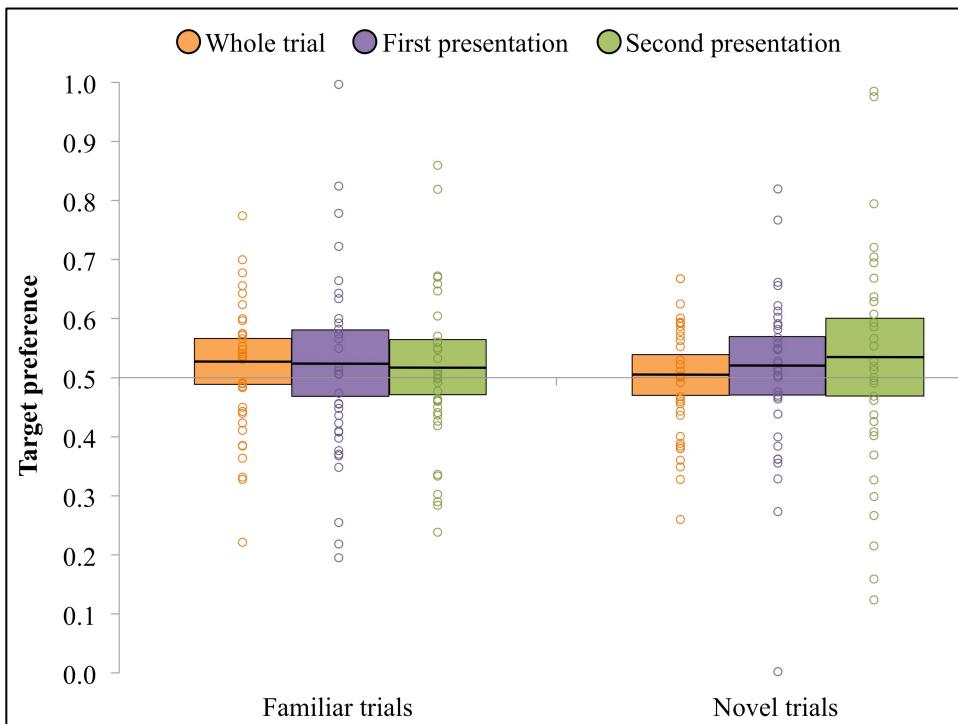


Figure 5. Experiment 1 *familiar* and *novel* test trial target preference. Target preference is defined as proportion of total looking to target and distractor objects spent looking to the target object. Each circle is the average target preference of an individual participant, boxes indicate 95% confidence intervals, and the horizontal line bisecting each box is the mean.

test trials. Although the window analysis allows us to compare infants' preference early and late in the trial, that analysis still involves collapsing over periods of time to calculate an average target preference. The time course analysis, in contrast, examines infants' target preference *on each observed sample* (i.e., 30 times per s) and thus gives a more fine-grained view of infants' behavior. The time course analysis in this experiment confirmed the results of the window analysis. Figure 6 depicts the target preference for each sample (with a line at .5 to illustrate whether this preference was above or below chance); it can be seen that infants' preference for

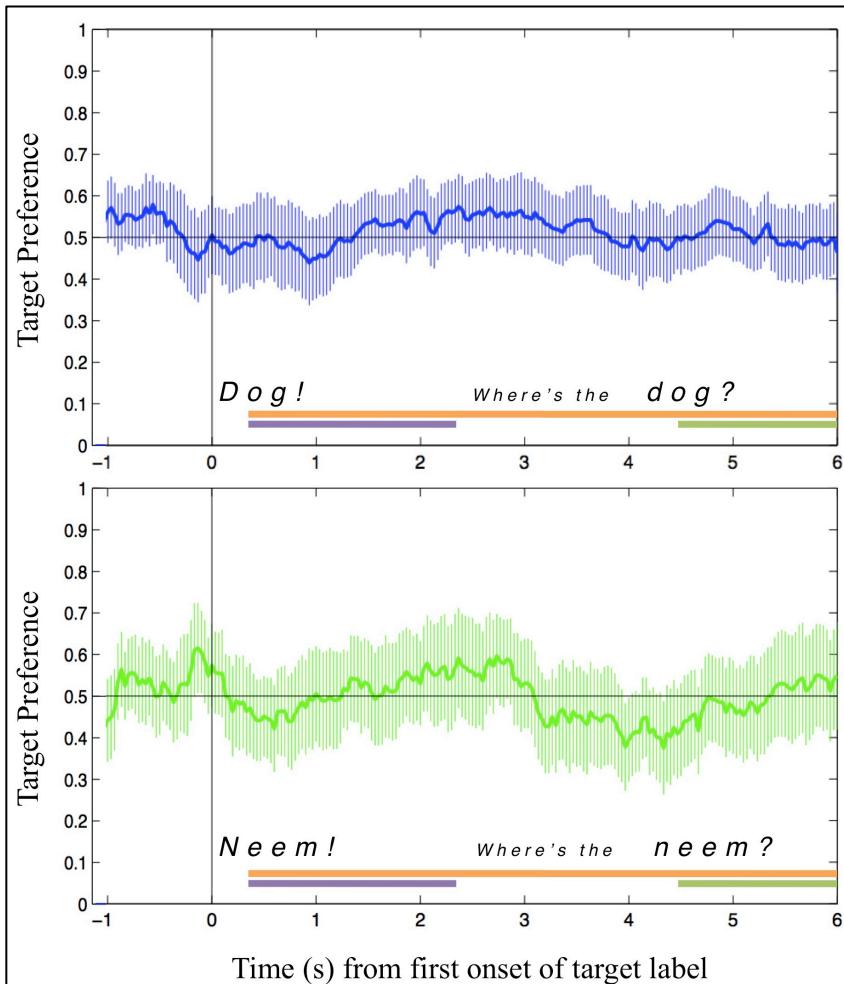


Figure 6. Time course of infants' preference for the target on *familiar* (upper) and *novel* (lower) test trials in Experiment 1 (1 = 100% of all looks aggregated across subjects were to the target; 0 = 0% of looks were to the target). The curve represents mean target preference at each sample and vertical lines represent the 95% confidence interval for each sample. Color bars along the x-axis indicate timing of analysis windows (orange: *whole trial*; purple: *first presentation*; green: *second presentation*).

the *familiar* and *novel* targets did not differ significantly from chance at any point over the course of the trial.

Finally, we asked whether infants' target preference differed in the two conditions, or if their preferences changed across the trial. Even if infants do not prefer the target more than expected by chance, it is possible that their preference for the target on (for example) familiar trials is greater than that on novel trials, indicating that they had a stronger bias to look at the target on those trials. Similarly, infants' target preference may increase or decrease over time. We addressed these questions by conducting an ANOVA on target preference scores with Window (first presentation, second presentation) and Test Condition (familiar, novel) as within-subjects factors. This analysis revealed no significant main effects or interactions, $Fs < 0.2$, $ps > .6$, indicating that target preference did not differ based on analysis window or test condition. The identical analysis on just the subset of infants who habituated also did not reveal any significant effects or interactions. Thus, infants did not respond differently to either type of trial or for either window of the trials.

The lack of a target preference on the *novel* trials just reported is taken to indicate that infants failed to learn the association between novel words and objects. However, infants also failed to show a consistent preference for the target on *familiar* trials, suggesting that 12-month-old infants do not respond to *any* label-object correspondences in this task. To increase the likelihood that infants would actually know these familiar label-object correspondences, these items were chosen based on vocabulary norms for 12-month-old infants. However, not all infants in our sample were reported by a parent to know all (or any) of the labels heard on *familiar* test trials. Thus, infants' responding in general on test trials may not be the best indicator of how sensitive this task is to infants' recognition of object-label mappings. As a follow-up, we asked

whether infants preferred the target on the subset of *familiar* trials for which parents reported that the words were understood by the infant. We classified the trials using the parent ratings on the short vocabulary questionnaire, and determined which *familiar* test trials fell within each infant's receptive vocabulary by identifying the trials that depicted object-label associations for words that parents gave a rating of 3 or higher. These trials will be subsequently be referred to as *known* trials. Twenty-six infants had at least one *known* trial, and an analysis of these infants' target preference on these *known* trials (e.g., preference for the ball over a distractor object when the label *ball* was heard, for those infants who parents reported they understood the word *ball*) revealed preference scores that still did not differ from chance in any of the three analysis windows, $ts < 0.98$, $ps > .33$ (Figure 7). This subset of infants also did not prefer the target on *novel* test trials, $ts < 1.16$, $ps > .26$. Analyses including only trials on which the infant was reported to "understand, maybe say" or "understand and say" the target word (a score of 4 or 5

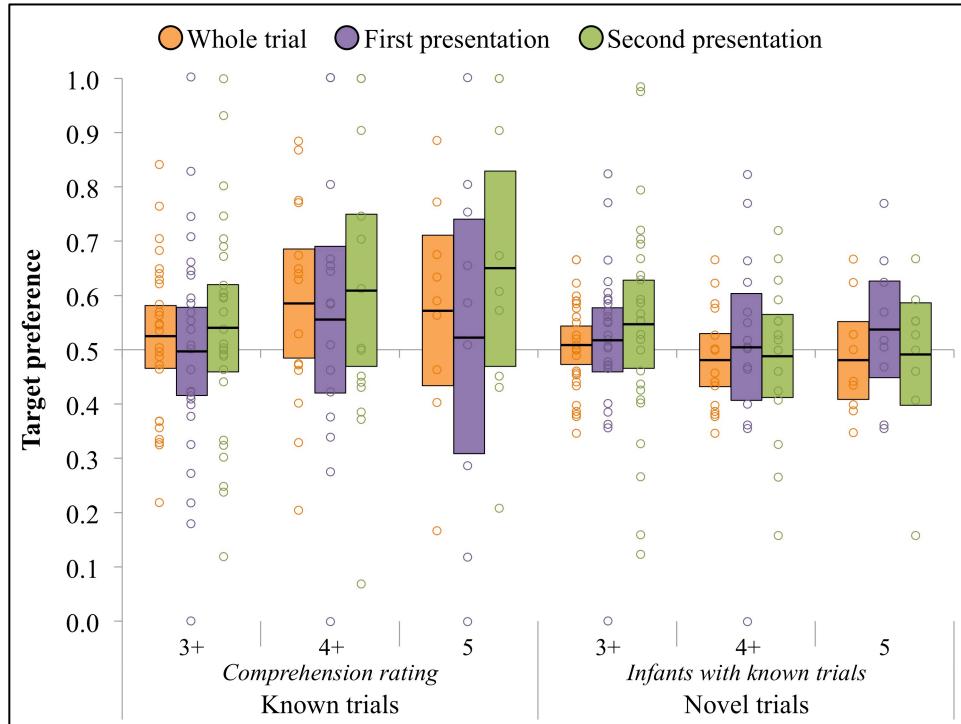


Figure 7. Experiment 1 target preference scores for *known* trials, divided by strength of parental rating of target words on the vocabulary questionnaire, and corresponding *novel* target preference scores from the same subset of participants (3+, $n = 26$; 4+, $n = 15$; 5, $n = 9$).

on the vocabulary questionnaire, respectively), revealed preference scores that were still not significantly different from chance, $p > .11$. Thus, when target preferences were averaged across the whole trial or within each pre-defined analysis window, infants did not show evidence of matching known labels to their referents, raising the possibility that this task is simply not sensitive enough to assess very young infants' linguistic knowledge and that infants' lack of a target preference on *novel* trials does not necessarily indicate that they failed to learn those items.

However, an analysis of the moment-to-moment time course of infants' looking on *known* trials (i.e., trials with a target word rating of 4 or higher on the vocabulary questionnaire) shows that infants did preferentially look to the target object more than would be expected by chance for a short period approximately 2 s after the first onset of the label (Figure 8). This period of significant preference for the target falls largely outside of the *first presentation* window (as indicated by the purple bar along the x-axis), and thus was not detected by the window analyses reported previously. The windows chosen for analysis were based on previous work using this procedure (Fernald et al., 2006; Swingley & Aslin, 2002; Yoshida et al., 2009), however those previous studies tested 14- to 15-month-old infants. It is therefore possible that the 12-month-old infants tested here are slower to selectively prefer the referent of the target and our chosen window underestimated infants' recognition of the label-object correspondence. In other words, the time course analysis provides some evidence that infants in this experiment did successfully identify referents of known words, but that they were slow to do so.

In addition, although window analyses indicated that infants as a group did not show a preference for the labeled object on either *familiar* or *novel* test trials, there was considerable variability in responding between participants and some infants clearly showed a preference for the target. Inspection of the responding of the individual infants shown in Figure 5 reveals that

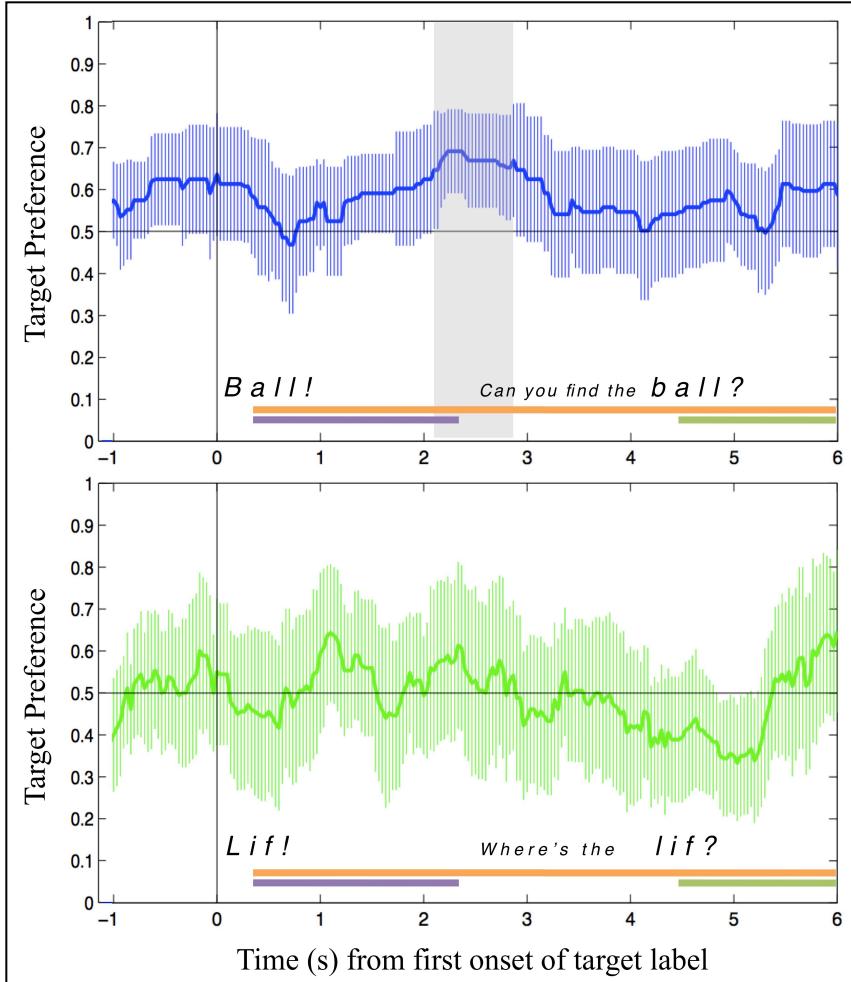


Figure 8. Time course of infants' ($n = 15$) preference for the target on *known* (upper) and *novel* (lower) test trials in Experiment 1. *Known* trials were the subset of *familiar* test trials for which individual infants were reported by a parent to understand the target word (i.e., a rating of 4 or higher on the vocabulary questionnaire). The curve represents mean target preference at each sample and vertical lines represent the 95% confidence interval for each sample. Color bars along the x-axis indicate timing of analysis windows (orange: *whole trial*; purple: *first presentation*; green: *second presentation*). The shaded area marks a run of samples that pass the 95% run mass criterion.

although the group means were close to chance (.5), many infants' individual mean responding differed considerably from chance, with several infants showing a preference for the *non-target* (i.e., scores below .5) and many infants showing a preference (in some cases a very strong preference) for the target. The question this raises is whether this variability is systematic and related to other indices of development or attention.

We addressed this question in several ways. We first asked if *familiar* or *novel* target preference was related to age or overall vocabulary. Within the narrow one-month age range of our sample, we found no correlation between age in days and target preference on either type of test trial, with the exception of an unexpected marginally significant negative correlation between age and *familiar* target preference for the *whole trial*, $r = -.33$, $p = .057$ (see Figure 9). This correlation between age and *familiar* target preference was present for the *second presentation* window, $r = -.34$, $p = .05$, but not the *first presentation* window, $r = -.21$, $p = .24$, suggesting that infants' late-trial preference for *familiar* word targets decreases with age. It is possible that this result reflects younger infants' slower orienting to the target relative to older infants, and/or older infants orienting toward the distractor object after identifying the target earlier in the trial. This is not inconsistent with other studies that have used the looking-while-listening procedure and have found interactions between preference scores and analysis window timing/length (see Fernald et al., 2008 for a discussion). There was no correlation between *familiar* or *novel* target preference and overall receptive vocabulary as measured by the MSFVC-I, $rs < .15$, $ps > .42$, nor was there any correlation between age and receptive vocabulary in this sample, $r = .06$, $p = .74$. In general, therefore, there was minimal evidence of any relation between age and/or vocabulary on learning in this task.

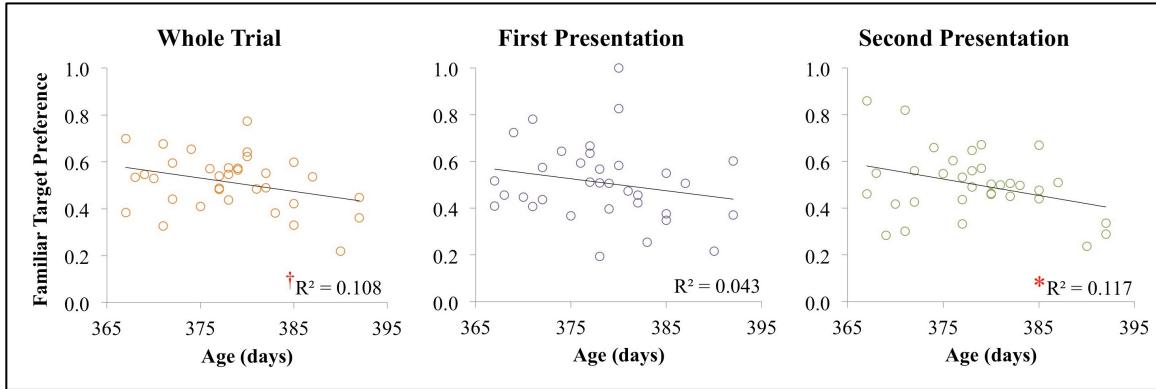


Figure 9. Scatterplots showing correlation between age and target preference on *familiar* test trials in three analysis windows in Experiment 1. $\dagger p < .10$, $* p < .05$.

Next, we examined the possibility that variation in target preference was related to infants' understanding of the task in general. For example, infants who experienced pretest trials with known words showed might have shown different patterns of learning during habituation than did infants who did not experience known words during pretest. Recall that during pretest infants saw a familiar object and heard it labeled five times (e.g., a picture of a ball accompanied by repetitions of the word *ball*). Pretest trials were designed to familiarize infants with the general goal of the task—to label visually presented objects. This manipulation may have been more effective for infants who received pretest trials that better matched their vocabulary. Knowing one or more of the words heard during pretest might have made it easier for infants to recognize the referential nature of the task. If so, these infants will have had an additional cue that the words heard during habituation were labels for the objects presented on the screen. We identified a subset of participants ($n = 22$) who were reported by a parent to understand at least one of the two words heard on pretest trials (as indicated by a rating of 3 or higher on the vocabulary checklist), and repeated the analyses for this group. The pattern of results was the same, with no preference scores differing from chance for either *familiar* or *novel* test trials, $ts < 1.72$, $ps > .10$. Thus, it appears that exposure to known labels during pretest seemed to have no effect on infants' learning of novel words or recognition of familiar words in this task.

We also examined the relation between infants' responding during *familiar* and *known* trials and *novel* trials. It is possible that the variability reflected in Figure 5 reflects infants who show stronger target preference on *known* or *familiar* trials also showing stronger target preference on *novel* trials. However, there was no correlation between infants' target preference on *familiar* or *known* trials and *novel* trials, $rs < .17$, $ps > .39$.

Finally, we asked whether this variability reflected infants' engagement in the task in general. To assess this possibility, we examined the relation between infants' duration of looking (or general interest) during pretest and their target preference (or measure of learning) on test trials. Interest during these pretest trials might reflect a general engagement in the task or an interest in object labeling events in particular. However, we found no correlation between the overall duration of looking on pretest trials and target preference on test trials, $rs < .15$, $ps > .42$, indicating that infants' ability to identify the target in this first experiment was not related to their general interest or length of looking during the pretest.

Discussion

The results of Experiment 1 seem to provide no evidence that these 12-month-old infants learned the object-label associations presented during habituation. Infants in this study showed no preference for the target object on *novel* test trials. This effect was observed for the group as a whole as well as for the subset of infants who met the habituation criterion, for infants whose experienced *known* words during the familiar test trials, and during all three analyses windows. However, infants' also did not show a preference for the target on *familiar* test trials. That is, not only did they not show evidence of learning associations between the completely novel labels and objects presented during habituation, they also did not look more to familiar objects after hearing them labeled. Even for the subset of trials for which parents reported that the infant knew the label with a high degree of confidence, the only evidence of target recognition came from an analysis of the time course of infants' looking, in which infants showed a slow-developing preference for the target that emerged over 2 s after the onset of the label.

These results are not completely surprising, given the difficulty of the task for infants of this age. Few laboratory studies using looking-while-listening or other procedures like the switch

task have demonstrated consistent group-level word recognition or word-referent mapping skills in infants younger than 14 months of age (Bergelson & Swingley, 2012; MacKenzie, Curtin, & Graham, 2012; Smith & Yu, 2008), even though on average parents report infants to have a receptive vocabulary of over 50 words by this age (Fenson, Dale, Reznick, Bates, & Thal, 1994). There are a number of possible explanations for 12-month-old infants' failure on word recognition tasks that have been discussed in the literature (e.g., Bergelson & Swingley, 2012; Fernald et al., 2008). Infants' linguistic skills might not yet be powerful enough to support word recognition or word-referent association learning in a relatively artificial laboratory setting. Another possibility is that young infants' word-referent mappings, both for familiar words heard in the real world and for novel words learned during the experiment, are too fragile or limited to allow them to demonstrate their knowledge through their looking behavior on test trials. And, of particular interest here, it is possible that young infants' failure to demonstrate a consistent preference for the target stems from an interaction between task demands, properties of the learning environment, and limits on attention in infancy.

In the following two experiments we experimentally manipulated aspects of the learning context, in particular the level of variability in the stimuli, to establish whether or not increasing variability in the features of the task would highlight the invariant relation between the objects and the labels and therefore make it easier for infants to attend to and learn the relevant associations. In Experiment 4, described in Chapter 3, we explored the relation between infants' selective attention and their responding in this task to determine whether individual differences in learning reflected systematic individual differences in infants' selective attention.

Experiment 2

In Experiment 1, we attempted to teach 12-month-old infants two novel object-label associations in the absence of stimulus variability. Infants were repeatedly exposed to a single token of each label paired with a single exemplar of each object, and then asked to identify the referent of each label during test. As predicted, infants in Experiment 1 showed no evidence of having formed associations between the novel labels and objects. They were also unable to identify referents of common familiar labels. Experiment 2 was designed to assess whether hearing multiple speakers produce a label for an object facilitates infants' ability to recognize previously-known and/or recently-learned object-label associations.

Method

Participants. Participants were 27 12-month-old infants ($M = 381.5$ days, $SD = 6.6$ days, range = 370–392 days, 14 females). Twenty-two infants were white, 1 was Asian, 3 were mixed race, and 1 infant's race was not reported. Across these groups, 5 infants were Hispanic. All mothers had graduated from high school, and 23 had at least a Bachelor's degree. An additional 29 infants were tested but excluded from the final analysis for the following reasons: fussiness ($n = 13$), parental interference ($n = 3$), equipment error ($n = 2$), parent choosing to end the study early ($n = 2$), test trials not codable because eyes were not visible ($n = 1$), or fewer than two test trials from each condition that met the inclusion criteria ($n = 8$). It is interesting to note that we observe here a higher attrition rate in Experiment 2 (52%) than in Experiment 1 (36%); a chi-squared test indicates that these rates are different than would be expected by chance, $\chi^2 = 13.1$, $p < .001$. This higher rate of attrition, however, does not appear to be due to a difference in rates of infant fussiness in the two experiments. We observed a 23.2% rate of fussiness in this experiment (13 out of 56 total participants), nearly identical to the 22.6% rate of fussiness

observed in Experiment 1 (12 out of 53 total participants). Even when other indicators of interest (e.g., fewer than 2 usable trials per condition, eyes not visible for coding) are included, rates of attrition due to infant interest in the task did not differ by study, $\chi^2 = 0.99, p = .32$. Still, this attrition rate is higher than normally observed in our laboratory for studies with infants of this age, raising the possibility that the added stimulus variability relative to Experiment 1 made this a more difficult task for infants to complete. The possible causes and implications of this higher rate of attrition will be considered in the Discussion.

Stimuli. Visual and auditory stimuli for test trials were identical to those used in Experiment 1.

Visual stimuli for pretest and habituation trials were identical to those used in Experiment 1 (with the exception of the ball exemplar, as previously noted), and consisted of a single exemplar from each familiar and novel object category. Pretest stimuli were limited to the ball, car, and shoe object categories for maximum consistency with Experiment 3.

Label tokens heard on test trials (and on pretest and habituation trials in Experiment 1) were removed from the set of auditory stimuli available for pretest and habituation trials. Five tokens (one token produced by each of five different speakers) of each familiar label (*ball, car, and/or shoe*) were randomly selected for each infant from the remaining set of 14 label tokens per object for use on pretest trials. Four tokens from each of seven different speakers (for a total of 28 tokens per label) were randomly selected for each infant from the set of 35 tokens per label for use on habituation trials (Table 4).

Procedure and design. The procedure of this study is as described in the General Method section.

Objects for the two pretest trials were randomly selected for each infant. On each 10-s pretest trial, infants saw a single exemplar from a familiar object category (ball, car, or shoe) and heard five unique tokens of its label, each produced by a different speaker.

On each 14-s habituation trial, infants saw a single exemplar from a novel object category (Category A or Category B) and heard seven unique tokens of its label (*lif* or *neem*) (Table 4). Tokens heard on each trial were randomly selected from the set of 28 tokens per label with the following constraints: 1) each of the seven tokens heard on a given trial were produced by different speakers, and 2) tokens were not repeated until all tokens had been used. Therefore, all infants heard all 28 tokens of each novel label once in each set of eight habituation trials (four trials per object category). This ensured that all infants were exposed to the same number of unique label tokens regardless of the number of habituation trials they received (minimum of eight, maximum of 24).

Test trials were identical to Experiment 1. Importantly, in contrast to the habituation stimuli, all the test trials involved only Speaker 1 (i.e., only one speaker was heard on these trials).

Data analysis. Across all included participants, seven *familiar* test trials and 22 *novel* test trials were excluded. On average, infants contributed 3.74 *familiar* test trials and 3.19 *novel* test trials to the analysis.

Results

Ten infants met the habituation criteria, requiring an average of 18 trials to do so. The remaining 17 infants did not meet the habituation criteria and saw the maximum number of habituation trials (24). On average, the infants who did not habituate also did not show a significant decrease in their duration of looking from the first habituation block ($M = 41.7$ s) to

the last habituation block ($M = 37.8$ s), $t(16) = 1.68$, $p = .11$, $d = .41$. Although a smaller percentage of participants met the habituation criteria in this experiment (37%) than in Experiment 1 (56%), a chi-squared test indicated that the proportion of infants who met the habituation criterion did not differ by experiment, $\chi^2 = 2.14$, $p = .14$. However, the lower rate of habituation combined with the higher rate of attrition observed here is consistent with the conclusion that the increased variability in the stimuli made Experiment 2 a more difficult task for infants than Experiment 1. Nonetheless, preliminary analyses comparing the habituators and non-habituators in Experiment 2 revealed no effect of habituation status on target preference scores, so our main analyses are collapsed across all infants regardless of habituation status.

As in Experiment 1, our first analyses was to calculate mean target preference scores for the *whole trial*, *first presentation*, and *second presentation* analysis windows for both *familiar* and *novel* test trials, and then compare those scores to chance using one-sample t tests (Figure 10). In contrast to Experiment 1, here we observed target preferences on *familiar* test trials that were greater than would be expected by chance for the *whole trial* following the first onset of the target word, $t(26) = 2.97$, $p = .006$, $d = .57$, and in the *first presentation* window, $t(26) = 3.88$, $p = .001$, $d = .75$. Infants' target preference was not different from chance in the *second presentation* window, $t(26) = 1.18$, $p = .25$, $d = .23$. The implication is that because these *familiar* test trials are identical to the *familiar* test trials of Experiment 1 (on which infants showed no significant preference for the target object even for known words), something about the inter- and intra-speaker variability in label tokens heard during the pretest and habituation trials of this experiment affected infants' ability to match familiar labels to referents.

As in Experiment 1, infants' preference for the target on *novel* test trials did not differ from chance over the course of the *whole trial*, $t(26) = 0.77$, $p = .45$, in the *first*

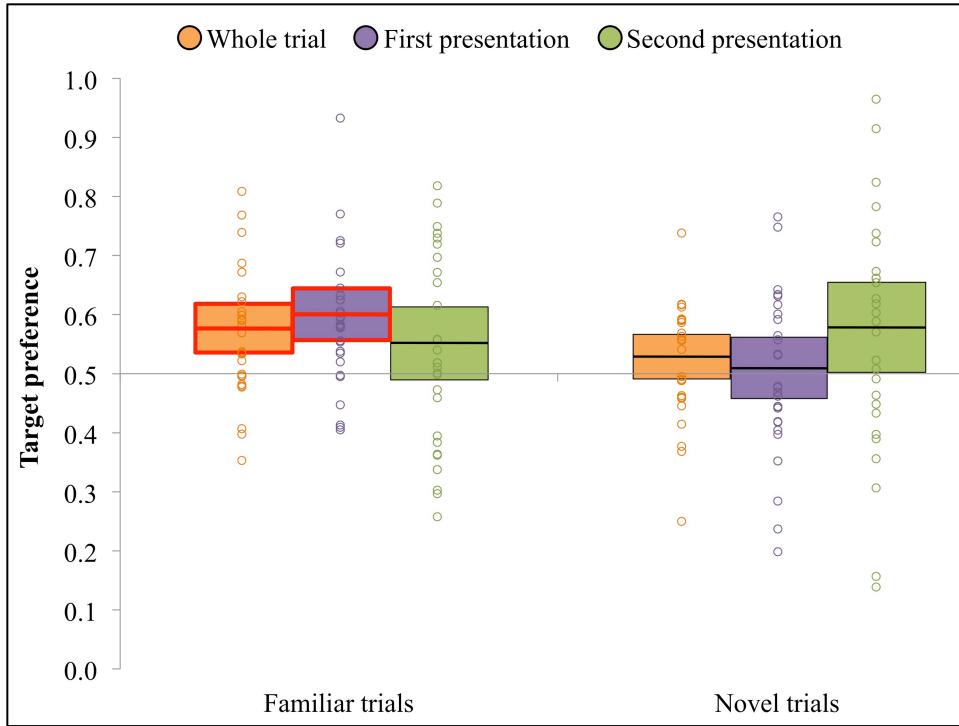


Figure 10. Experiment 2 *familiar* and *novel* test trial target preference. Target preference is defined as proportion of total looking to target and distractor objects spent looking to the target object. Each circle is the average target preference of an individual participant, boxes indicate 95% confidence intervals, and the horizontal line bisecting each box is the mean. Means different from chance (50%) are indicated by a bold red line.

presentation window, $t(26) = 0.19$, $p = .85$, $d = .04$, or in the *second presentation window*, $t(26) = 1.65$, $p = .11$, $d = .32$. Thus, although the increased stimulus variability in the form of labels produced by multiple speakers apparently enhanced infants' target preference on *familiar* trials, it did not result in a significant target preference on *novel* trials.

To confirm that these results were not driven by infants who met the habituation criterion, we conducted separate supplemental analyses on the groups of habituators and non-habituators. These results confirmed that both groups of infants showed a significant preference for the target on *familiar* test trials in the *first presentation window* (habituation: $t(9) = 3.58$, $p = .006$, $d = 1.13$, non-habituators: $t(16) = 2.82$, $p = .01$, $d = .68$), but not the *second presentation*

window, $ts < 1.09$, $ps > .29$. Target preference for the *whole trial* was significantly above chance for non-habituators, $t(16) = 2.28$, $p = .04$, $d = .55$, and marginally significant for habituators, $t(9) = 2.05$, $p = .07$, $d = .65$ (this marginal effect is likely due to small sample size). Also, neither habituators nor non-habituators showed a preference for the target on *novel* test trials in any of the three analysis windows, $ts < 0.81$, $ps > .44$. Thus, infants' ability to demonstrate word knowledge on test trials appears to be unrelated to whether or not they met the habituation criteria in this study.

The time course of infants' looking on *familiar* and *novel* trials can be seen in Figure 11, and confirms the comparisons of the mean preferences to chance. The curve indicating infants' mean preference for the target on *familiar* test trials is above 50% at nearly every point in the trial. There are also multiple individual samples for which infants' target preference was greater than chance (as indicated in the figure by samples for which the vertical 95% confidence interval bar does not cross 0.5). Although the observed pattern of looking did not meet the conservative threshold of statistical significance for number of consecutive samples above chance set by the Monte Carlo analysis (18 samples), it is clear that infants looked more to the target object than the distractor on *familiar* trials, especially during the first 3 s following the first onset of the label.

Infants' looking on *novel* test trials, however, was more variable, with mean preferences at or near chance for most samples. Consistent with the results of the window analyses, infants did not preferentially look to the target in either period following the presentations of the label on *novel* trials, and looking over the course of the trial did not meet our thresholds of significance in the permutation analysis. There is one potentially interesting element of the time course of infants' looking on novel trials that was not evident from the window analyses. In the period

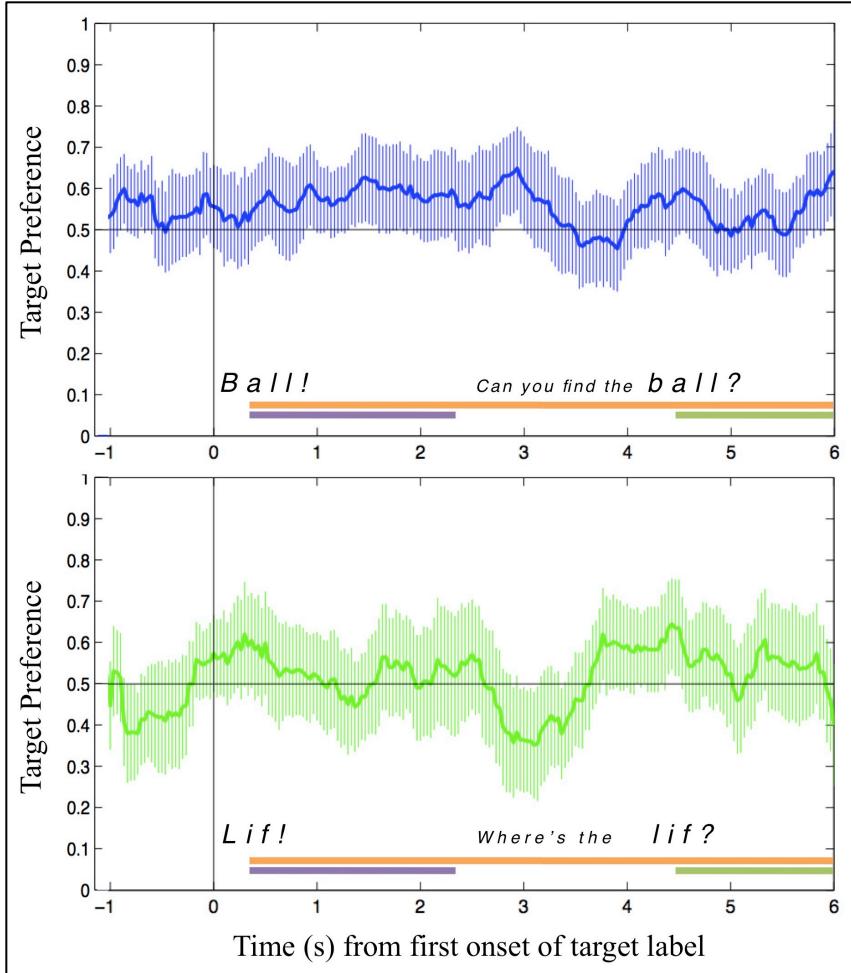


Figure 11. Time course of infants' preference for the target on *familiar* (upper) and *novel* (lower) test trials in Experiment 2. The curve represents mean target preference at each sample and vertical lines represent the 95% confidence interval for each sample. Color bars along the x-axis indicate timing of analysis windows (orange: *whole trial*; purple: *first presentation*; green: *second presentation*).

following the onset of the sentence frame, infants tended to shift away from the target and toward the distractor and actually showed a preference (as indicated by independent t tests) for the distractor on multiple samples. Infants' preference then shifts back toward the target, and is above chance for a few samples. We can not make any claims about the meaning of this pattern of looking as it did not meet our thresholds of significance, but it provides some indication that

infants were engaged in the task and somewhat systematically shifting their attention between novel targets and distractors at different points in the trial.

Finally, we conducted an ANOVA comparing infants' preference for *familiar* and *novel* targets to determine whether those preferences differed from each other early in test trials versus late in test trials. An ANOVA with Window (first presentation, second presentation) and Test Condition (familiar, novel) as within-subjects factors revealed a significant interaction between Window and Test Condition, $F(1, 26) = 5.86, p = .02, \eta_p^2 = .18$, indicating that early and late target preference measures differed as a function of test condition. Follow-up paired-sample t tests revealed that target preference in the *first presentation* window was significantly higher for *familiar* trials ($M = .59$) than *novel* trials ($M = .50$), $t(26) = 2.97, p = .006, d = .57$. In the *second presentation* window, target preference was slightly—but not significantly—higher for *novel* trials ($M = .56$) than *familiar* trials ($M = .54$), $t(26) = 0.58, p = .57, d = .11$. Thus, it appears that infants' preference for the target on *familiar* test trials was driven by a strong preference for the target in the window following the first presentation of the label. In addition, although infants' preference for the target on *novel* trials was not different from chance, it was as high as their preference for familiar targets by the end of the trial. Separate analyses of the infants who did and did not habituate showed essentially this same pattern for both groups.

As in Experiment 1, we examined infants' target preferences on *known* trials. Parent ratings on the vocabulary questionnaire were used to determine the subset of *familiar* test trials for which the label for the target object was known by the infant (i.e., rating of 3 or higher on the vocabulary questionnaire). Three infants had no *known* trials and were thus excluded from this analysis. Of the remaining 24 infants with at least one *known* trial, an average of 1.29 *familiar* trials were excluded for each infant because the target word was not reported to be known by the

infant, resulting in on average 2.46 *known* trials per infant included in this analysis. As we observed for the *familiar* trials as a whole, target preference on *known* trials was significantly above chance for the *whole trial*, $t(23) = 2.78, p = .01, d = .57$, and in the *first presentation* window, $t(23) = 3.01, p = .006, d = .61$. Unlike the analysis including all *familiar* trials for all infants, target preference for *known* trials in the *second presentation* window was also significant, $t(23) = 2.16, p = .04, d = .44$. Preference scores were even higher when only familiar targets with a parent rating of 4 or higher were included (Figure 12).

An examination of the time course of infants' looking on *known* trials provides additional information about infants' identification of known referents in response to a labeling event (Figure 13, upper). When parents report that their infants actually know the object label, infants' preference for the labeled object increases steadily in the period following the first presentation

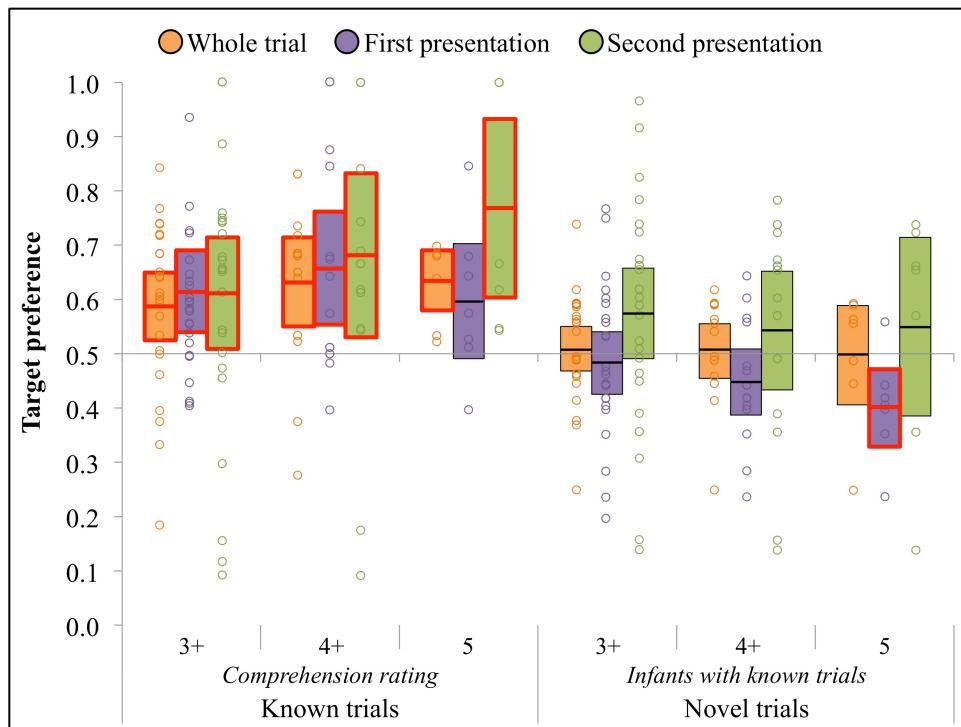


Figure 12. Experiment 2 target preference scores for *known* trials, divided by strength of parental rating of target words on the vocabulary questionnaire, and corresponding *novel* target preference scores from the same subset of participants (3+, $n = 24$; 4+, $n = 14$; 5, $n = 7$). Means different from chance (50%) are indicated by a bold red line.

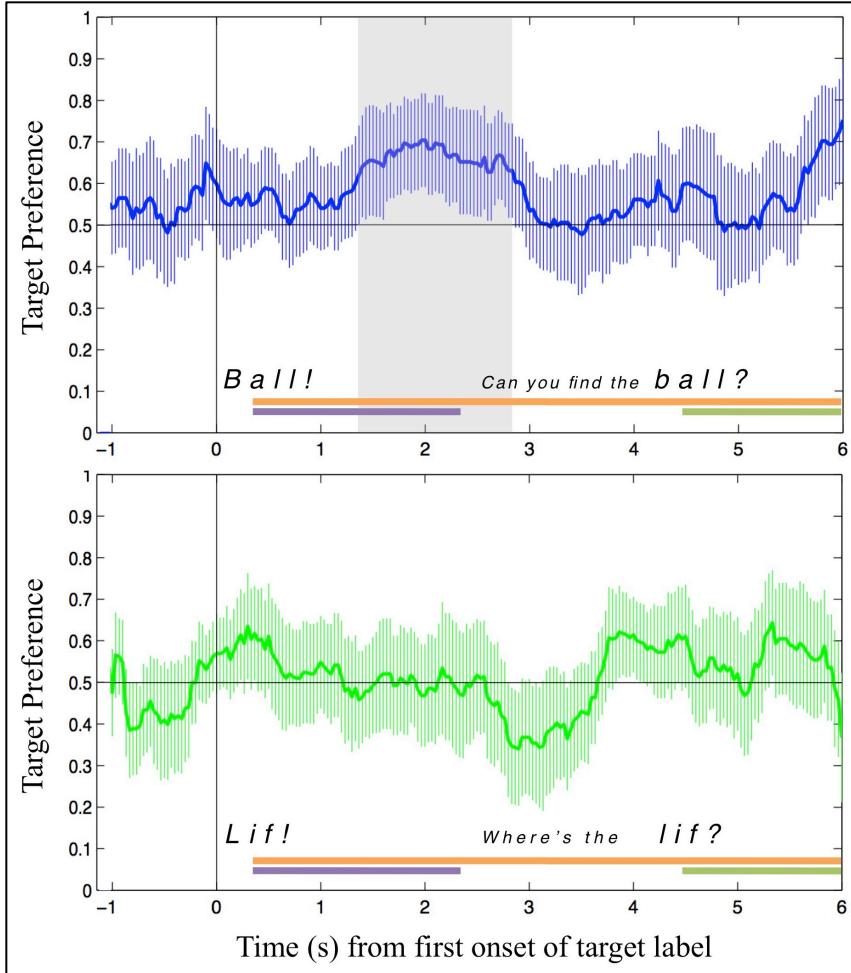


Figure 13. Time course of infants' preference for the target on *known* (upper) and *novel* (lower) test trials in Experiment 2. *Known* trials were the subset of *familiar* test trials for which individual infants were reported by a parent to understand the target word (i.e., a rating of 3 or higher on the vocabulary questionnaire). The curve represents mean target preference at each sample and vertical lines represent the 95% confidence interval for each sample. Color bars along the x-axis indicate timing of analysis windows (orange: *whole trial*; purple: *first presentation*; green: *second presentation*). The shaded area marks a run of samples that pass the 95% run length criterion.

of the label and remains well above chance for an extended period of time until the onset of the sentence frame. Infants' preference for the target also increases sharply at the end of the trial following the second presentation of the label, and it is not hard to imagine that a robust preference for the target would have been observed if the trial had continued. Thus, when

isolating just the *familiar* trials for which infants were reported by parents to know the target words, we observed an even more robust target preference. Moreover, compared to Experiment 1, in which infants demonstrated a *late* preference for the target on *known* trials, the target preference here emerges early and well within the analysis window identified by previous work. Comparison of the time course results for *known* trials in these two experiments suggests that exposure to variability during pretest and habituation facilitated infants' more rapid recognition of the targets of known words during the test trials.

An analysis of target preference scores for *novel* test trials using this same subset of participants who had at least one *known familiar* trial revealed a non-significant preference for the target in the *second presentation* window, $t(23) = 1.73, p = .097, d = .35$; infants' target preferences for the *whole trial*, $t(23) = 0.41, p = .68, d = .08$, and the *first presentation* window, $t(23) = 0.55, p = .59, d = .11$, were near chance (Figure 12). As observed for the group as a whole, this subset of infants showed higher preference for the target on *known* trials than *novel* trials in the *first presentation* window, $t(23) = 2.26, p = .03, d = .46$, while target preference scores in the *second presentation* window did not differ by condition, $t(23) = 0.55, p = .59, d = .11$. The time course of this subset of infants' target preference on *novel* trials looks nearly identical to that of the whole sample (Figure 13, lower).

As in the previous experiment, we were interested in whether individual differences in responding during test were related to other indices of development or attention such as age and vocabulary. We found no significant correlations between age or receptive vocabulary (as measured by the MSFVC-I) and target preference on *familiar* or *novel* test trials, $rs < .35, ps > .07$. Age and receptive vocabulary were also not significantly within this group, $r = .30, p = .13$.

Next, we examined the relation between infants' responding during *familiar* and *known* trials and *novel* trials. As in the previous experiment, we again found no correlation between *familiar* and *novel* responding, $rs < .07$, $ps > .77$, indicating that the variability observed on these trials is not explained by infants with stronger target preferences on one kind of trial also having stronger target preferences on the other kind of trial.

We next examined whether infants' target preference on *familiar* or *novel* test trials was influenced by their knowledge of the words heard on pretest trials. It is possible that pretest trials facilitated word recognition and/or learning more for infants who received pretest trials that better matched their vocabulary. To investigate a possible relation between knowledge of pretest labels and performance on test trials, we repeated the analyses for the subgroup of infants who were reported by a parent to know at least one of the labels heard on pretest trials ($n = 21$). Infants in this group showed above-chance target preference only during the *first presentation* window on *familiar* test trials, $t(20) = 2.92$, $p = .008$, $d = .64$ (*familiar* target preference for the *whole trial* was marginally significant, $t(20) = 2.02$, $p = .058$, $d = .44$). Target preference did not differ from chance for *novel* test trials in any analysis window, $ts < 1.48$, $ps > .15$. These results closely mirror those obtained from the full sample, and thus provide no evidence that hearing known labels on pretest trials affected infants' performance on test trials.

Finally, we asked whether infants' performance on the task could be predicted by their general interest in the task, as measured by the amount of time spent looking at the pretest stimuli. Duration of looking to the pretest stimuli could indicate general engagement in the task or a specific interest in object-labeling events, either of which could lead to enhanced performance on test trials. However, there was no relation between the duration of looking on pretest trials and target preference, $rs < .09$, $ps > .66$.

Discussion

Experiment 2 revealed a more robust recognition of the referents of familiar labels than did Experiment 1. This was especially true for labels that were reported by parents to be understood by their infants, demonstrating that increasing variability in acoustic features enhances the sensitivity of this task to 12-month-old infants' demonstration of their linguistic knowledge. One possible explanation of these results is that increased acoustic variability in the form of multiple label tokens produced by multiple speakers increased infants' attention to the stimuli and/or highlighted the referential nature of the task. As a result of this variability, infants might have been more prepared to engage in the visual-match-to-referent "game" of test trials than infants who did not hear variable label tokens during pretest and habituation trials in Experiment 1. Another possibility is that the increased variability in the label tokens heard during pretest and habituation trials increased infants' attention to the task, and they were therefore more engaged with the task when the test phase began. Although, as noted previously, a smaller (though not significantly so) percentage of participants met the habituation criteria in this experiment than in Experiment 1, the number of habituation trials completed was not different across experiments (Experiment 1: $M = 20.4$ trials; Experiment 2: $M = 21.8$ trials; $t(59) = 1.21, p = .23$). On average, infants in this experiment accumulated slightly more looking to habituation stimuli ($M = 191.7$ s) than did infants in Experiment 1 ($M = 178.0$ s), but this difference was not significant, $t(59) = 0.82, p = .42$. Therefore, these comparisons of indicators of infants' general level of interest in and attention to the task do not show differences consistent with the conclusion that observed effects were due to infants in Experiment 2 being more engaged in the task than were infants in Experiment 1.

However, as in Experiment 1, infants in Experiment 2 showed little evidence on *novel* test trials of having learned the associations between labels and objects presented during habituation. There are many possible explanations for the lack of a significant target preference on these trials. One possibility is that infants may have been unable to learn the novel associations over the course of the habituation phase. It is also possible, however, that infants may have learned something about the object-label associations but were unable to demonstrate their knowledge on novel test trials, perhaps because of limited attentional abilities.

The results also cannot be discussed without considering the higher rate of attrition observed in this study relative to Experiment 1. The parameters of the task (habituation criteria, maximum number of habituation trials, etc.) were chosen based on previous studies of word learning in infancy, as well as pilot testing using the stimuli from Experiment 1. It appears that something about the increased stimulus variability in the form of label tokens produced by multiple speakers made this a more difficult task for infants to complete. Perhaps the habituation criteria that were appropriate for Experiment 1 were not optimal for maximizing the number of participants who would meet the habituation criteria and therefore complete the study before getting fussy in this experiment, but comparisons between experiments required that the procedure remain constant. Exploring how study parameters might interact with the effect of stimulus variability on word learning is outside of the scope of this dissertation, but certainly worthy of further investigation.

Experiment 3

In the previous experiment, we introduced speaker variability and measured how this change in the learning context influenced infants' ability to match labels to their referents. With

the addition of between- and within-speaker variability, infants in Experiment 2 recognized the referents of familiar labels, demonstrating that increasing variability in acoustic features enhances the sensitivity of this task to infants' demonstration of linguistic knowledge. It is unknown from the results of Experiment 2 if the observed facilitation of familiar target recognition is an effect of variability in general or if it is specific to variability in the label tokens.

In Experiment 3, we asked if increased object variability will also affect infants' performance in this task. Experiment 3 was designed to assess the influence of variability in an object property that is unrelated to category membership, color, on infants' learning of object-label associations. If infants' improved performance observed in Experiment 2 was due to a facilitative effect of variability in general, we would expect variability in surface features of objects to facilitate performance in the same way as speaker variability. If it is not variability in general but variability in label tokens in particular that led to improved performance during Experiment 2, then we would expect infants in this experiment to behave like infants in Experiment 1. It is also possible that introducing object variability is an even more effective way to facilitate infant target recognition in this task, in which case we would expect to see stronger preferences for the target on both *familiar* and *novel* test trials.

Method

Participants. Participants were 24 12-month-old infants ($M = 379.7$ days, $SD = 7.5$ days, range = 369-394 days, 12 females). Fifteen infants were white, 2 were black/African American, and 7 were mixed race. Across these groups, 4 infants were Hispanic. All mothers had graduated from high school, and 14 had at least a Bachelor's degree. An additional 18 infants were tested but excluded from the final analysis for the following reasons: fussiness ($n = 10$), parental

interference ($n = 3$), or fewer than two test trials from each condition that met the inclusion criteria ($n = 5$).

Stimuli. Visual and auditory stimuli for test trials were identical to those used in Experiments 1 and 2.

The object exemplars used for test trials (and for pretest and habituation trials in Experiments 1 and 2) were removed from the set of available visual stimuli and not seen during pretest or habituation trials in this experiment. Five object exemplars (one from each of five color categories) were randomly selected for each participant from the remaining set of 10 color exemplars for each of three familiar object categories (ball, car, shoe). As previously mentioned, pretest stimuli were limited to these object categories in Experiments 2 and 3 due to practical limitations regarding creating multiple category exemplars varying in color for the other familiar categories (baby, dog, kitty). Four exemplars from each of the seven color categories (for a total of 28 color exemplars per novel object category) were randomly selected for each infant from the remaining set of 35 exemplars per object category for use on habituation trials. Colors were consistent across categories and therefore completely non-diagnostic of category membership.

Auditory stimuli for pretest and habituation trials were identical to those used in Experiment 1: a single token of each label, all produced by the same speaker.

Procedure and design. The procedure of this study is as described in the General Method section.

Objects for the two pretest trials were randomly selected for each infant. On each 10-s pretest trial, infants saw five exemplars of a familiar object category (ball, car, or shoe) from different color categories, and heard a single token of the object's label repeated five times.

On each 14-s habituation trial, infants saw seven exemplars from a novel object category (Category A or Category B) and heard a single token of its label (*lif* or *neem*) repeated seven times (Table 4). Color exemplars seen on each trial were randomly selected from the set of 28 exemplars per object category with the following constraints: 1) all seven exemplars seen on a given trial were from different color categories, and 2) object exemplars were not repeated until all exemplars had been used. Therefore, each infant saw 28 color exemplars from each novel object category once in each set of eight habituation trials (four trials per object category). This ensured that all infants were exposed to the same number of unique object exemplars regardless of the number of habituation trials they received (minimum of eight, maximum of 24).

Test trials were identical to Experiments 1 and 2.

Data analysis. Across all included participants, three *familiar* test trials and 14 *novel* test trials were excluded. On average, infants contributed 3.88 *familiar* test trials and 3.42 *novel* test trials to the analysis.

Results

Eleven infants met the habituation criteria, requiring an average of 18.9 trials to do so. The remaining 13 infants did not meet the habituation criteria and saw the maximum number of habituation trials (24). As in the previous experiment, infants who did not habituate did not show a significant decrease in looking from the first habituation block ($M = 41.9$ s) to the last habituation block ($M = 37.8$ s), $t(12) = 1.22$, $p = .25$, $d = .34$. Note that the attrition and failure-to-habituate rates are lower than Experiment 2, but still higher than Experiment 1 (although not significantly so according to chi-square tests, χ^2 s < 0.57, $ps > .45$). These are additional indicators that increasing the variability in the learning context makes it more difficult for infants to meet the habituation criteria and complete the task. In addition, and unlike in the previous

experiments, preliminary analyses indicated a marginally significant interaction between habituation status and target preference during the first and second presentation windows, $F(1, 22) = 3.43, p = .077$, suggesting that infants who met the habituation criteria performed differently on test trials than infants who did not. For consistency with the other experiments our main analyses are collapsed across all infants regardless of habituation status, but differences based on habituation status are explored in follow-up analyses.

As in the previous experiments, we first compared mean preference scores for both *familiar* and *novel* test trials to chance (50%) using one-sample t tests (Figure 14). Infants' preference for the target object did not differ from chance on either *familiar* or *novel* test trials, both across the *whole trial* and within the *first* and *second presentation* windows, $ts < 1.34, ps > .19$. Thus, as in Experiment 1—and unlike Experiment 2—we did not see a significant

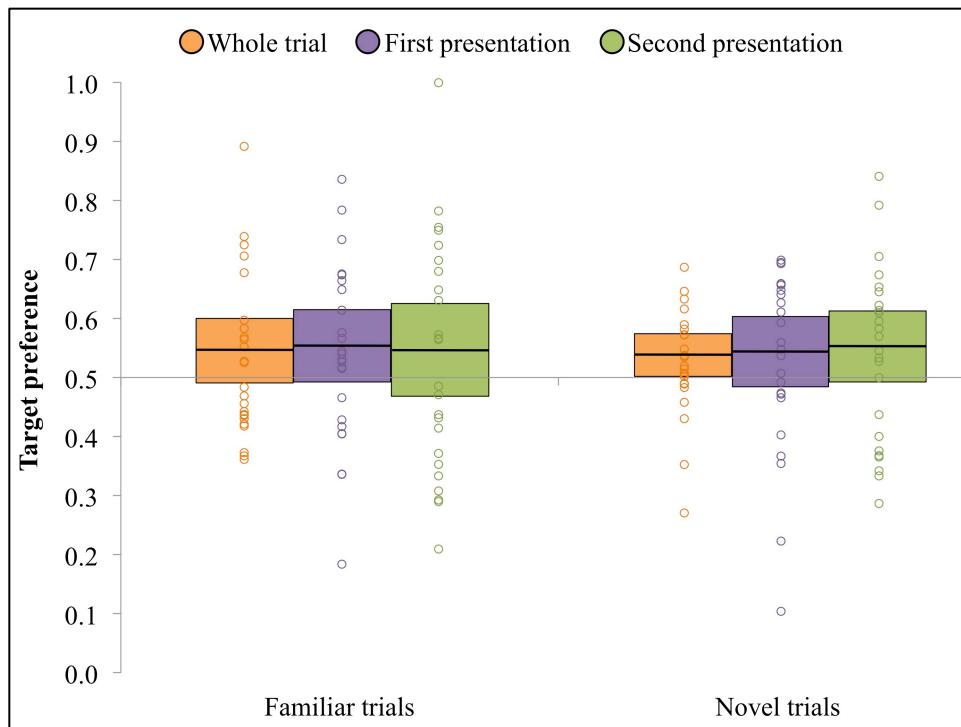


Figure 14. Experiment 3 *familiar* and *novel* test trial target preference. Target preference is defined as proportion of total looking to target and distractor objects spent looking to the target object. Each circle is the average target preference of an individual participant, boxes indicate 95% confidence intervals, and the horizontal line bisecting each box is the mean.

preference for familiar targets within the analysis windows in this task, suggesting that increasing variability in the objects presented during pretest and habituation did not increase the sensitivity of this task to infants' recognition of familiar object-label associations.

Because preliminary window analyses suggested that infants who met the habituation criteria might have performed differently at test than those who did not, we repeated the analyses for separately for the groups of habituators and non-habituators. Mean target preference scores for both *familiar* and *novel* test trials were again compared to chance (50%) using one-sample *t* tests. For infants who met the habituation criteria, target preference did not differ from chance for *familiar* test trials, $ts < 0.91$, $ps > .38$, but were greater than chance for *novel* test trials in both the *whole trial*, $t(10) = 2.50$, $p = .03$, $d = .75$, and *second presentation* analysis windows, $t(10) = 2.34$, $p = .04$, $d = .71$. Thus, infants who actually habituated to the novel object-label pairs appeared to correctly identify the target object on *novel* test trials. Interestingly, these infants did not respond more to the target on the *familiar* trials, suggesting that habituation with multiple object category exemplars allowed infants to respond to the label-object associations presented during habituation, but not to label-object associations they may have learned in their everyday lives. For infants who did not habituate, preference scores did not differ from chance for either type of trial, $ts < 1.45$, $ps > .17$.

Although initial analyses of infants' mean target preference within each analysis window indicate that the group of infants as a whole did not recognize targets on *familiar* test trials, a visual inspection of the time course of infants' looking provides some evidence that infants were able to identify familiar objects (Figure 15, upper left). There is a noticeable increase in infants' preference for the target in the period following the first label onset, with mean preferences of over 60% for multiple consecutive samples. Although this run does not meet the threshold of

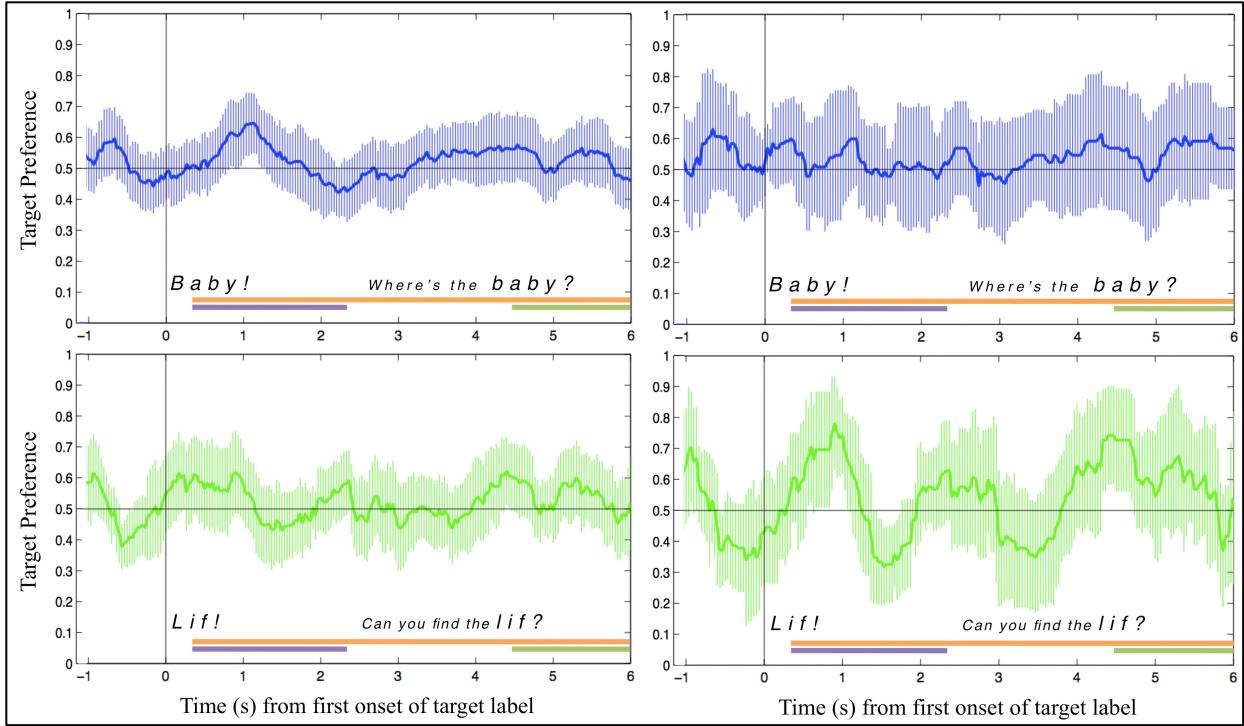


Figure 15. Time course of infants' preference for the target on *familiar* (upper) and *novel* (lower) test trials in Experiment 3. The panels on the left show target preferences for the whole sample, and the panels on the right show the target preferences for the subset of infants who met the habituation criteria. The curve represents mean target preference at each sample and vertical lines represent the 95% confidence interval for each sample. Color bars along x-axis indicate timing of analysis windows (orange: *whole trial*; purple: *first presentation*; green: *second presentation*).

significance determined by the permutation analysis, it is an indication that infants in this experiment responded differently on familiar test trials than did infants in Experiment 1, whose mean preference scores never reached 60% at any point in the trial. When the time course of looking is examined for the subset of infants who met the habituation criteria, it is clear that this group responded differently than the sample as a whole to *novel* test trials (Figure 15, lower right). In addition to having a general preference for the target over the whole trial and especially late in the trial, these infants are showing a systematic tendency to alternate between the target and distractor object (as shown by the more “wavelike” pattern of looking). Consistent with the

window analyses, this supports the idea that infants who habituated to the novel object-label pairs were able to demonstrate their learning on novel test trials in a way that the sample as a whole was not.

Finally, the ANOVA on target preference with Window (first presentation, second presentation) and Test Condition (familiar, novel) as within-subjects factors also revealed no significant main effects or interactions, $Fs < 0.07$, $ps > .79$, indicating that target preference did not differ based on analysis window or test condition.

We next examined infants' responding on *known* familiar trials to determine whether the kind of variability introduced here enhanced infants' responding on those trials. Parent ratings on the short vocabulary questionnaire were used to determine the subset of *familiar* test trials on which the label was known by the infant (i.e., a rating of 3 or higher). Three infants were excluded from this analysis because they had no *known* trials. Of the remaining 21 infants with at least one *known* trial, an average of 1.62 *familiar* trials were excluded for each infant because the target word was not reported to be known, resulting in on average 2.33 *known* trials per infant included in this analysis. For this subset of *known* trials, target preference for the *whole trial* was marginally significant, $t(20) = 2.03$, $p = .056$, $d = .44$ (Figure 16). The time course of infants' looking on these *known* trials also shows that target preference was above 50% for practically the entire trial following the onset of the label and was highest in the period following the first presentation of the label (Figure 17, upper left).

The relation between receptive vocabulary and target preference was even more pronounced when only trials with target words with a rating of 4 or higher were included ($n = 12$). For these trials, window analyses revealed that preference scores on *known* test trials exceeded chance over the *whole trial*, $t(11) = 6.42$, $p < .0001$, $d = 1.85$, and in the windows

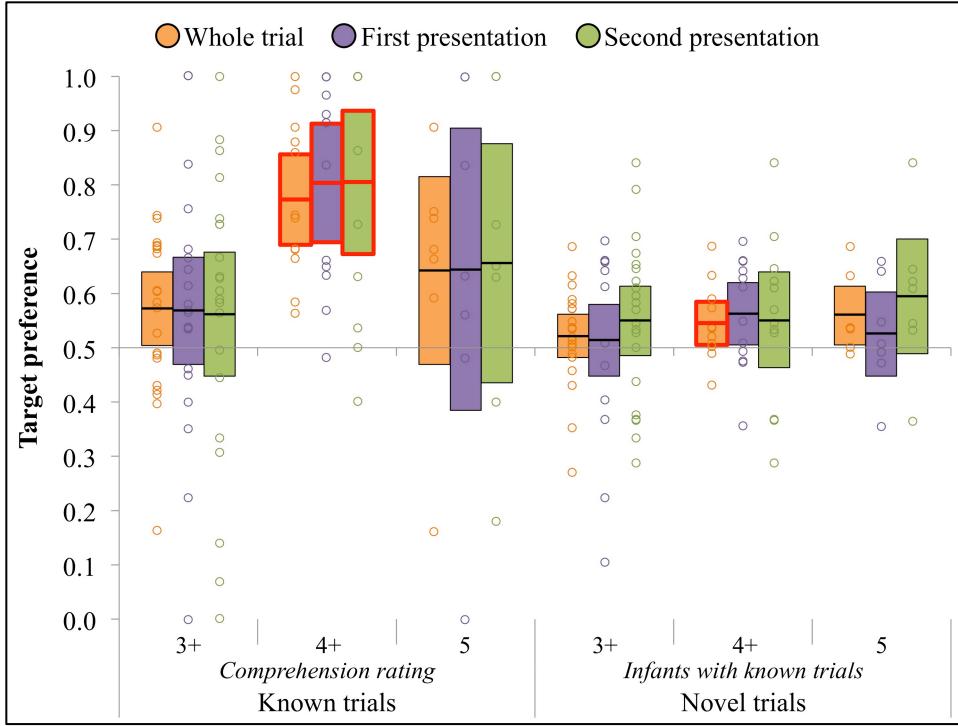


Figure 16. Experiment 3 target preference scores for *known* trials, divided by strength of parental rating of target words on the vocabulary questionnaire, and corresponding *novel* target preference scores from the same subset of participants (3+, $n = 21$; 4+, $n = 12$; 5, $n = 7$). Means different from chance (50%) are indicated by a bold red line ($p < .001$ for *known* 4+ trials, $p < .05$ for *novel* 4+ trials).

following both the *first presentation*, $t(11) = 5.49, p < .001, d = 1.58$, and the *second presentation* of the target word, $t(11) = 4.54, p < .001, d = 1.31$. In addition, the permutation analysis of the time course of infants' looking to these trials identified two runs of consecutive samples on which infants showed a significant sustained preference for the target object (Figure 17, upper right). However, the time course analysis also reveals that infants showed a significant preference for the target object in the period *preceding* the first labeling event. This indicates that these infants had an *a priori* preference for the target object on these trials, perhaps (or likely) because these were objects for which they knew a label. Similarly, this preference could be due to infants being more likely to know (or be reported by a parent to know) the labels for objects that draw their attention. Because targets and distractors were randomly paired, we do not have

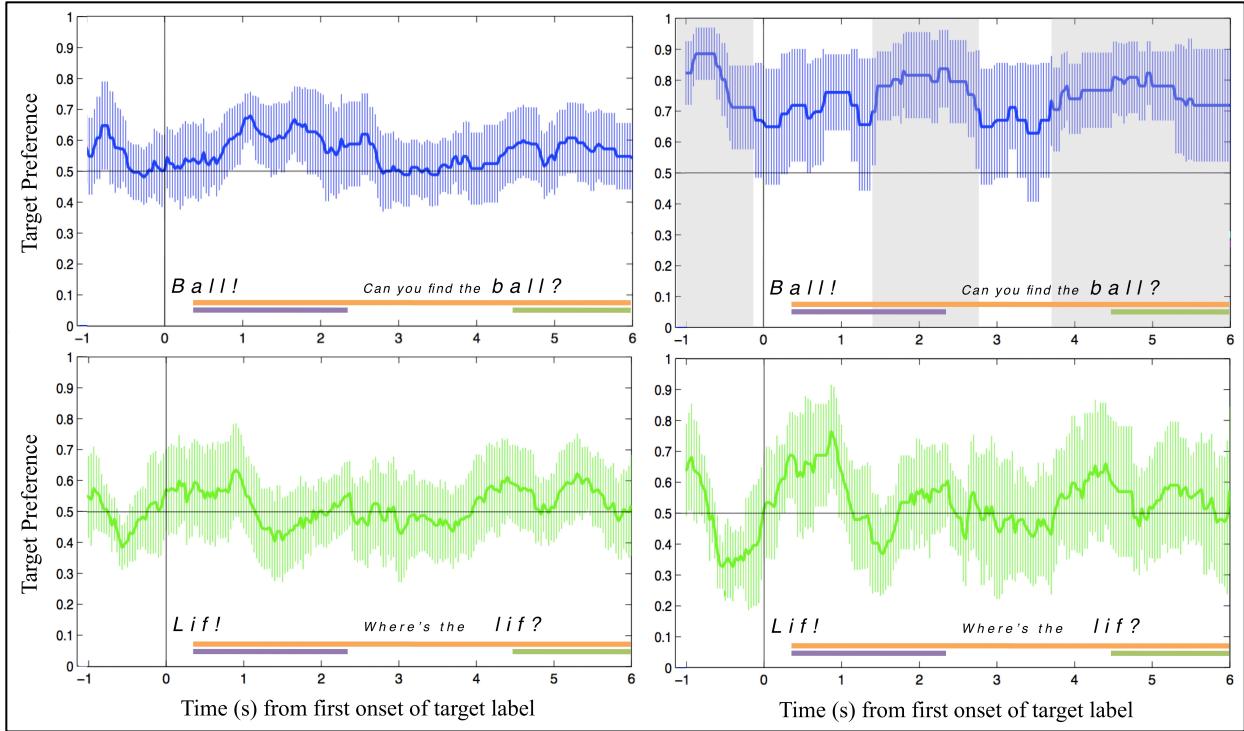


Figure 17. Time course of looking on *known* (upper) and corresponding *novel* (lower) test trials for the subsets of infants in Experiment 3 with at least one familiar target word rating of 3 or higher (left; $n = 21$) or 4 or higher (right; $n = 12$) according to parental report. The curve represents mean target preference at each sample and vertical lines represent the 95% confidence interval for each sample. Color bars along x-axis indicate timing of analysis windows (orange: *whole trial*; purple: *first presentation*; green: *second presentation*). The shaded areas mark runs of samples that pass the 95% run length criterion.

the power here to test for infants' preference for known targets compared to known or unknown distractors, but this is certainly a question that merits further investigation. Nonetheless, the permutation analysis indicates that infants' preference for the target was not different from chance at the onset of the target label, and exceeded chance in the periods following the labeling events.

To further investigate this relation between word knowledge and responding on test trials, we compared target preferences on *known* trials from this subset of infants to target preferences on *familiar* (but not *known*) trials from the remaining infants who had no familiar target words

with a rating of at least 4. A two-tailed independent samples t test confirmed that target preference on *familiar* test trials was higher for known target words, $t(22) = 4.19, p < .001, d = 1.71$. Therefore, as was observed in Experiment 2, this task is sensitive to infants' recognition of *known* object-label associations, especially for labels reported by a parent to be known with a high degree of confidence.

Unlike the previous two experiments, infants who knew at least one of the labels presented on *familiar* test trials also showed a preference for the target on *novel* test trials. The subgroup of infants with familiar target word ratings of 4 or higher successfully located the target object on *novel* test trials, as shown by above-chance target preference scores across the *whole trial* window, $t(11) = 2.31, p = .04, d = .67$ (target preference in the *first presentation* window for *novel* trials was also marginally significant, $t(11) = 2.11, p = .059, d = .61$) (Figure 16). A permutation analysis of the time course of this group's looking on *novel* test trials did not identify any runs of significant length, but a visual inspection of infants' preference over time shows infants' preference for novel targets was highest in the period immediately following the first presentation of the label (Figure 17, lower right). While infants with at least one known trial had an above-chance preference for *novel* targets and infants without any known trials did not, a direct comparison of these groups' *novel* preference scores across the whole trial revealed no significant difference, $t(22) = 1.14, p = .13, d = .47$.

We again followed up on these initial analyses by asking if individual differences in responding during test were related to general indices of development, age and vocabulary. We did find relations between these developmental measures and target preference scores. Age was positively correlated with target preference on *novel* test trials in the *second presentation* window, $r = .45, p = .03$, indicating that as infants got older they looked more to the labeled

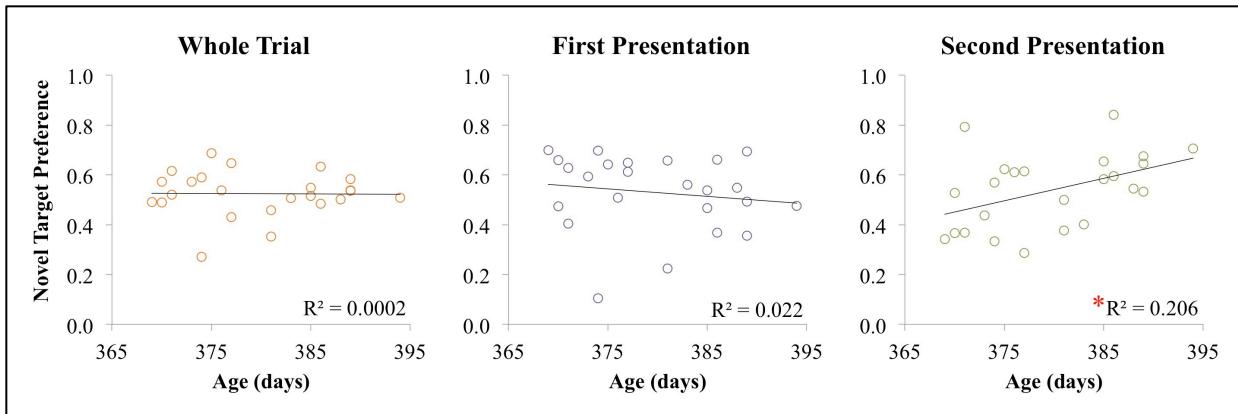


Figure 18. Scatterplots showing correlation between age and target preference on *novel* test trials in three analysis windows in Experiment 3. * $p < .05$.

novel object late in test trials (Figure 18). Age was not related to *familiar* target preference scores, $rs < .11$, $ps > .59$, or *novel* target preference scores in the *whole trial* or *first presentation* windows, $rs < .15$, $ps > .14$. Recall that in Experiment 1, which appears to have been less sensitive to infants' knowledge of familiar words overall, age was negatively correlated with target preference in the *second presentation* window of *familiar* test trials. Taken together, these results suggest possible interactions between age, sustained attention on test trials, and trial difficulty. One possibility is that younger infants in Experiment 1 were slower to direct their attention to the target on *familiar* trials, and thus showed a higher preference for the target during the second window than older infants who had previously identified the target and directed their attention elsewhere. Older infants' responding on the more difficult *novel* test trials in this experiment might have followed a pattern similar to that of younger infants' responding on *familiar* trials in Experiment 1, with more sustained looking to the target later in the trial after hearing the label for a second time. Of course, this is only one possible interpretation of the observed correlations.

We also observed a relation between our general measure of receptive vocabulary (MSFVC-I score) and target preference on *familiar* test trials in the *second presentation* window, $r = .40, p = .05$, meaning that infants with larger receptive vocabularies looked more to familiar targets after hearing them labeled for a second time than did infants who knew fewer words. As in the previous experiments, age and receptive vocabulary were not correlated, $r = .18, p = .40$, indicating that they provide relatively independent measures of developmental status. There was again no correlation between infants' preference for the target on *familiar* and *novel* trials, $r = .01, p = .96$, and also no significant relation between target preference on *known* trials (rating of 4 or 5) and *novel* trials, $r = -.26, p = .41$.

Lastly, we investigated a possible relation between infants' behavior during pretest trials and at test. We found significant correlations between duration of looking to pretest stimuli (which can be interpreted as a possible indicator of general interest or interest in object-labeling events) and target preference on *novel* test trials, both across the *whole trial*, $r = .57, p = .003$, and in the *first presentation* window, $r = .42, p < .001$. Therefore, it appears that infants who were more interested in viewing object-labeling events from the outset of the experiment were more likely to learn the novel object-label associations presented during the habituation phase than infants who were less interested in the pretest stimuli.

Discussion

Initial window and time course analyses of infants' preference for a labeled object relative to a distractor following exposure to object-labeling events containing multiple color exemplars from each object category indicated that infants in this experiment were unable to identify familiar or recently-learned referents. This suggests that the facilitative effect of stimulus variability on familiar label recognition observed in Experiment 2 is specific to

variability in features of the words, and not variability in general. However, supplemental analyses showed that infants in this experiment were able to demonstrate their linguistic knowledge. Infants showed an extremely robust preference for the target when they were reported by a parent to know the target label. In addition, infants who met the habituation criteria (which in itself is an indicator that some learning has taken place) successfully identified the target on novel test trials. Therefore, it appears that exposure variable object exemplars helped infants demonstrate their knowledge of known and newly-learned labels.

As discussed in the context of Experiment 2, it is still an open question as to exactly why variable object exemplars during pretest habituation led to improved performance on test trials. It is possible that pairing a label with object exemplars varying in a feature that is often not relevant to category membership (color) might highlight the unvarying relation between object category and label, and thus the referential nature of the task. Or, it might be that the variable object exemplars are more interesting and engaging for the infant to look at, and as a result they might get more exposure to the object-label pairs over the course of habituation. As was the case in Experiment 2, infants in this experiment accumulated non-significantly more looking to habituation stimuli on average ($M = 201.7$ s) than did infants in Experiment 1 in which the stimuli did not vary ($M = 178.0$ s), $t(56) = 1.53, p = .13$. Thus, the difference does not appear to be due to robust differences in overall levels of interest in the task.

Cross-Experiment Analyses and Summary

Taken individually, the results of Experiments 1, 2, and 3 indicate that increased variability in both labels and objects during pretest and habituation helped infants demonstrate knowledge of familiar labels during test. In addition, infants in Experiments 2 and 3 showed a

preference for the target on *novel* test trials in some circumstances, particularly when they heard known labels on familiar test trials. To further investigate the effect of increased stimulus variability on infants' performance on these tasks, we conducted additional analyses to compare sample characteristics and behaviors across studies.

First, we asked if infants' differential responding by experiment could be attributable to unintended a priori differences in age or vocabulary size between samples. As mentioned previously, attrition rates varied by experiment and were particularly high in Experiment 2. It is therefore possible that the infants included in the analyses of Experiments 2 and 3 were different in some way from the infants included in Experiment 1. However, one-way ANOVAs revealed no differences by study in age, $F(2, 82) = 1.67, p = .19$, or receptive vocabulary, $F(2, 82) = 0.17, p = .85$. In addition, there was no difference between studies in the average amount of time infants spent looking to the stimuli on pretest trials, $F(2, 82) = 0.38, p = .68$, a measure that can be interpreted as an index of infants' initial general interest in the task. Thus, it appears that the observed effects were not due to differences in our samples.

Next, we asked how infants as a whole responded to *familiar* and *novel* test trials, collapsed across studies. Overall, were infants able to identify referents of familiar labels and learn associations between novel labels and objects? Target preference scores on *familiar* and *novel* trials for all infants were compared to chance using two-tailed t tests. As a group, infants' preference scores were above chance for *familiar* trials in both the *whole trial*, $t(84) = 2.58, p = .01, d = .28$, and *first presentation* window, $t(84) = 2.61, p = .01, d = .28$, just as they were in Experiment 2. In addition, preference scores for *novel* trials were marginally significant in the second presentation window, $t(84) = 1.96, p = .054, d = .21$. When we compare infants' preference for the target on *known* trials (rating of 3 or higher) to chance, the effect is even

stronger. Infants significantly preferred *known* targets in all three analysis windows (*whole trial*: $t(70) = 3.24, p = .002, d = .38$, *first presentation*: $t(70) = 2.32, p = .02, d = .28$, *second presentation*: $t(70) = 2.48, p = .02, d = .29$), and these same infants' preference for *novel* targets exceeded chance in the *second presentation* window, $t(70) = 2.53, p = .01, d = .30$. It therefore appears that by collapsing across experiment, and increasing power, we show that overall the task is sensitive to infants' ability to recognize known words. In addition, we detected a small but reliable preference for novel targets that was not above chance for any of the three experiments individually, demonstrating that infants were able to learn novel associations over the course of habituation and show this learning during test, and raising the possibility that the failure to detect these effects in the individual experiments reflects a lack of power and not necessarily an absence of learning.

However, the question of interest (i.e., how variability contributes to infants' recognition of label-object correspondences) can best be addressed by comparing infants' performance in the different experiments reported here. In particular, we compared infants' performance on test trials following habituation to non-variable stimuli (Experiment 1) with performance on test trials following habituation to variable stimuli (Experiments 2 and 3). Recall that in the absence of variable stimuli, infants' preference for labeled objects did not differ from chance on *familiar* or *novel* test trials. Target preferences differed from chance for *familiar* trials in Experiment 2, but not Experiment 3. However, a visual inspection of the data indicates that the performance on *familiar* trials of infants in Experiment 3 was more similar to that of infants in Experiment 2 in which infants were also exposed to variable stimuli prior to test, than Experiment 1 in which habituation stimuli did not vary. We therefore conducted follow-up analyses in which we collapsed across variability conditions, and found that target preference on *familiar* test trials was

significantly above chance for the *whole trial*, $t(50) = 2.81, p < .01, d = .39$, and *first presentation* window, $t(50) = 3.41, p = .001, d = .48$, but not the *second presentation* window, $t(50) = 1.40, p = .17, d = .20$. Target preference on *novel* trials was not different from chance for the *whole trial* or *first presentation* window, $ts < 1.47, ps > .14$, but it was above chance for the *second presentation* window, $t(50) = 2.09, p = .04, d = .29$. An analysis of the time course of infants' looking collapsed across variability condition confirms the results of the window analyses (Figure 19). These infants' preference for the target on familiar test trials was at or above 50% for nearly the entire trial. In addition, infants showed a sustained preference for the

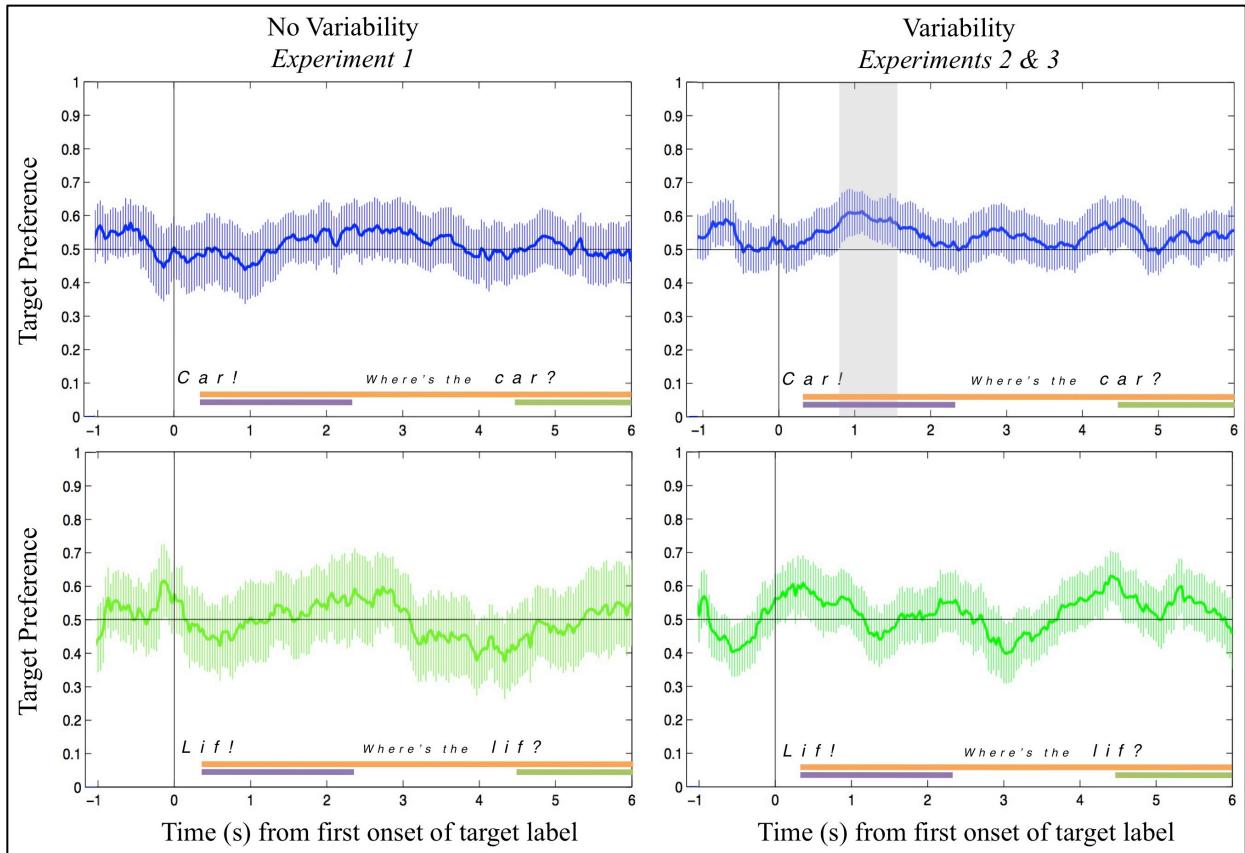


Figure 19. Time course of infants' target preference on *familiar* (upper) and *novel* (lower) test trials following habituation to non-variable (Experiment 1, left) or variable (Experiments 2 and 3, right) stimuli. The curve represents mean target preference at each sample and vertical lines represent the 95% confidence interval for each sample. Color bars along x-axis indicate timing of analysis windows (orange: *whole trial*; purple: *first presentation*; green: *second presentation*). The shaded area marks a run of samples that pass the 95% run length criterion.

target following the first presentation of the label that was statistically significant according to the 95% confidence threshold determined by random permutations of the data. We can therefore conclude that the task was sensitive to infants' recognition of familiar referents, but only when test trials followed exposure to variable object labeling events.

We then compared target preference scores on *familiar* and *novel* trials by experiment. Although preference scores were above chance for *familiar* trials in Experiment 2 but not for *familiar* trials in Experiments 1 or 3, an ANOVA with Experiment (1, 2, 3) as a between-subjects factor found that *familiar* target preference scores did not differ by experiment, $F(2, 82) = 1.30, p = .28$. Infants' average preference for the target on *novel* test trials also did not differ by experiment, $F(2, 82) = 0.98, p = .38$. While preference scores for *known* trials did not differ by experiment when *known* trials were defined as having a target word with a parental rating of 3 or higher on the vocabulary questionnaire, $F(2, 68) = 1.14, p = .33$, preference scores for *known* trials with ratings of 4 or higher (indicating a higher level of parental confidence) did differ by study, $F(2, 38) = 4.28, p = .02$. Follow-up *t* tests reveal that this effect is due to a higher average *known* target preference in Experiment 3 than Experiment 1, $t(25) = 2.75, p = .01, d = 1.06$, and Experiment 2, $t(24) = 2.35, p = .03, d = .92$. Known target preferences in Experiments 1 and 2 were not different, $t(27) = 0.71, p = .49, d = .26$.

Summary

The results of Experiments 1 through 3 reveal that exposure to variable stimuli influences 12-month-old infants' subsequent performance on tests of recognition for both familiar and recently-learned words. Only infants who habituated to labeling events in which objects varied in color (Experiment 3) demonstrated robust learning of these new words at test, suggesting that category-irrelevant object variation in particular supports word learning. Exposure to variable

label tokens (Experiment 2), on the other hand, appeared to boost infants' ability to demonstrate recognition of familiar words. When we analyzed the data collapsing across variability conditions, we observed a general improvement in performance for infants habituated to variable stimuli relative to those who were not. Possible causes and implications for these findings will be explored further in the General Discussion (Chapter 4).

Chapter 3: The Role of Variability in Visual Attention

Experiment 4

The primary goal of Experiment 4 was to investigate the relation between selective attention during visual search and label learning in 12-month-old infants. To achieve this goal, we developed a visual search task with two conditions, one with variability in a task-irrelevant feature (color), and one without. We tested infant participants of Experiment 1 with this task, allowing us to (1) examine the effect of task-irrelevant variability on selective attention, and (2) compare infants' performance in the label-learning task in Experiment 1 with their visual search in these two conditions.

Method

Participants. Forty-six of the 53 infants who participated in Experiment 1 also participated in this experiment (they were tested in the task in Experiment 1 first). The final sample consisted of 28 12-month-old infants ($M = 378.6$ days, $SD = 7.6$ days, range: 367-392; 14 females). Fourteen infants were white, 10 were mixed race, 2 were African American, and race was not reported for 2 infants. Across these groups, 3 infants were Hispanic. All mothers had graduated from high school, and 22 had at least a Bachelor's degree. Eighteen infants were excluded from the analyses for this experiment for the following reasons: fussiness or lack of interest ($n = 5$), could not calibrate/poor gaze tracking ($n = 6$), equipment error ($n = 1$), or fewer than two test trials from each condition that met the inclusion criteria ($n = 6$).

Twenty-three infants met the inclusion criteria for both Experiment 1 and Experiment 4 ($M_{age} = 378.1$ days, $SD = 7.5$ days, range: 367-392; 13 females). This sample will be used to explore a possible relation between selective visual attention as measured in this experiment and word learning as assessed in Experiment 1.

Stimuli. The primary stimuli were the *training* and *search* displays. Each type of display had an associated *reward* stimulus. Both types of displays involved a selection of colored shapes appearing on a black background. Training displays consisted of a single triangle (2.8h x 2.8w cm, subtending 2.7° x 2.7° visual angle at a distance of 60 cm) (Figure 20, top left panel). The triangle could be one of eight possible colors with the following CIE [x, y, luminance] coordinates: blue [0.14, 0.04, 26.0], cyan [0.23, 0.32, 101.0], green [0.31, 0.62, 106], orange [0.51, 0.45, 57.6], pink [0.29, 0.29, 60.9], purple [0.26, 0.11, 16.6], red [0.63, 0.34, 18.4], yellow [0.42 0.53 95.8]. The triangle could appear in one of eight possible locations (at 0°, or directly above the center of the screen, 45°, 90°, 135°, 180°, 225°, 270°, and 315° clockwise from 0°) equidistant from the center of the display (9 cm, or 8.5° visual angle, from center of screen). The reward display for each training display involved a slightly enlarged version of the triangle (4.5 x 4.5 cm, subtending 4.3° x 4.3° visual angle), presented in the same location as in the training display, with one of 12 possible cartoon characters overlaid (Figure 20, top right panel), accompanied by a short, engaging sound clip. Thus, when the reward was presented, the triangle appeared to grow somewhat and was enhanced by the inclusion of interesting auditory and visual features.

In the search displays, a single triangle was presented with seven circles. The target (triangle) appeared in one of the eight possible locations, and a distractor (circle) appeared in each of the other seven locations. In *same color* search displays, all eight items in the array were the same color (Figure 20, middle left panel). In *variable color* search displays, each item in the array was a unique color (Figure 20, bottom left panel). In the reward displays, the target triangle was slightly enlarged and overlaid with a cartoon character. In addition, the distractors remained

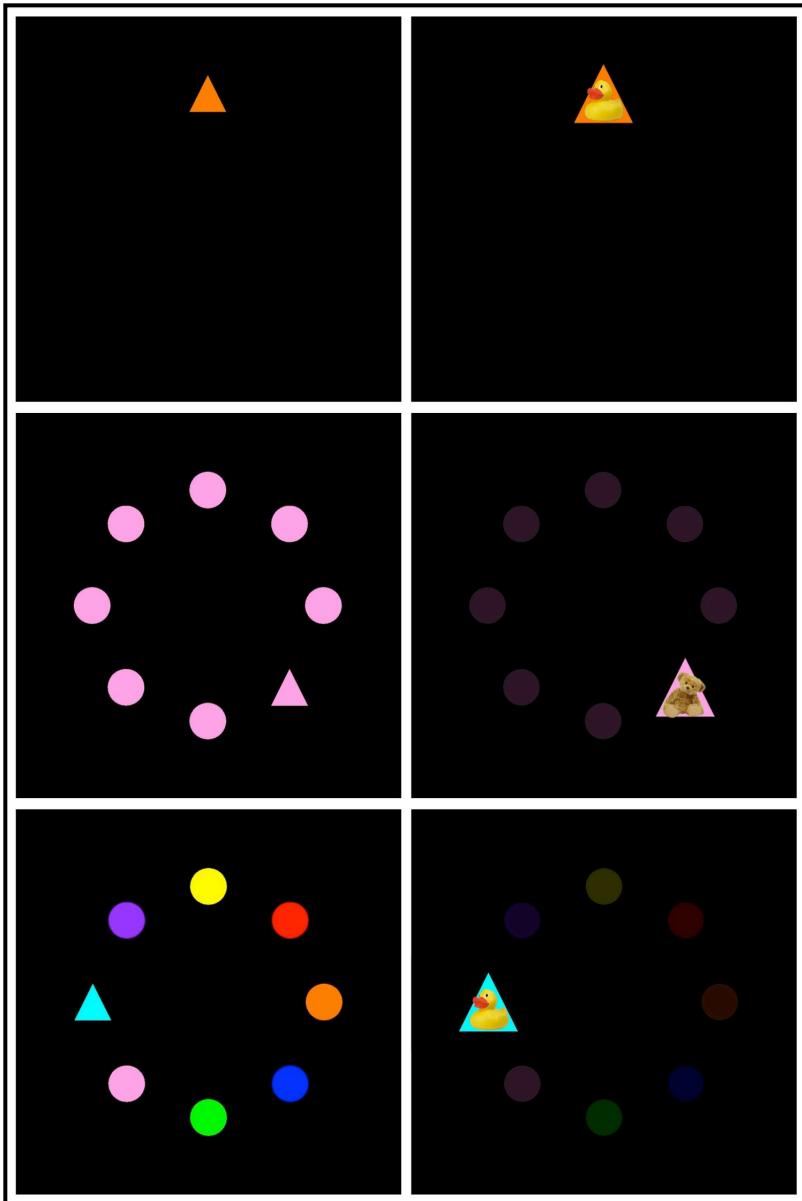


Figure 20. Example search (left) and reward (right) stimuli seen on training (top row), *same color* (middle row), and *variable color* (bottom row) trials in Experiment 4.

visible during the reward display but were dimmed relative to the search display (Figure 20, middle and bottom right panels).

We created two stimulus sets, each consisting of 8 *training* displays, 20 *same color* search displays, and 20 *variable color* search displays. Within each set, the target triangle's location and color were pseudo-randomly selected for each display, with the constraint that the

target appeared in each location and in each color an equal number of times across all displays. For *variable color* displays, the color of each circle distractor was also randomly selected from the remaining set of colors. The cartoon character that appeared on each reward display was also randomly assigned.

Short 15- to 30-s video clips from children's television shows were used as filler stimuli. Clips were "The Rubber Ducky Song" and "C is for Cookie" from *Sesame Street*, "The Mail Time Song" from *Blue's Clues*, and a cartoon hippopotamus singing "The Lion Sleeps Tonight."

Apparatus. Infants were tested in a dimly-lit experimental room (different from the room used for Experiment 1). Stimuli were presented on a 55.88-cm LCD monitor. Eye movement data was collected at a rate of 120 Hz using a SensoMotoric Instruments (SMI) RED eye tracker, controlled by a PC operated by a trained experimenter. The eye tracker used an eye camera attached to the bottom of the monitor to detect the infant's pupil and corneal reflections and use the difference between them to determine the location of the infant's point of gaze (POG). A wide-angle camera above the monitor provided a full view of the infant's behavior. The monitor and eye camera were mounted on a flexible arm that allowed the experimenter to adjust the position of the monitor as needed for individual infants. Stimuli were presented using SMI ExperimentCenter software.

Design and procedure. Infants sat on a parent's lap approximately 60 cm from the stimulus monitor. To minimize bias, parents wore felt-lined glasses and were asked not to talk to their infant or point to the screen during the experiment. A video of nature scenes accompanied by music was presented on the monitor to attract the infant's attention toward the screen. Once the eye tracker was focused on the infant's eyes, a short calibration procedure was initiated in which an attention-getting stimulus (a cartoon duck that jiggled and made a ringing sound) was

presented sequentially in five locations on the monitor. The eye tracker automatically collected information regarding the infant's estimated POG for each location. The accuracy of the calibration was validated by presenting a second stimulus (a rotating, multi-colored ball) sequentially in four locations and comparing POG with the stimulus location. If the calibration was deemed inaccurate (e.g., POG was more than approximately 2° from the validation stimulus), the calibration procedure was repeated. Once an accurate calibration was obtained, the experiment began.

Each trial was preceded by an approximately 4 cm² (3.2°) attention-getting stimulus (a checkered, striped, or rainbow-colored square or circle) presented in the center of the monitor as a series of attention-getting sounds (e.g., chirps, beeps, and whistles) played. Trials were infant-controlled; they were initiated when the eye tracker recorded that the infant looked to the central attention-getter for 200 ms (Figure 21, left panel). Once the trial was initiated, a training or search display (depending on the phase of the experiment) was presented on the monitor until the infant looked at the target item (i.e., the triangle) for a minimum of 100 ms, or until a maximum of 5 s had elapsed (Figure 21, right panel). The search display was then replaced by the reward display, which remained visible for a period of 3 s.

Each participant was randomly assigned to receive one of the two stimulus sets; note, however, that within the stimulus set the order of training and search trials was randomly ordered within blocks, so no two infants received the stimuli in precisely the same order. Infants were first presented with a series of eight *training* trials. On each trial, a training display was presented and then followed by a reward stimulus. The triangle appeared once in each of the eight possible locations and once in each of the eight possible colors over the course of this training phase. As there was only one item on the screen for the infant to look at during each training display, this

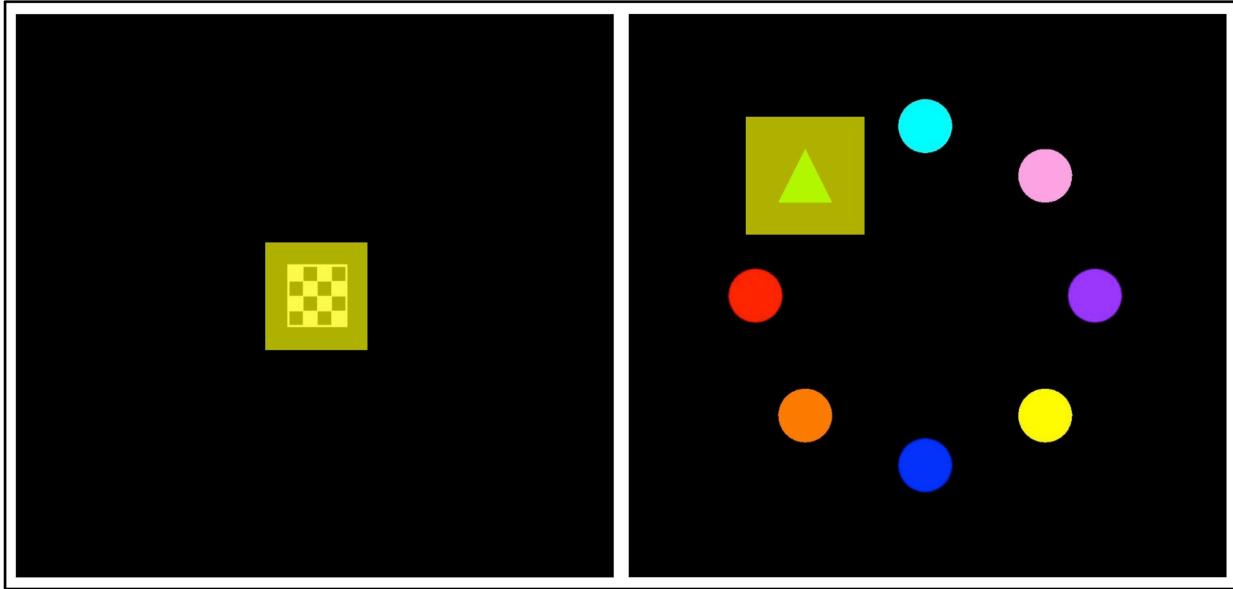


Figure 21. Example pre-trial attention-getting stimulus (left), with overlaid trigger area of interest (AOI). 200 ms of looking within this AOI initiated a trial. Example *variable color* search trial (right), with trigger AOI overlaid on target. 100 ms of looking within this AOI ended the search period and triggered the reward stimulus.

phase was designed to introduce infants to the task and provide positive feedback for looks to the triangle.

Following the training phase, infants were presented with up to 40 search trials. Again, each trial consisted of a search display followed by a reward display. Search trials were presented in blocks of eight, with four *same color* and four *variable color* trials per block. Within each block, the target appeared once in each possible location and once in each possible color, with a different random order for each subject. After each block of eight search trials, infants were shown a short video clip to break up the task and hold their interest.

Data analysis. Areas of interest (AOIs) were defined for the target and distractor items in each search array using SMI BeGaze analysis software (Figure 22). AOIs were drawn larger than the actual items (6.16 cm^2 , 5.9° visual angle) to allow for small inaccuracies in calibration and include looks to the edges of items. Recall that the eye tracker sampled infants' gaze 120 times

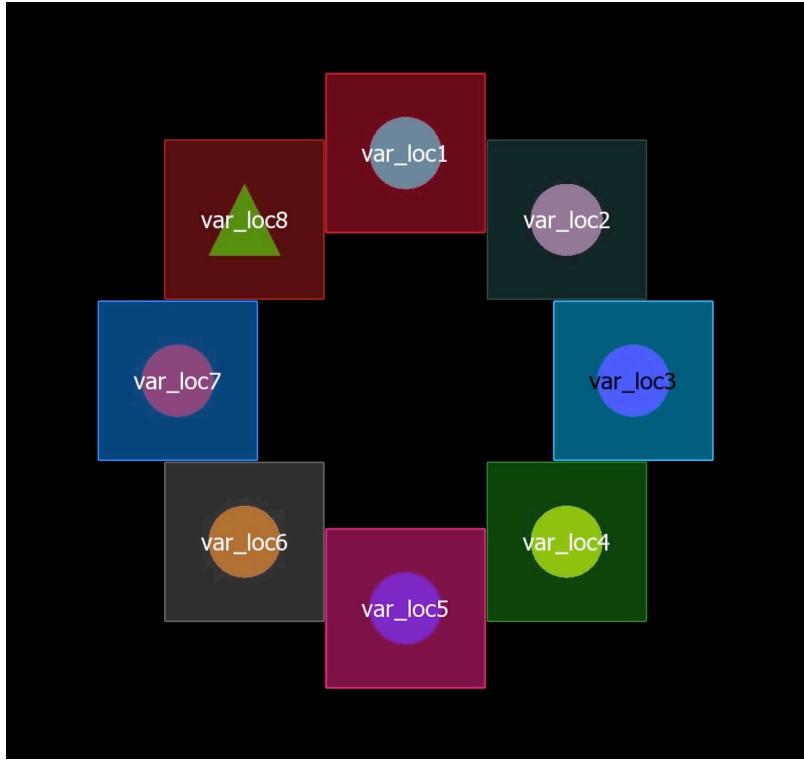


Figure 22. Analysis AOIs overlaid on an example *variable color* search display.

each second. To evaluate infants' looking to the targets and distractors here, we first filtered this moment-to-moment record of infants' gaze into individual fixations. Fixations were defined by the BeGaze software as 80 ms of continuous looking within a 100 pixel (2.7°) dispersion radius (a fixation length of 80 ms was chosen to ensure that a maximal number of reward-triggering looks to the target would meet fixation criteria). We then used the record of these fixations on each trial to determine the following measures of interest: accuracy (fixation to target during search display, yes or no); latency to target (start time of first target fixation relative to start of trial, for trials with a fixation to target only); target first fixation percentage (proportion of trials with a first fixation to target); and search efficiency (number of distractors fixated).

To maximize the likelihood that results reflect infants' ability to selectively direct their visual attention, we used the following exclusion criteria. Trials were excluded if: 1) cumulative

looking to all the search stimuli was less than 25% of the search display period, 2) there were no fixations to the reward target during the first 2000 ms of the reward display, and/or 3) although there were sufficient individual samples to the target to trigger the reward display before the 5-s maximum trial length, those samples did not meet the criteria for a fixation to the target (this occurred occasionally due to inconsistencies between how the stimulus presentation software and analysis software filtered the data into fixations). These criteria were chosen as indicators of either low interest in the task (1, 2) or inconsistent/inaccurate tracking of POG by the eye tracker (1, 2, 3). Five infants were excluded from the analysis because they did not have at least 2 trials per condition that met these criteria. Across all subjects in the final sample, 229 out of 785 search trials were excluded based on these criteria. On average, infants had 9.75 *same color* trials and 10.11 *variable color* trials that met the inclusion criteria. Because some infants quickly lost interest in the task—and their responding became variable and inconsistent after the first two blocks—we limited our analyses to the first two blocks of search trials (8 trials from each condition). After applying this cut-off, infants contributed on average 5.21 *same color* trials and 5.21 *variable color* trials to the analyses.

Results

Infants' responding in the visual search task. Our first set of analyses examined the effect of task-irrelevant variability on selective attention and included the full sample of 28 infants. To examine differences in infants' ability to locate the target triangle on *same color* versus *variable color* search trials, we compared each of our measures of interest using two-way, paired-samples *t* tests. These comparisons revealed, in general, that infants responded differently to the two types of trials (Figure 23). Specifically, a comparison of the latency to fixate the target revealed that infants located the target triangle significantly faster on *variable color* trials than

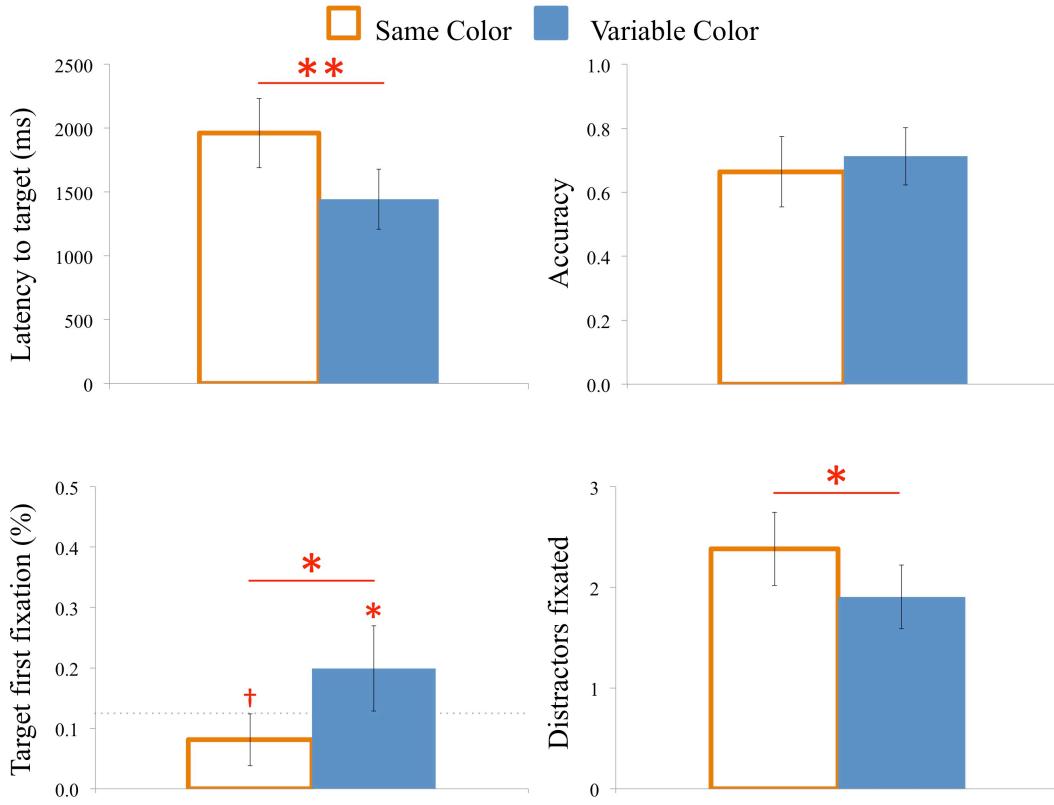


Figure 23. Comparison of search behavior on *same color* and *variable color* trials in Experiment 4 ($n = 28$). Error bars are 95% confidence intervals. $^{**}p < .01$, $^{*}p < .05$, $^{\dagger}p < .10$.

same color trials, $t(26) = 3.13, p = .004, d = .60^5$. They also had a higher proportion of first fixations to the target item on *variable color* trials, $t(27) = 2.63, p = .014, d = .50$. Finally, infants fixated fewer distractors on *variable color* trials than *same color* trials, $t(27) = 2.61, p = .014, d = .49$. Thus, infants apparently found it easier to locate the target when each item in the search array was a unique color than when all items were the same color. Accuracy (i.e., whether or not they ever fixated the target in a trial) did not differ by search condition, $t(27) = 0.92, p = .37, d = .17$, indicating that infants eventually located the target on a similar percentage of trials in each condition.

⁵ Note that this comparison excluded one infant who never located the target on a *same color* trial.

A second analysis compared infants' percentage of first fixations to the target on each search condition to chance (given that there were 8 items in the array, chance was 12.5%; Figure 23, lower left). Infants had more first fixations to the target item than expected by chance for *variable color* trials, $t(27) = 2.06, p < .05, d = .39$. On *same color* trials, however, infants' rate of first fixations to the target was marginally significant in the other direction, with fewer than chance first fixations to the target than, $t(27) = 2.01, p = .055, d = .38$, suggesting that their initial attention was drawn away from the target and toward the distractors on those trials. Thus, taken together, these results provide converging evidence that search performance differed by condition, with higher performance on *variable color* trials than *same color* trials.

Selective attention and word learning. The main goal of this experiment was to determine whether infants' learning of the object-label associations in Experiment 1 was related to general individual differences in selective attention. Our second set of analyses, therefore, investigated the relation between selective attention during visual search and label learning in 12-month-old infants. These analyses included the 23 infants who successfully completed both Experiment 1 and Experiment 4.

First, we confirmed that the results for the whole samples in each experiment were obtained with this subset of infants. Infants in this group contributed on average 5.17 *same color* trials and 5.22 *variable color* trials to the analysis. Like the full sample, these infants were faster to fixate the target, $t(21) = 2.79, p = .01, d = .60^6$, fixated fewer distractors, $t(22) = 2.24, p = .04, d = .47$, and fixated the target first on a (marginally significant) higher percentage of trials, $t(22) = 1.87, p = .07, d = .39$, in the *variable color* condition than in the *same color* condition. As before, overall accuracy did not differ by condition, $t(22) = 0.86, p = .40, d = .18$. Similarly, as

⁶ One infant was again excluded from this comparison because she did not locate the target on any *same color* trials.

observed for the whole sample in Experiment 1, these infants' target preference was not different from chance for either *familiar* or *novel* test trials in any of the three analysis windows, $ts < 1.34$, $ps > .20$. In addition, this sample featured high levels of variability in individual infants' average target preferences, with *familiar* trial target preferences ranging from 33-70%, and *novel* trial target preferences ranging from 26-67%.

Our next step was to conduct a series of bivariate correlations between measures of infants' attention in the *same color* and *variable color* conditions of this experiment with their target preferences on test trials in Experiment 1 (see Table 5). Infants' *novel* target preference in Experiment 1 was related to their performance on *same color* search trials. Target preference on *novel* test trials was significantly correlated with *same color* latency to the target and number of distractors fixated (Figure 24). In general, infants who were faster and more efficient at finding the target on *same color* trials preferred the referents of the novel words in the looking-while-listening task; infants who were slower and less efficient at finding the target on *same color* trials tended to divide their looking between the two objects on the looking-while-listening task. Measures of performance on *same color* trials, specifically latency to the target and number of distractors fixated, were also correlated with age. To ensure that the relations reported above were not due to this potentially mediating variable, we performed partial correlations controlling for age. The observed relation between *novel* target preference in Experiment 1 and latency to the target on *same color* search trials remained statistically significant, indicating that performance in the two tasks was related even beyond the contribution of age. Infants' performance on the easier *variable color* search trials was not consistently related to their target preference on *novel* word trials in Experiment 1.

Table 5. Correlations Between Target Preference (Experiment 1) and Visual Search Measures (Experiment 4)

| Search Condition (Expt. 4) | Search Measures | Target Preference (Expt. 1) | | | | | |
|-------------------------------|------------------------|-----------------------------|---------------------|-----------------------|---------------------|---------------------|---------------------|
| | | Novel trials | | Familiar trials | | partial correlation | partial correlation |
| | | bivariate correlation | partial correlation | bivariate correlation | partial correlation | | |
| Same Color | Latency to Target (ms) | -.539** | -.447* | .188 | .076 | .557** | |
| Accuracy | | .027 | .079 | -.113 | -.150 | -.136 | |
| Distractors Fixated | | -.457* | -.379 [†] | .288 | .225 | -.370 [†] | |
| Target First Fixation % | | .334 | .223 | -.291 | -.221 | .422* | |
| Variable Color | Latency to Target (ms) | -.162 | -.066 | .251 | .196 | -.302 | |
| Accuracy | | -.018 | .066 | .009 | .040 | -.229 | |
| Distractors Fixated | | -.268 | -.220 | .301 | .269 | -.193 | |
| Target First Fixation % | | .166 | .173 | -.290 | -.296 | .009 | |

Note. Values are Pearson Correlation coefficients. ** $p < .01$, * $p < .05$, [†] $p < .10$.

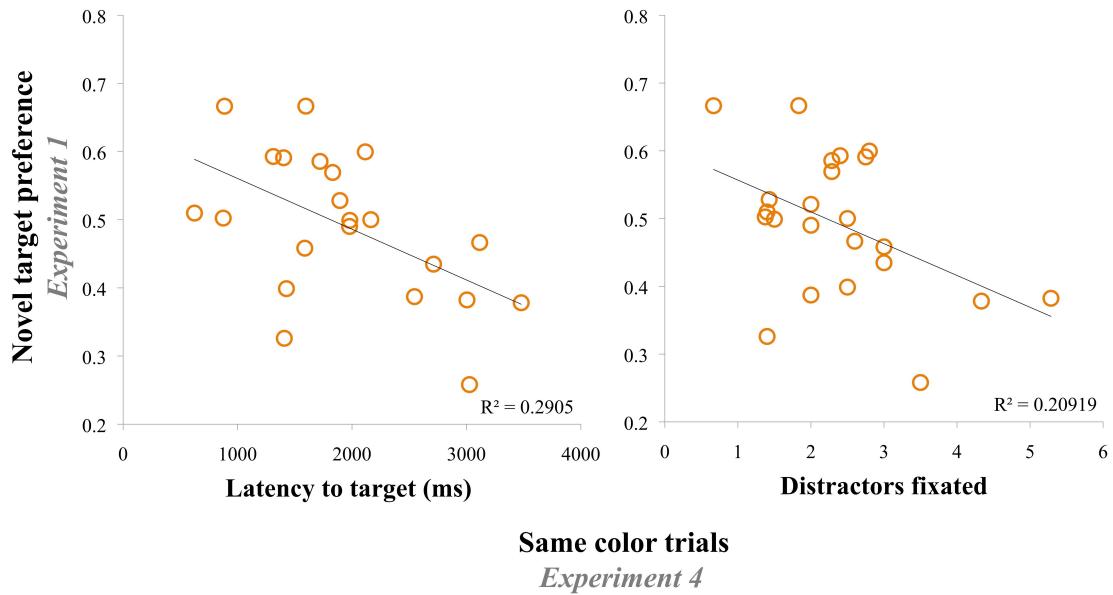


Figure 24. Scatter plot of correlations between search behavior on *same color* trials (x-axis) and *novel* target preference in Experiment 1 (y-axis).

Infants' preference for the target on *familiar* test trials in Experiment 1 was not significantly related to measures of visual attention on *same color* trials or *variable color* trials in Experiment 4 (Table 5). This therefore suggests that the observed relation between visual search behavior on the more difficult same color condition is related to infants' ability to learn new object-label associations, and not just to their ability to identify word referents in general.

Discussion

The results from the visual search task show that infants found it easier to locate the shape-defined target when distractor items varied in color than when all items were the same color. This pattern of results is somewhat counter-intuitive and seemingly at odds with results from the adult visual search literature, which has established that targets that differ from distractors by a single feature (in this case, shape) "pop-out" from search arrays and automatically draw attention in a bottom-up fashion (e.g., Treisman & Gelade, 1980). There is

some evidence that infants exhibit a pop-out effect under some conditions as well (Adler & Gallego, 2014; Adler & Orprecio, 2006). In our task, the target differs from distractors by a single feature (shape) in both conditions, and we would therefore not expect to see differences in performance by search condition if infants' visual attention here was being driven by the pop-out of a unique feature, as Adler and colleague's work would suggest. However, our comparison of the proportion of trials within each condition for which infants fixated the target item first revealed a significant (if not overwhelming) pop-out effect for *variable color* trials only. There are important differences between Adler's studies and the current experiment in methodology, participant age, and data analysis, however. Most notably, the authors of those studies compared the latency of infants' first saccade from a central fixation point to a target item across set sizes, and excluded all trials on which the first fixation was to a distractor item. They report that their measure of accuracy (analogous to our measure of proportions of first fixations to the target) did not differ by set size, but they do not compare this measure to chance. Therefore, it is difficult to know if the results of the present experiment are inconsistent with these previous studies of visual search in infancy or merely addressing a different question altogether.

It might also have been predicted from the visual search literature that if performance were to differ by condition, it would be the *same color* condition that would be easier for infants. Because color was constant across all items, the shape feature (and the difference in shape between the target and distractors) might have stood out more and made it easier for infants to locate the one item that differed on that feature. In fact, when we conceived of this task, we expected that this would be the case and that individual differences in search behavior on *variable color* trials would be more relevant to performance on the word-learning task in Experiment 1.

However, the results reveal the opposite pattern; infants found the target more quickly and efficiently when search items varied on a dimension that was irrelevant to the distinction between target and distractors, and it was performance on the more difficult *same color* trials that was related to target preferences in Experiment 1. As discussed in the contexts of Experiments 2 and 3, it also possible that *variability* in general induces greater levels of interest, and as a result infants were simply more attentive to the variable color displays. However, infants contributed a nearly identical number of trials from each condition to the analysis and accuracy did not differ by condition. In addition, infants actually looked at *fewer* distractors in the variable color condition, suggesting that they were not simply more interested in examining individual items that varied in color. Thus, we have no evidence that infants' general interest in the stimuli differed by search condition.

In fact, the present results considered along with the results of Experiments 1 through 3 are consistent with the conclusion that variability in task-irrelevant features can be beneficial for learning across domains. Color variability in the items in the search array might have heightened infants' attention *toward* (rather than distracting them *away from*) the feature that was consistent among distractors (shape) and therefore made it easier to locate the one item that differed along that dimension. Having all items vary along a task-irrelevant feature might have also helped infants to individuate the items in the array, whereas having all items be the same color might have led to a grouping effect in which infants treated the array as a single item (i.e., a large circle around the center of the monitor) rather than eight individual items. If infants found it difficult to individuate the items in the *same color* condition, they would be less likely to engage in goal-directed search behaviors.

Moreover, individual differences in selective attention in the more difficult *same color* condition were predictive of learning in Experiment 1, in which there was little variability in the stimuli. Specifically, we found that infants who performed better on *same color* trials demonstrated greater learning of the novel words presented without variation during habituation. Therefore, although infants did not demonstrate novel word learning on a group level in Experiment 1, we identified a measure of visual attention that explains at least some of the variance in individual responding.

This pattern of results is strikingly similar to those from Amsel and Johnson's (2006) previously discussed study of the relation between visual attention and object perception. Following exposure to a partially occluded moving rod, infants as a group showed no preference for an unoccluded completed or broken rod. However, individual differences in this task were related to measures of visual attention, as infants who perceived the unity of the rod parts in the occlusion task (as inferred from a dishabituation response to the broken rod) also showed superior performance on the more difficult search condition than did infants who looked longer at the complete rod. These results therefore add to a growing body of research showing that indices of visual attention can predict perception and learning in disparate contexts.

Chapter 4: General Discussion

The experiments presented here investigated the effects of stimulus variability on word learning and visual attention. We were specifically interested in the effect of *non-contrastive* variability, or variability in features that are not relevant for word meaning or category membership. We hypothesized that an increase in non-contrastive variability would facilitate infants' ability to learn novel object-label associations and to demonstrate that learning in a looking-while-listening task. We also explored whether infants' ability to learn these novel associations was related to visual attention as assessed in a visual search task.

In Experiments 1, 2, and 3, we examined infants' preference for a target object relative to a distractor in response to hearing the object labeled. Prior to test trials, infants were habituated to two labeling events; on each trial a novel object was paired with a novel label. In Experiment 1, the objects and labels presented during habituation did not vary; a single exemplar of each object was paired with a single token of each label throughout. In Experiments 2 and 3 respectively, we introduced variability into non-contrastive features of the labels (speaker voice and intonation) or objects (color). In general, across these three conditions, 12-month-old infants showed only weak learning of the novel object-label associations presented during habituation, with infants who met the habituation criteria in Experiment 3 being the only group to reliably prefer the target on tests of these new words. Infants in Experiments 1 and 2 did not demonstrate having learned the novel associations on a group level. Thus, these experiments failed to extend previous findings that variability in non-contrastive, indexical features of speech facilitates word learning in infancy (Galle et al., 2014; Rost & McMurray, 2010) to this younger age group. It is important to note those findings came from studies of slightly older (14-month-old) infants' ability to learn phonemically similar words. Thus, it is possible that the language abilities of 12-

month-old infants are still too fragile or underdeveloped to allow for successful target recognition on this task even with the aid of non-contrastive variability, or the task was not sufficiently sensitive to infants' knowledge. However, infants' successful recognition of familiar targets in Experiment 2 shows that infants were able to demonstrate their pre-existing linguistic knowledge in this context. It is also possible that non-contrastive variation is beneficial when the to-be-learned words are very similar to one another and require disambiguation, such as the words used by McMurray and colleagues, but this type of variability is not helpful (or even harmful) when the words are dissimilar, as were the words in the present investigation. This remains an open question that could be addressed in future work by replicating Experiment 2 and replacing the novel labels used here (*lif* and *neem*) with a minimal pair (e.g., *buk* and *puk*).

Although we did not see an effect of non-contrastive label variability on novel word learning, we did see an effect of variability in a non-contrastive object feature, color. Infants who habituated to labeling events featuring multiple color exemplars from each novel object category demonstrated learning of these associations at test. This finding adds to a small but growing body of work showing that 12-month-old infants are capable of learning words after a brief period of exposure and of demonstrating this learning in a laboratory setting (e.g., MacKenzie et al., 2012). In addition, infants' ability to demonstrate learning in Experiment 3 (and failure to demonstrate learning in Experiments 1 and 2) suggests that exposure to variable object tokens is particularly beneficial for forming associations between novel objects and words. Why might this be the case? There is evidence that labeling objects facilitates categorization for infants at this age (Booth & Waxman, 2002; Waxman & Markow, 1995). The results of Experiment 3 suggest that the converse is also true: providing a category of objects (rather than an individual object) facilitates word learning. In the same way that labeling a multiple objects might provide a cue

that those objects should be grouped together into a category, pairing multiple objects with a label might provide an additional cue that the association between the label and objects is important and worthy of attention. This interpretation fits nicely with the observation that infants are robust categorizers by this age, and many of infants' first words are simple nouns that refer to basic object categories (Tardif et al., 2008).

Our primary motivation for these studies was to investigate the effect of variable stimuli on infants' learning of *novel* associations, but the experiments revealed what may seem to be a somewhat surprising effect of variability during habituation on infants' responding on *familiar* trials; i.e., those trials that (presumably) tested knowledge of associations infants formed *before* coming to the lab. Infants showed stronger target preferences on trials with familiar labels and referents (e.g., ball, baby, shoe) following exposure (during habituation) to novel object labeling events that varied either in speaker (Experiment 2) or in the color of the objects (Experiment 3). This relation between stimulus variability and responding on *familiar* trials is intriguing. As previously discussed, few studies using traditional looking-while-listening or switch procedures have demonstrated consistent laboratory word recognition by infants younger than 14 months of age, even when parents report that infants know the words being tested. Twelve-month-old infants' failure on traditional comprehension tasks has been suggested to reflect fragile linguistic knowledge that does not support robust preferences for referents, narrow word representations that are mapped to specific referents (and thus do not lead to recognition of unfamiliar exemplars in laboratory experiments), and/or limits on attentional abilities such as the ability (or desire) to sit still or to inhibit attention to distractions (see Bergelson & Swingley, 2012; Fernald et al., 2008 for discussions). These explanations are supported by the fact that studies showing successful target recognition before 14 months have tended to do so by modifying procedures to

reduce task demands (e.g., by having a parent label items in a picture book for their infant; Bergelson & Swingley, 2012) or by testing infants' ability to match words to specific familiar referents (e.g., looking to a video of their parents in response to the word "mommy" or "daddy"; Tincoff & Jusczyk, 1999).

Given younger infants' ability to demonstrate word knowledge in modified contexts, it makes sense that infants' responding to *familiar* object-label pairings would be influenced by the broader context and seemingly unrelated aspects of the task. That is, it is typically assumed that trials like the familiar trials included here are "pure" assessments of infants' knowledge, and are included as control trials to evaluate the efficacy or sensitivity of the task. Clearly, the results presented here show that the context of the testing environment is critical in young infants' ability to demonstrate word knowledge even on such trials, and that success or failure on those control trials (at least in the present case) reflects something about what infants have learned about the task or the testing context *in general*. We have no reason to believe that the infants that constituted the sample of each experiment were different from each other prior to their participation—they were nearly identical in age, reported vocabulary, and interest in the task as indexed by duration of looking and habituation. And yet infants who were exposed to variable object labeling events performed differently on test trials with familiar words than did infants who were exposed to non-variable events. Despite the fact that the rate of habituation and accumulated looking to habituation stimuli did not differ by experiment, increased stimulus variability during those habituation trials may have prevented infants from losing interest in the task prior to test trials.

Alternatively, the differences in variability may have provided infants with different information about the *linguistic* nature of the task. At the very early stages of language learning,

demonstrating knowledge of word-object associations may require that infants are in the right “frame of mind”—that is, that they understand (in some sense) that these are *labeling* events. Indeed, researchers have adopted strategies to make the object-labeling aspect of tasks clearer to young infants, such as providing pretest trials with familiar referents or presenting words in the context of familiar sentence frames (Fennell & Waxman, 2010; Fennell, 2012; Fernald & Hurtado, 2006). These studies show that infants show greatest sensitivity to object-label associations when such strategies are adopted. The manipulations discussed above that allowed infants as young as 6 months to recognize familiar words (Bergelson & Swingley, 2012; Tincoff & Jusczyk, 1999), may demonstrate infants’ sensitivity to object-relations in these highly familiar object labeling contexts. In the present experiment, the variability in the novel object labeling events might have facilitated familiar target recognition by providing similar cues about the linguistic and referential nature of the task.

A second goal of this project was to explore a possible relation between infants’ ability to demonstrate their learning of novel words at test and selective attention. As a group, infants in Experiment 1 did not show a significant target preference on novel test trials, providing no evidence that they had learned the novel associations. However, we found that these infants’ ability to locate the target (and inhibit looks to distractor items) in the more difficult *same colors* trials was correlated with their ability to demonstrate their learning of novel labels. Infants who were better able to selectively direct their visual attention to a shape-defined target on the search task without the aid of color variation were also better able to direct their attention to the referent of a word learned in the absence of stimulus variability. It therefore appears that the ability to detect, attend to, and/or demonstrate learning of meaningful associations between invariable objects and words is closely linked to or perhaps driven by the same mechanism that supports

visual search across invariable distractors. The relation between visual search and word learning in the more difficult, variability-free condition of each task is consistent with other studies of attention and learning in infancy, in which search performance is associated with more advanced performance on other measures of cognition such as object perception (e.g., Amslo & Johnson, 2006). Advanced visual search performance is indicative of the ability to gather relevant information from visual scenes (and inhibit attention to irrelevant information), a skill that is vitally important for successful word learning.

Taken together, the experiments presented here consider the effects of stimulus variability on attention and learning, and suggest that increased variability in task-irrelevant features is beneficial across domains. In addition, the observed relation between measures of visual attention and word learning highlight the importance of considering how multiple aspects of an infant's cognitive system develop together, rather than examining abilities in isolation.

References

- Adler, S. A., & Gallego, P. (2014). Search asymmetry and eye movements in infants and adults. *Attention, Perception & Psychophysics*, 76(6), 1590–1608.
- Adler, S. A., & Orprecio, J. (2006). The eyes have it: Visual pop-out in infants and adults. *Developmental Science*, 9(2), 189–206.
- Amso, D., & Johnson, S. P. (2005). Selection and inhibition in infancy: Evidence from the spatial negative priming paradigm. *Cognition*, 95(2), B27–B36.
- Amso, D., & Johnson, S. P. (2006). Learning by selection: Visual search and object perception in young infants. *Developmental Psychology*, 42(6), 1236–1245.
- Apfelbaum, K. S., & McMurray, B. (2011). Using variability to guide dimensional weighting: Associative mechanisms in early word learning. *Cognitive Science*, 35(6), 1105–1138.
- Bahrick, L. E., Gogate, L. J., & Ruiz, I. (2002). Attention and memory for faces and actions in infancy: The salience of actions over faces in dynamic events. *Child Development*, 73(6), 1629–1643.
- Bahrick, L. E., Hernandez-Reif, M., & Flom, R. (2005). The development of infant learning about specific face-voice relations. *Developmental Psychology*, 41(3), 541–552.
- Barsalou, L., Sloman, S., & Chaigneau, S. (2005). The HIPE Theory of Function. In L. Carlson & E. van der Zee (Eds.), *Functional features in language and space: Insights from perception, categorization, and development* (pp. 131–147). New York: Oxford University Press.

- Baumgartner, H. A., & Oakes, L. M. (2011). Infants' developing sensitivity to object function: Attention to features and feature correlations. *Journal of Cognition and Development*, 12(3), 275–298.
- Behrmann, M., & Williams, P. (2007). Impairments in part-whole representations of objects in two cases of integrative visual agnosia. *Cognitive Neuropsychology*, 24(7), 701–730.
- Bergelson, E., & Swingley, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *PNAS*, 109(9), 3253–3258.
- Berkson, G., & Fitz-Gerald, F. L. (1963). Eye fixation aspect of attention to visual stimuli in infant chimpanzees. *Science*, 139(3555), 586–587.
- Bhatt, R. S., & Quinn, P. C. (2011). How does learning impact development in infancy? The case of perceptual organization. *Infancy*, 16(1), 2–38.
- Booth, A. E., & Waxman, S. R. (2002). Object names and object functions serve as cues to categories for infants. *Developmental Psychology*, 38(6), 948–956.
- Brennan, W. M., Ames, E. W., & Moore, R. W. (1966). Age differences in infants' attention to patterns of different complexities. *Science*, 151(3708), 354–356.
- Carey, S. (2009). *The origin of concepts*. New York, NY: Oxford University Press.
- Casasola, M. (2005). When less is more: How infants learn to form an abstract categorical representation of support. *Child Development*, 76(1), 279–290.
- Casey, B. J., & Richards, J. E. (1988). Sustained visual attention in young infants measured with an adapted version of the visual preference paradigm. *Child Development*, 59(6), 1514–1521.
- Cohen, L. B., & Oakes, L. M. (1993). How infants perceive a simple causal event. *Developmental Psychology*, 29(3), 421–433.

- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193–222.
- Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science*, 146(3644), 668–670.
- Fennell, C. T. (2012). Object familiarity enhances infants' use of phonetic detail in novel words. *Infancy*, 17(3), 339–353.
- Fennell, C. T., & Waxman, S. R. (2010). What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development*, 81(5), 1376–1383.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., & Thal, D. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59(5), 1–185.
- Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics*, 21(01), 95–115.
- Fernald, A., & Hurtado, N. (2006). Names in frames: infants interpret words in sentence frames faster than words in isolation. *Developmental Science*, 9(3), F33–F40.
- Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, 42(1), 98–116.
- Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernandez, & H. Clahsen (Eds.), *Developmental*

- psycholinguistics: On-line methods in children's language processing* (pp. 97–135). Amsterdam, The Netherlands: John Benjamins Publishing.
- Fukuda, K., & Vogel, E. K. (2009). Human variation in overriding attentional capture. *The Journal of Neuroscience*, 29(27), 8726–8733.
- Fukuda, K., & Vogel, E. K. (2011). Individual differences in recovery time from attentional capture. *Psychological Science*, 22(3), 361–368.
- Fukuda, K., Vogel, E. K., Mayr, U., & Awh, E. (2010). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review*, 17(5), 673–679.
- Galle, M. E., Apfelbaum, K. S., & McMurray, B. (2014). The Role of Single Talker Acoustic Variation in Early Word Learning. *Language Learning and Development*. doi:10.1080/15475441.2014.895249
- Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13(5), 431–436.
- Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, 18(3), 254–260.
- Hernandez-Reif, M., & Bahrick, L. E. (2001). The development of visual-tactual perception of objects: Amodal relations provide the basis for learning arbitrary relations. *Infancy*, 2(1), 51–72.
- Hood, B. M., & Atkinson, J. (1993). Disengaging visual attention in the infant and adult. *Infant Behavior and Development*, 16(4), 405–422.

- Jankowski, J. J., Rose, S. A., & Feldman, J. F. (2001). Modifying the distribution of attention in infants. *Child Development*, 72(2), 339–351.
- Johnson, M. H. (1990). Cortical maturation and the development of visual attention in early infancy. *Journal of Cognitive Neuroscience*, 2(2), 81–95.
- Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1991). Components of visual orienting in early infancy: Contingency learning, anticipatory looking, and disengaging. *Journal of Cognitive Neuroscience*, 3(4), 335–344.
- Johnson, S. P., Slemmer, J. A., & Amso, D. (2004). Where infants look determines how they see: Eye movements and object perception performance in 3-month-olds. *Infancy*, 6(2), 185–201.
- Kannass, K. N., Oakes, L. M., & Shaddy, D. J. (2006). A longitudinal investigation of the development of attention and distractibility. *Journal of Cognition and Development*, 7(3), 381–409.
- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS One*, 7(5), e36399. doi:10.1371/journal.pone.0036399
- Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, 83(2), B35–B42.
- Kloos, H., & Sloutsky, V. M. (2008). What's behind different kinds of kinds: Effects of statistical density on learning and representation of categories. *Journal of Experimental Psychology: General*, 137(1), 52–72.
- Lewkowicz, D. J., & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. *Trends in Cognitive Sciences*, 13(11), 470–478.

- MacKenzie, H., Curtin, S., & Graham, S. A. (2012). 12-month-olds' phonotactic knowledge guides their word-object mappings. *Child Development*, 83(4), 1129–1136.
- Madole, K. L., & Cohen, L. B. (1995). The role of object parts in infants' attention to form-function correlations. *Developmental Psychology*, 31(4), 637–648.
- Markant, J., & Amso, D. (2013). Selective memories: Infants' encoding is enhanced in selection via suppression. *Developmental Science*, 16(6), 926–940.
- Mather, E., & Plunkett, K. (2011). Same items, different order: Effects of temporal variability on infant categorization. *Cognition*, 119(3), 438–447.
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82(3), B101–B111.
- Namy, L. L., & Waxman, S. R. (1998). Words and gestures: Infants' interpretations of different forms of symbolic reference. *Child Development*, 69(2), 295–308.
- Oakes, L. M. (2010). Using habituation of looking time to assess mental processes in infancy. *Journal of Cognition and Development*, 11(3), 255–268.
- Oakes, L. M., Baumgartner, H. A., Barrett, F. S., Messenger, I. M., & Luck, S. J. (2013). Developmental changes in visual short-term memory in infancy: evidence from eye-tracking. *Frontiers in Psychology*, 4(697), 1–13. doi:10.3389/fpsyg.2013.00697
- Oakes, L. M., Coppage, D. J., & Dingel, A. (1997). By land or by sea: The role of perceptual similarity in infants' categorization of animals. *Developmental Psychology*, 33(3), 396–407.
- Oakes, L. M., & Madole, K. L. (2008). Function revisited: How infants construe functional features in their representation of objects. *Advances in Child Development and Behavior*, 36, 135–185.

- Oakes, L. M., & Ribar, R. J. (2005). A comparison of infants' categorization in paired and successive presentation familiarization tasks. *Infancy*, 7(1), 85–98.
- Oakes, L. M., & Tellinghuisen, D. J. (1994). Examining in infancy: Does it reflect active processing? *Developmental Psychology*, 30(5), 748–756.
- Palmer, S. B., Fais, L., Golinkoff, R. M., & Werker, J. F. (2012). Perceptual narrowing of linguistic sign occurs in the 1st year of life. *Child Development*, 83(2), 543–553.
- Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science*, 296(5571), 1321–1323.
- Pegg, J. E., & Werker, J. F. (1997). Adult and infant perception of two English phones. *The Journal of the Acoustical Society of America*, 102(6), 3742–3753.
- Perone, S., & Oakes, L. M. (2006). It clicks when it is rolled and it squeaks when it is squeezed: What 10-month-old infants learn about object function. *Child Development*, 77(6), 1608–1622.
- Plunkett, K. (2011). The role of auditory stimuli in infant categorization. In L. M. Oakes, C. H. Cashon, M. Casasola, & D. H. Rakison (Eds.), *Infant perception and cognition: Recent advances, emerging theories, and future directions* (pp. 203–221). New York, NY: Oxford University Press.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32(1), 3–25.
- Quinn, P. C., Adams, A., Kennedy, E., Shettler, L., & Wasnik, A. (2003). Development of an abstract category representation for the spatial relation between in 6- to 10-month-old infants. *Developmental Psychology*, 39(1), 151–163.

- Quinn, P. C., Cummins, M., Kase, J., Martin, E., & Weissman, S. (1996). Development of categorical representations for above and below spatial relations in 3- to 7-month-old infants. *Developmental Psychology, 32*(5), 942–950.
- Quinn, P. C., Eimas, P. D., & Rosenkrantz, S. L. (1993). Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. *Perception, 22*(4), 463–475.
- Rakison, D. H. (2004). Infants' sensitivity to correlations between static and dynamic features in a category context. *Journal of Experimental Child Psychology, 89*(1), 1–30.
- Rakison, D. H. (2005). A secret agent? How infants learn about the identity of objects in a causal scene. *Journal of Experimental Child Psychology, 91*(4), 271–296.
- Rakison, D. H., & Poulin-Dubois, D. (2002). You go this way and I'll go that way: Developmental changes in infants' detection of correlations among static and dynamic features in motion events. *Child Development, 73*(3), 682–699.
- Reznick, J. S., & Kagan, J. (1983). Category detection in infancy. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 2., pp. 79–111). Norwood, NJ: Ablex.
- Richards, J. E. (1989). Development and stability in visual sustained attention in 14, 20, and 26 week old infants. *Psychophysiology, 26*(4), 422–430.
- Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382–439.
- Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science, 12*(2), 339–349.

- Rost, G. C., & McMurray, B. (2010). Finding the signal by adding noise: The role of noncontrastive phonetic variability in early word learning. *Infancy*, 15(6), 608–635.
- Ruff, H. A., & Rothbart, M. K. (1996). *Attention in early development*. New York, NY: Oxford University Press.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70(1), 27–52.
- Scott, L. S., Pascalis, O., & Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. *Current Directions in Psychological Science*, 16(4), 197–201.
- Singh, L. (2008). Influences of high and low variability on infant word recognition. *Cognition*, 106(2), 833–870.
- Smith, L. B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3), 1558–1568.
- Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, 388(6640), 381–382.
- Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, 13(5), 480–484.
- Tardif, T., Fletcher, P., Liang, W., Zhang, Z., Kaciroti, N., & Marchman, V. A. (2008). Baby's first 10 words. *Developmental Psychology*, 44(4), 929–938.
- Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, 10(2), 172–175.

- Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *The Quarterly Journal of Experimental Psychology*, 37(4), 571–590.
- Tipper, S. P., Bourque, T. A., Anderson, S. H., & Brehaut, J. C. (1989). Mechanisms of attention: A developmental study. *Journal of Experimental Child Psychology*, 48(3), 353–378.
- Toscano, J. C., & McMurray, B. (2010). Cue integration with categories: Weighting acoustic cues in speech using unsupervised learning and distributional statistics. *Cognitive Science*, 34(3), 434–464.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12(1), 97–136.
- Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology*, 29(3), 257–302.
- Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology*, 34(6), 1289–1309.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7(1), 49–63.
- Wolfe, J. M. (2007). Guided search 4.0: Current progress with a model of visual search. In W. D. Gray (Ed.), *Integrated models of cognitive systems* (pp. 99–119). New York, NY: Oxford University Press.
- Woodward, A. L., & Hoyne, K. L. (1999). Infants' learning about words and sounds in relation to objects. *Child Development*, 70(1), 65–77.

- Yoshida, K. A., Fennell, C. T., Swingley, D., & Werker, J. F. (2009). Fourteen-month-old infants learn similar-sounding words. *Developmental Science*, 12(3), 412–418.
- Younger, B. A. (1985). The segregation of items into categories by ten-month-old infants. *Child Development*, 56(6), 1574–1583.
- Younger, B. A., & Cohen, L. B. (1983). Infant perception of correlations among attributes. *Child Development*, 54(4), 858–867.
- Younger, B. A., & Cohen, L. B. (1986). Developmental change in infants' perception of correlations among attributes. *Child Development*, 57(3), 803–815.
- Younger, B. A., & Furrer, S. (2003). A Comparison of Visual Familiarization and Object-Examining Measures of Categorization in 9-Month-Old Infants. *Infancy*, 4(3), 327–348.
- Zangl, R., Klarman, L., Thal, D., Fernald, A., & Bates, E. (2005). Dynamics of word comprehension in infancy: Developments in timing, accuracy, and resistance to acoustic degradation. *Journal of Cognition and Development*, 6(2), 179–208.

Appendix

Vocabulary Questionnaire

Children understand many more words than they say. We are interested in which words your child understands, in addition to words that your child says.

Please choose the option that best describes your child's ability to understand and say each word, and place a check mark in the appropriate box. If your child uses a different pronunciation of a word, mark the word anyway.

| | | Does not understand, does not say | Maybe understands, does not say | Understands, does not say | Understands, maybe says | Understands and says |
|---|---------------|-----------------------------------|---------------------------------|---------------------------|--------------------------|--------------------------|
| 1 | ball | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 2 | baby | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 3 | book | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 4 | bottle | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5 | car | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 6 | cookie | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 7 | dog | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 8 | kitty | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 9 | shoe | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Does your child use or understand a different word for any of the words listed above (e.g., "cat" instead of "kitty")? If so, please list them here:
