

**The Role of Contingency and Ostensive Cues on Infants' Cognitively Demanding Word-Object Learning**

Laura Mills-Smith

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy in Psychology

Committee Members:

Robin K. Panneton, Ph.D., Committee Chair  
Anthony D. Cate, Ph.D., Committee Member  
Craig T. Ramey, Ph.D., Committee Member  
Susan W. White, Ph.D., Committee Member

April 27, 2016

Blacksburg, VA

Keywords: referential learning, infancy, eye-tracking, ostensive cues

# **The Role of Contingency and Ostensive Cues on Infants' Cognitively Demanding Word-Object Learning**

Laura Mills-Smith

## **ABSTRACT**

Older infants are good referential learners. That is, at around 14-months of age, they begin to learn the verbal labels of objects and events around them. However, referential learning can be made more challenging by increasing the lexical similarity between labels. The primary goal of this study was to examine whether an adult speaker's ostensive cues and eye gaze-object contingency could augment referential learning in 14-month-old infants under difficult conditions (i.e., minimal pair labels). In Experiment 1, infants were familiarized and tested on two word-object associations with minimal pairs (e.g., "bin" and "din"), presented on an eye-tracker. Importantly, each session began when infants made eye contact with a female speaker on the screen, and she continually looked at and verbally referenced each object in an infant-directed style. On test trials when the familiar object+label was switched, infants significantly increased their visual scanning of the speaker's mouth compared to control trials. In Experiment 2, the same procedure was followed with a new group of 14-month-olds, except that the speaker now looked in the opposite direction from the objects on the screen, but continued to label them with minimal pairs in an infant-directed style. In contrast to the results of Experiment 1, infants in this latter experiment did not differentially attend to any area of her face during the switch trials. This pattern of results shows that the ostensive nature of a social partner augments infants' referential learning under cognitive challenge, but it is the contingent nature of the speaker's regard to what is being labeled that is a necessary factor in promoting learning.

## TABLE OF CONTENTS

Introduction.....	7
Emergentist Coalition Model.....	8
Resource Limitation Hypothesis.....	10
Integrating the Emergentist Coalition Model and the Resource Limitation Hypothesis.....	11
Is Contingency a Necessary Feature of Word-Object Learning?.....	12
Can Ostensive Cues and Contingency Attenuate Cognitive Load?.....	17
Purpose of this Study.....	23
Contribution to the Literature.....	24
General Method for Experiments 1 and 2.....	25
Measures and Materials.....	25
Procedure.....	28
Data Preparation and Areas of Interest.....	31
Primary Dependent Measures.....	32
Experiment 1.....	33
Participants.....	33
Results for Experiment 1.....	34
Discussion for Experiment 1.....	40
Experiment 2.....	42
Participants.....	42
Results for Experiment 2.....	43
Discussion for Experiment 2.....	49
General Discussion.....	50
Limitations and Conclusions.....	58
References.....	60

## LIST OF TABLES

Table 1. Means, standard deviations and ranges for language skill and temperament metrics for infants in both experiments. ....	67
Table 2. Correlations between percent fixation duration on test trials, language skill, and temperament (Experiment 1) .....	68
Table 3. Correlations between percent fixation duration on test trials, language skill, and temperament (Experiment 2) .....	69

## LIST OF FIGURES

Figure 1. Dynamic AOIs for Hi Baby, Forward and Turned segments .....	70
Figure 2. Average percent fixation duration by Trial Type (Switch, Same) and AOI (Object, Eye, Mouth) for Experiment 1 .....	71
Figure 3. Average percent fixation duration by Trial Type (F4, Switch) and AOI (Object, Eye, Mouth) for Experiment 1. ....	72
Figure 4. Average percent fixation duration by AOI (Object, Eye, Mouth) for Experiment 2 ...	73

## LIST OF APPENDICES

Appendix A - i-LEAP Laboratory Family Information Sheet .....	74
Appendix B - Infant Behavior Questionnaire-Revised Very Short Form.....	75
Appendix C - Ages and Stages Questionnaire .....	76
Appendix D - MacArthur Communicative Development Inventory-Short Form .....	77
Appendix E - Maternal Sensitivity Task.....	78
Appendix F - Diagram of Experiments.....	80
Appendix G - Objects Used in Word-Object Learning Task.....	81
Appendix H - Informed Consent.....	82

## Introduction

Toward the end of their first postnatal year, infants are transitioning from being primarily passive recipients of language to active users as their comprehension and production skills increase (Golinkoff, Mervis, & Hirsch-Pasek, 1994). Infants are motivated to share experiences with the people around them and, having observed language as fundamental to this sharing, they must master the communication system therein (Locke, 2001). As infants begin active use of language, they are dependent on a variety of cues to guide their association of labels and their referents. These cues require infants' attention across general, linguistic, and social domains (Hollich et al., 2010). For example, infants may attend to the temporal synchrony (general domain) between the presentation of an object and hearing a label (linguistic domain) when a caregiver is looking at the infant and object while producing the label (social domain). It is well established that infants are adept perceivers of cues from these three domains and can use the cues to differing degrees as necessary. It is not well known, however, if specific challenge in one domain (e.g., linguistic) can be abated if attention can be increased to other domains (e.g., general and social). This study investigated whether infants' attention to the cues provided by communicative partners and the contingency between a speaker's object regard and their labeling of the object when associative word-object learning was challenging. Extending previous work showing significantly diminished associative learning between novel objects and novel labels when the labels are highly similar, the current study tested whether mutual gaze, infant-directed speech (IDS) and eye gaze-object regard contingency augmented 14-month-olds' attention to difficult object-label relations.

Various models have been developed to explain the necessary and sufficient conditions under which such associative learning occurs in infancy. Two such models, *emergentist*

*coalition and resource limitation hypothesis*, share ideas about how attention and critical features impact infants' word learning. The emergentist coalition model focuses on how the developmental status of the child interacts with available cues across general, linguistic, and social domains to affect lexical growth. However, the emergentist coalition model does not explicitly address how the issue of cognitive competition or challenge at a given age, arising from demands of immediate processing contexts, compromises successful associative learning. If integrated (especially across linguistic and social domains), these two models can more effectively be used to investigate specific properties (both within the child and within the learning context) of referential learning in young toddlers.

### **Emergentist Coalition Model**

The Emergentist Coalition Model (ECM) was proposed by Hollich and colleagues (2010) to organize and synthesize a wide variety of influences on infant word learning. ECM posits two phases of development required for associating a given referent with its label: in the first phase, infants are considered language “perceivers” whereas in the second phase, infants are considered language “users” (i.e., learning to associate what they perceive with other meaningful events). During these phases the information available to infants is differentially weighted based on developmental skills. Phase I is characterized by infants' sensitivity to cues that generally elicit and maintain attention (e.g., salience, temporal synchrony, contiguity). For example, when labels are synchronized with objects, 7- to 8-month-olds are better able to discriminate the labels from each other than when the object-label presentations are asynchronous (Gogate, 2010). Phase II is characterized by infants' emerging processing within two additional domains: social and linguistic. Social cues include both verbal and nonverbal information that directs attention (e.g., pointing, eye gaze, the use of infant-directed speech style - collectively called ‘ostensive

cues'), whereas linguistic cues direct attention to specific information about native-language features (e.g., word stress, sentence prosody, phonotactics). A communicative partner delivers all of these cues in ways that scaffolds infant attention toward objects, events, and/or features of interest about which the partner will share information.

In ECM, ostensive cues are relegated to Phase II because infants' prioritization and understanding of these does not take place until after the first postnatal year. Typically-developing infants' bias toward ostensive cues is evident in episodes of joint attention in which infants' follow (and later initiate) bids to share attention on a common target. For example, within episodes of joint play with toys, mother and infants produce more vocalizations and engage in longer exchanges than when they are not jointing attending (Tomasello & Farrar, 1986). In ECM, linguistic cues are relegated to Phase II because infants' knowledge about what kinds of sounds comprising native language as well language-specific phonotactics is emerging late in the first postnatal year (Werker & Tees, 2002). Infants must be perceptually attuned to the sounds, prosody and statistical probabilities of their native language. This knowledge promotes infants' ability to segment streams of fluent speech into meaningful units (e.g., words) and to understand the morphological and syntactic patterns of their language. For example, 12-month-old English-learning infants can associate multisyllabic words differing only by syllable stress (e.g., BEdoka vs. beDOka) with two distinct objects because they are familiar with the stress patterns of English (Curtin, 2009).

Across this transition (i.e., from perceiver to user), ECM predicts that successful word learning requires 1) sensitivity to all three domains of cues, 2) that the importance of and focus on cues within each domain vary as words are learned and associated with their referents, and 3) utilization of cues moves from domain-general to domain-specific as infants develop the

knowledge that referents have labels (i.e., nominal insight; Hollich et al., 2000). As mentioned above, one strength of ECM comes from its recognition that how an infant attends to and extracts information from the environment is a combination of the level of expertise of the infant and the cues that are available to the infant that can assist them based on their developmental status. More experienced infants are able to better attend to and utilize cues from all three domains, whereas emerging word learners' attention is focused on general information as those are the only cues they are sensitive to and prioritize at this early stage of language development. Although this model acknowledges infants' emerging abilities, it does not specifically speak to how infants struggle to learn word-object relations as a function of the limits of their cognitive resources (i.e., how infants' distributed attention may be a part of the challenge they face when associating words and objects).

### **Resource Limitation Hypothesis**

The Resource Limitation Hypothesis (RLH) was initially proposed by Werker, Fennell, Corcoran, and Stager (2002) and speculated that as infants transition from being passive consumers to active users of their language, there is a qualitative shift in the demands language learning places on them. When becoming accustomed to the sounds and features of their native language, infants' attention is more likely focused on phonotactic cues compared to more experienced infants. In contrast, infants who are transitioning out of this stage and into the mapping of those native sounds onto meaning face a distinctly different cognitive task. Associating words with their referents requires the unification of two respective cognitive representations: one for the object and one for the label itself. This shift in task demands and infants' subsequent performance is the basic premise of the RLH (Werker et al., 2002).

These ideas are echoed by Naigles (2002) who discusses discrepancies in infants' perception and production of speech-like stimuli. Early in development, infants are adept at detecting auditory patterns consisting of both linguistic and non-linguistic sounds, whereas later in the first postnatal year, producing complex linguistic patterns appears challenging. That is, the richness of events that can be accurately perceived falls short of that which can be produced. To resolve this paradox, Naigles (2002) suggests that younger infants are attending and processing form only, whereas older infants are attending, processing, and integrating form *and* meaning. In other words, the cognitive demands of linguistic tasks change as a function of infants' stage of language learning. When infants are seeking to produce language and attempting to integrate representations of words and their referents, their attention is divided. Of the many variables known to influence both amount and distribution of attention in the word learning context, which are likely available and effective in lessening growing cognitive demand?

### **Integrating the Emergentist Coalition Model and the Resource Limitation Hypothesis**

Interestingly, recent studies have explored infants' ability to attend to and remember aspects of visual scenes ranging from low to high complexity. For example, in one study it was found that neither presentations perceived as "too simple" nor as "too complex" sustain infants' attention enough to warrant effective learning (called the "Goldilocks Effect"; Kidd, Piantadosi, & Aslin, 2012). The Goldilocks Effect illustrates a mechanism through which ECM cue differentiation and reliance can be explained while at the same time considers the allocation of cognitive resources. With respect to cognitive load, tasks requiring little to no allocation of cognitive resources do not elicit attention, whereas highly complicated or difficult tasks require more cognitive and attentional resources than are available to the perceiver. Accordingly,

implicit in the ECM is the notion that infants differentially rely on cues from a given domain when cues from other domains are too simple or too complex. For example, when infants hear subtly different labels for objects (e.g., “bin” vs. “din”), they rely more on social cues to assist them in creating the required association. It is important to remember, however, that perception of linguistic and social (ostensive) cues is dependent upon knowledge of domain-general cues. For example, one particular domain-general pattern that infants observe in their environment and that holds their attention is contingent relations between events. That is, infants are able to detect patterns of cause and effect that facilitate their learning, as the outcome of the precipitous event is probabilistically certain.

### **Is Contingency a Necessary Feature of Word-Object Learning?**

The ECM minimizes the role of contingency learning as a mechanism through which infants experience reinforcement and come to associate labels with their referents. Infants’ sensitivity to contingencies is often illustrated through experiments in which a change in infants’ behavior is consequated with a known reinforcing outcome. Examples of such contingencies in newborns include sucking on a pacifier and hearing either the maternal voice (DeCasper & Fifer, 1980) or a woman singing (DeCasper & Carstens, 1981). These experiments demonstrate that contingent outcomes based on infant behavior increase the rate/frequency of that behavior. Decreases in behavior, frustration, and negative affect follow removals of rewarding contingent outcomes (DeCasper & Carstens, 1981).

Many experiments assessing the learning and memory of slightly older infants have featured ‘contingency mobiles’ that move as a function of the infants’ behavior. For example, Watson and Ramey (1972) configured mobiles to rotate when 8-week-old infants made small head movements. Compared to two non-contingent control groups, infants in the contingent

condition activated their mobiles more within experimental sessions and across the duration of the study. All three groups of infants moved their heads to similar degrees at the beginning of the study, but increases in head movements were not observed for infants in the non-contingent groups. Also, complementary to DeCasper and Carstens (1981), positive affect following contingency learning was evidenced through anecdotal reports from mothers that their infants were happier, better focused in other kinds of play, and sleeping better than they were prior to the study (Watson & Ramey, 1972). Further, Rovee-Collier's research utilized mobiles to explore the importance of contingency for infants' learning and memory (e.g., Rovee-Collier, 1984).

Johnson, Posner, and Rothbart (1991) examined the development of young infants' learning of contingent relations in the visual domain through measures of anticipatory looking. Infants were shown two orienting stimuli in the center of a screen. During training, depending on which stimulus was present, one of two secondary stimuli then appeared. For example, when object A came on at center, object B would appear on the left side of the screen, and when object C was shown, object D would appear on the right side of the screen. After training and during test, the experimenters measured whether infants would disengage their attention from the center stimulus and orient their attention in anticipation of the associated side of its contingent stimulus. Testing infants at various ages, the authors found that 4-month-olds, but not 2- or 3-month olds, were able to release their attention and re-orient their attention to the cued location (Johnson, Posner, & Rothbart, 1991).

Eye-tracking studies with infants between 6 and 8 months have explored how infants would respond to presentations of contingency on their own gaze. As with newborns, productions of behaviors leading to rewarding contingent outcomes increased in rate and

frequent. In addition, reaction times to produce the behaviors increased under contingent conditions (Kenward, 2010; Wang et al, 2012). Studies of this variety demonstrate that infants are capable of recognizing that their own eye gaze can generate rewarding outcomes, and the contingency only needs to be demonstrated a few times for expectation of reinforcement to be observed.

Collectively, infants' ability to perceive contingency appears early in the first year and includes sensitivity for both explicit (i.e., reinforce depends on infant's response) and implicit (i.e., contingent relations between events not requiring a response) relations. Contingency detection remains critically important for infant learning, as contingent relations are embedded within infants' interactions with their environments and, inevitably, within their interactions with people. Importantly, the social domain is a context in which infants experience explicit contingencies stemming from their behavior in ways that have been demonstrated to be critical for general, as well as specific (e.g., language) learning. For example, Bhat, Galloway, and Landa (2010) explored social attention and associative learning in 6-month-old siblings of typically developing children and children with autism spectrum disorder (ASD). Seated in a high chair with a joystick attached to their hands, infants viewed their mother and an engaging toy. The engaging toy was only activated when it was under joystick control. During the first and last two minutes of the experiment, infants' movement of the joystick had no effect on the activation of the toy. However, during acquisition, the toy lit up and played music contingent on infants' bending of the joystick. For each of the three stages, mothers were instructed to be silent at the beginning, only smiling when the infant initiated interaction. After a specified amount of time, mothers were to initiate scripted, social engagement that prompted manipulation of the joystick. During acquisition, the mothers' script also included verbal reinforcement for toy

activations. Across all three phases of the experiment, the timing and nature of the encouragement provided by mothers was controlled, and analyses confirmed that there were no significant differences in the duration of mothers' verbal feedback. Regardless of their sibling's diagnostic status, infants in both groups readily learned the contingency between their hand behavior and the actions of the toy (Bhat, Galloway, & Landa, 2010).

Most relevant to the current study, siblings of typically developing children gazed at their mothers significantly more than siblings of children with ASD (Bhat, Galloway, & Landa, 2010). Thus, although both groups exhibited effective learning, they may have relied on different cues given that the authors did not include a mother absent condition. That is, the siblings of typically-developing children may have required the presence of their own mother in order for contingency learning to be effective, whereas siblings of children with ASD are not as reliant on ostensive cues provided by their mothers for contingency learning. In other words, most infants and children should have access to contingent relations in the general domain and show effective learning under some conditions. Importantly, both infants and children who do not perceive contingencies in the social domain may experience specific deficits regarding tasks that greatly benefit from abilities to perceive and act on cues from others.

Goldstein, King, and West (2003) examined the impact of contingent reinforcement from mothers on their 8-month-old infants' vocalizations. During a play interaction, half of the mothers in the study were instructed to smile, touch, and move closer to their infants when they babbled in a way that produced a sound that may be perceived as lexical (e.g., repeating the phoneme "ma" and its resemblance to "mom"). The other half of mothers interacted with their infant only when directed to by the experimenter. The timing of the directives provided by the experimenter were identical to the timing of interactive episodes from a mother who was able to

interact based on their infant's behavior. This manipulation maintained the social contingency for some infants and intentionally disrupted contingent social reinforcement for the infants' whose mothers' reinforcement was not the result of their own behavior. Infants whose mothers responded contingently when infants vocalized produced more and higher quality canonical speech sounds compared to infants whose mothers' responses were under experimental control. This study demonstrated that in human infants' contingent social reinforcement of canonical vocalizations was necessary to improve infants' speech-like productions (Goldstein, King, & West, 2002). These findings could be extrapolated to include situations in which a caregiver reinforces vocalizations that sound like words, not just phonemes. In addition to speech production, social contingencies also have critical implications of the reception and acquisition of language.

In a study of non-native phoneme learning, Kuhl, Tsao, and Liu (2003) exposed 9- and 10-month-old English-learning infants to a woman speaking infant-directed Mandarin Chinese. Infants either observed the woman live (and in-person) or recorded. Learning of non-native phonemes was only demonstrated for the infants who experienced the live social interaction, and those infants in the recorded condition attended to the woman significantly less than the live group. When interpreting these results, Kuhl, Tsao, and Liu (2003) attribute the benefits of the live learning condition to the infant-directed speaking style and specific ostensive cues that promote attention and facilitate learning. This interpretation, however, does not explicitly consider the importance of contingent social interaction. During naturalistic social interactions, the gaze and gestures provided by social partners is dependent upon the receiving partner. When watching a recording, the ostensive cues provided did not change, but the specific contingency between the adult and child was removed. As the finding by Goldstein, King and West (2003)

suggests, if infants in the recorded condition imitated or spontaneously produced a speech sound like one found in Mandarin, it would not have been possible for the woman to respond to and reinforce that vocalization. Together these studies highlight the ways in which social contingencies are crucial for infants' development of receptive and expressive language.

### **Can Ostensive Cues and Contingency Attenuate Cognitive Load?**

As mentioned earlier, the gestures, movements and speaking style used between social partners can be labeled as *ostensive cues*, which include both verbal and nonverbal gestures that elicit, direct, and maintain the attention of a social partner (Csibra, 2010). These effects on the social partner's attention are especially important because they also communicate to the recipient intentionality on behalf of the speaker (Csibra, 2010). Accordingly, ostensive cues may serve as an indicator to the recipient that information communicated by their social partner is for them, further recruiting and directing attentional resources of the recipient. Ostensive cues include such behaviors as eye gaze, head-turn, pointing, positive affect and, for infants, the use of an infant-directed speaking style.

In particular, infants' attention to eyes and the target of their gaze has been demonstrated in the first few months of life, and eye gaze is consistently followed by typically developing infants by the age of 6 months (e.g., Hood, Willen, & Driver, 1998; Senju & Csibra, 2008). Frank, Vul, and Saxe (2012) measured infants' attention to a variety of dynamic, social visual displays between the ages of 3 and 30 months to examine what features of people are most salient. Infants watched videos of faces alone, whole individual people, and multiple people. When faces were presented in isolation, the eye and mouth areas occupied almost all of the infants' attention. When attention to the eyes and mouth were compared, attention to the eyes decreased with age, whereas attention to the mouth increased, such that infants younger than 16

months attended to the eyes more than the mouth and vice versa. When it was possible to view the whole body of a person, infants continued their attention to faces but also increased their attention to others' hands (this increased steadily with age). This pattern was even more robust in scenes containing multiple people in interaction. Further analyses considered these patterns as a function of mouth movement. If a mouth was static, although attention decreased with age, attention to the eyes was still greater than attention to the mouth. If a mouth was smiling or speaking, eyes attended more to the mouth only for infants less than 10 months old (Frank, Vul, & Saxe, 2012). Overall, these findings indicate that infants attend to eyes fairly early in infancy and this attention and preference persists until later in the first year when mouths and hands take on increased importance.

Returning to gaze, Wu and Kirkham (2010) familiarized 4- and 8-month-old infants to two object-sound pairings (e.g., a toy cat and "bloop") on an eye-tracker to examine developmental changes in infants' social attention and gaze following. First, a woman's face appeared in the center of the screen, smiled and spoke to the baby in IDS, and looked to one of the two bottom corners where an object might appear. Next, the sound "bloop" was heard and identical pictures of a toy cat appeared in the top left and bottom right corners of the screen. The woman's gaze always served as a cue to the location at which an object would appear (i.e., in the example just given, the right lower corner). When tested, infants simply heard a target word (the woman's face was not provided). Gaze to each of the four corners was measured. Evidence was found that 8-month-olds, but not 4-month-olds, would look at the corner signaled by the word *and* the direction of the woman's gaze. Surprisingly, a subsequent experiment found that when object location was cued by color instead of gaze, both ages of infants successfully acquired correct target locations (Wu & Kirkham, 2010). This latter finding demonstrates that the

younger infants have the ability to learn location-object relations, but do not use adult gaze as a reliable cue to location. Building on these findings, Paulus and Fikkert (2014) examined 14- and 24-month-old infants' preference for gaze and pointing, and found that 14-month-olds were more reliant on gaze than pointing in a word learning task, whereas 24-month-olds preferred to follow pointing.

Together, the findings of these studies demonstrate that the ostensive properties of social cues become apparent to infants later in the first postnatal year and become increasingly differentiated into the second year. Infants transition from passive consumer to active user of language, including both ostensive and linguistic cues, requiring their attention to be dynamic and differentially reliant on available information across domains. Social interactions that feature co-regulation of attention may assist in infants' learning by introducing contingency between ostensive and linguistic cues. Such co-regulation is transparent in the emergence of joint attention (JA) which is generally defined as a triadic configuration of two social partners and a target of their shared attention (Dunham & Moore, 1995). Joint attention requires at least two social partners to initially establish mutual regard (e.g., eye contact), and then, contingent upon the attention of partner A, partner B directs the attention of the other toward a particular target, facilitating a shared and co-regulated experience (Tomasello, 1995).

A significant positive correlation between JA toward the end of the first year and expressive language in the second year of life has been found repeatedly (e.g., Baldwin, 1995; Morales et al., 2000; Mundy, Fox, & Card, 2003; Salley & Dixon, 2007; Salley, Panneton, & Colombo, 2013). Accordingly, the development of joint attention expertise is considered to be a critical component of language acquisition. When shared attention is successfully co-regulated between social partners to a particular event by attention-contingent ostensive cues, infants are

more likely to form associations between what they observe and the label they are provided with. Decoupling caregiver's shared attention and the labeling of objects does not facilitate infants' language learning nearly as much as when they are coupled (Baldwin et al., 1996). Given this finding, contingent interactions embedded within episodes of joint attention should be most beneficial when infants are struggling to associate objects with difficult labels.

For example, words that differ by a single phonetic element (e.g., bat v. cat; baby v. vaby) fill a special niche within infant language learning research, as they require acute perception and attention to subtle phonemic similarities in their language. That is, the use of minimal pair labels in referential learning tasks presents infants with a deliberate cognitive challenge. For instance, infants must learn the association between a given object and its label amidst other objects with labels that share similar sounds and compete for similar cognitive resources (Swingley & Aslin, 2002). It has been repeatedly demonstrated that 14-month-olds fail to show evidence of object-label learning when the labels are highly similar, even though these infants can clearly discriminate these similar-sounding words. Interestingly, younger infants (e.g., 8-month-olds) successfully discriminate highly similar labels used in the presence of one object but 14-month-olds fail this task (Stager & Werker, 1997). The RLH was put forth as a possible explanation for these seemingly disparate findings. Younger infants' (e.g., 8-month-olds) attention has not yet shifted to the meaning of the words, so they focus on and encode the subtle phonemic contrasts. As such, the cognitive load of the task does not exceed the resources necessary for discrimination. However, older infants (e.g., 14-month-olds) have made such a transition, and in doing so, these infants do not have the cognitive resources to attend sufficiently to both the phonemic contrast and the characteristics of the label's referent.

To date, some studies have successfully attenuated cognitive load in minimal pair object-word learning (e.g., familiarization to the labels prior to the associative task; Werker & Fennell, 2004; increased salience of the phonetic contrast; Fennell & Waxman, 2010; Fennell & Byers-Heinlein, 2014). However, only one study has attempted to assist minimal pair associative learning through the use of cues in the social domain. Fais and colleagues (2012) conducted two experiments to address whether the addition of a social partner would assist 14-month-old infants with the challenge of associating minimal pair labels to objects. In the first experiment, infants were seated on the laps of their caregivers facing screen. Objects were presented on the screen by the experimenter in a typical visual habituation protocol. What sets this experiment apart from others, however, is the use of a live interactor. This potential social partner was seated perpendicularly next to the infant so her gaze could overtly shift back and forth between the infant and the visual display. This live experimenter also produced the minimal pair labels for the objects when they appeared. Interestingly, and potentially unnaturally, the labels were provided in isolation, as opposed to being presented in a carrying phrase. After the established habituation criterion had been met, two test trials were conducted, as like the previously described Switch studies, consisted of one trial like those during habituation and the critical Switch test trial in which one of the objects was presented but labeled with the label provided for the other object. Pre- and posttests were also administered to demonstrate infants' general attention to presentations of a third object and non-minimal pair label was measured. Two cameras were utilized to capture both the infants' looking to the visual presentation (i.e., a camera facing the infant) and the looking patterns between the experimenter and the infant (i.e., a camera facing the experimenter with the infant visible). Looking time was the primary dependent measure of interest. Infant movement was also assessed but will not be discussed

here. As demonstrated through a mixed, repeated measures ANOVA, there were no significant main effects of or interactions between test trial type, test trial order (Same- or Switch-first), and gender. Thus, 14-month-old infants' difficulty for minimal pair referential learning was again demonstrated at the group level. Further analyses, however, were utilized to address individual differences in performance.

For all test trials, including pre- and posttests, total accumulated looking time to experimenter, time spent looking jointly at the object with the experimenter, and time spent in mutual gaze with the experimenter were determined for each infant. Median splits were applied to these data to classify infants as "high" or "low" for each looking parameter. Examination of these individual differences demonstrated systematic effects of accumulated mutual gaze time on minimal pair Switch performance such that infants with high mutual gaze time showed recovering of attention on Switch trials, whereas infants with low mutual gaze time did not.

In a second experiment, infants viewed a woman who looked to and labeled objects much like before, but in this version, the objects and the woman's face were viewed on the screen. No live experimenter was present. The timing of the woman's looking to objects and looking to the infant were based on the timing from the live experimenter in the first experiment. In doing this, the timing was maintained but the gaze shifting was no longer contingent on the infants' behavior. Following this protocol, infants' difficulty with minimal pairs was maintained, regardless of individual differences.

Removing the in-person production of object labels also removed the specific contingency between the experimenter and the infant. In the recorded version of the experiment, the woman's eyes begin facing the infant but her gaze shifts to the object based on a specified period of time. Mutual attention in the live version triggers the gaze shift to the object. In this

way, mutual gaze provides a means through which infants' attention can be attracted, directed and maintained. The ostensive cues experienced and reciprocated by infants in the high mutual gaze group may have experienced decreased the cognitive load of the task either through increased attention across domains as a function of increased attention to ostensive cues or the removal of the cognitive requirements of focusing their attention to the object and the label themselves. The question remains as to whether attention to ostensive cues or the need for environmental contingencies serves as a mechanism for abating cognitive load.

### **Purpose of this Study**

The purpose of the experiments comprising this study was to examine the potential impact of ostensive cues (i.e., mutual gaze and infant-directed speech) and eye gaze-object regard contingency from a simulated social partner on 14-month-old infants' object-label learning under increased cognitive load. Experiment 1 investigated whether ostensive cues from a simulated social partner augmented 14-month-old infants' associative learning of minimal pair object labels when eye gaze-object regard contingency was maintained. Infants were presented with a still image of a female on a screen, with her gaze forward (toward the infant). Once the infant looked at her eyes, familiarization videos began in which she greeted the baby (i.e., "Hi, Baby!") in an infant-directed style, and then turned her head and eyes toward an object on the screen. Following her regard of the object, her face returned forward and she produced its associated verbal label (this was repeated 8 times during a given trial). Familiarization trials were provided for two different object-label pairs. After familiarization and during test, infants were presented with each of the familiar combinations (i.e., ObjectA+LabelA; ObjectB+LabelB), and also to a 'switch' trial in which Object A was presented with the speaker saying Label B. In this paradigm, infants increase their attention on switch relative to familiar test trials if they perceive

the object-label violation (e.g., Werker, Cohen, Lloyd, Casasola, & Stager, 1998). It was hypothesized that the attention that is elicited by the ostensive cues available from a dynamic face and voice would facilitate infants' processing of object-label presentations, and learning of the word-object associations. Accordingly, infants would attend differentially on switch trials in spite of their familiarization with minimal pair labels.

Experiment 2 investigated the importance of eye gaze-object contingency for infants' associative learning. Infants' attention was elicited in the same way as Experiment 1, but a dynamic face breaks eye gaze-object contingency by not looking at the object being labeled. In studies of infants' eye gaze following like those previously described (i.e., Fais, et al, 2012; Paulus & Fikkert, 2014; Wu & Kirkham, 2010), the social partner looks to an object. In Fais et al (2012), the experimenter shifted her eye gaze to the object to regard it before providing the label. If speaker ostensive cues are sufficient for the regulation of attention, then the violation of eye gaze-object contingency should not attenuate associative learning. However, if eye gaze-object contingency is a necessary component of associative learning, infants should not show evidence of associative learning in Experiment 2.

### **Contribution to the Literature**

The study by Fais and colleagues (2012) was not the first study to demonstrate successful referential learning with minimal pairs in this age group, but it was the first to utilize an interactive female partner presenting both facial and vocal cues supporting infant learning. Importantly, the face+voice alone was not enough to benefit all the infants in their sample, as success was dependent upon gaze sharing between a subset of the infants and the adult female. Although the timing was similar across the Fais et al.'s live and simulated conditions, the recorded face did not have the ability to interact contingently as did the live experimenter. It was

also not possible to specifically quantify the locations of infant's viewing of the visual presentation in the simulated condition. The current study approximated naturalistic ostensive cues with a social partner while utilizing an eye-tracker. This technology allowed for precise and specific measurement of the locations to which the infant attended. In particular, specific fixation duration on facial features/regions and the object to be labeled have been quantified and compared in this study.

## **General Method for Experiments 1 and 2**

### **Measures and Materials**

**Questionnaires.** Caregivers completed questionnaires addressing demographics (Appendix A), infant temperament (the Infant Behavior Questionnaire; Appendix B), the Ages and Stages Communication and Personal-Social Scales (Appendix C), and the short form of the MacArthur-Bates Vocabulary Checklist Level 1 (Appendix D). Psychometric properties of the measures are provided below.

***Infant behavior questionnaire.*** The Infant Behavior Questionnaire (IBQ; Gartstein & Rothbart, 2003) is designed to assess three dimensions of infant temperament: positive affectivity/surgency (PAS), negative emotionality (NEG), and orienting/regulatory capacity (ORC). Parents rate their child's behavior on a Likert scale (1=never to 7=always, or not applicable) on 38 items (i.e., 12 items per dimension with two non-target questions). Parents were instructed to consider their child's behavior over the last two weeks. A score for each dimension is calculated by averaging the items for that dimension, reverse scoring where indicated. Omitted items or those rated as not applicable were omitted from these calculations. These dimensions have good internal consistency (PAS  $\alpha = .77$ ; NEG  $\alpha = .78$ ; ORC  $\alpha = .75$ ; Putnam, et al (2013)).

***Ages and stages questionnaire.*** The Ages and Stages Questionnaire (ASQ; Bricker & Squires, 1999) is designed to assess developmental delays in children at many intervals from one month to five years of age. The Communication (ASQ-C) and Personal-Social (ASQ-PS) subscales of the 14-month-old ASQ was utilized to assess infants' communicative and social motivation. Each subscale consists of 6 questions answered with "yes," "sometimes," or "not yet." Points are awarded for each answer (i.e., 10 for "yes", 5 for "sometimes," and 0 for "not yet") and totaled for each subscale. These subscales have good to acceptable internal consistency (Communication  $\alpha = .73$ ; Personal-Social  $\alpha = .63$  for 14-month-old infants (Squires, Twombly, Bricker, & Potter, 2009)). These subscales are correlated  $r = .54$  (Squires et al., 2009).

***MacArthur-Bates vocabulary checklist.*** The short form of the MacArthur-Bates Vocabulary Checklist for infants (MCDI; Fenson et al., 1996) is a questionnaire designed to measure receptive (MCDI-RV) and expressive vocabulary (MCDI-EV) size in children 8-18 months of age. Caregivers indicate whether infants can understand and/or say any of the 89 words listed. This measure has been shown to be a reliable (i.e., Cronbach's alphas  $> .95$ ) and valid assessment of receptive and expressive vocabulary (Pearson correlations of .88 for reception and .90 for expression; Fenson et al., 2000). Fenson et al. (2000) provides normative percentiles by gender for receptive and expressive vocabularies. Fourteen-month-old male infants in the 50<sup>th</sup> percentile should understand 30 words and produce 5 words, whereas 14-month-old female infants in the 50<sup>th</sup> percentile should understand 38 words and produce 8 words (Fenson et al., 2000).

**Equipment.** The auditory+visual events shown to the infants were displayed on the 17" monitor of a Tobii © T-60 binocular eye-tracking system using Tobii Professional Studio 3.3.2

software. A separate desktop computer (Dell Optiplex 755) and monitor were used to observe infants' scan patterns during the procedure and to output data from the eye-tracking system. Infants were seated in a high chair (Experiment 1: n=11; Experiment 2: n=5) or on their caregiver's lap<sup>1</sup> (Experiment 1: n=6; Experiment 2: n=13) approximately 65 cm from the eye-tracking monitor. Either way, the caregiver was instructed to not interact with the infant during the procedure unless the infant began to cry. If the infant was seated on the caretaker's lap, dark non-translucent goggles were worn by the caretaker so as to not confuse the eye tracker. Sounds heard by the infants during the procedure were standardized to 65-70 dB SPL (A scale) at the head of the infant, and played through speakers positioned on either side of the eye-tracker.

**Audiovisual events.** Visual Studio and Adobe Premiere © were used to create all video clips by superimposing an object onto a recording of a female speaker with her facing forward and then shifting her head and eye-gaze to one side of the viewing area and then returning her head and eye-gaze back to the forward position while speaking a label. During all trials, the object moved in a counterclockwise direction. These videos of the female speaker began with her saying "Hi, baby" before alternating her gaze from forward (looking at the infant) to a side (toward an object for Experiment 1 and toward empty space for Experiment 2) and were presented during familiarization and test trials. When looking forward, the woman produced an object label in a typical infant-directed register. The woman was visible from her shoulders to the top of her head (Appendix F), with her face subtending 20° of visual angle.

Four single-syllable, consonant-vowel-consonant (CVC) word forms were used as minimal pairs in the experiments (counterbalanced across infants). One pair was /bin/ as in *cabin* and /din/ as in *dinner* and the second pair was /gem/ as in *gemote* and /pem/ as in *Pembroke*.

---

<sup>1</sup> The video recordings of three infants in Experiment 1 were lost and could not be reviewed due to experimenter error.

Both pairs have short, center vowels and nasal glide ending consonants so that they differ in first consonant only. All of the initial stop consonants differ in production as a function of mouth placement (i.e., /b/ is labial and /d/ is dental; /g/ is velar and /p/ is labial).

The objects used in this task were four unfamiliar objects with unusual geometric shapes and bold colors (Appendix G). Each infant was familiarized to only two of these objects (Object A/Object B or Object C/Object D). Each object was presented on the left side of the screen in one trial and on the right side in one trial (just off the ear level of the female speaker). All objects were the same size and subtended 8° of visual angle.

The order of presentation of the object-label pairs were counterbalanced across infants (i.e., each infant was familiarized to only one of the following word orders: Bin/Din, Din/Bin, Gem/Pem, Pem/Gem). In addition, each minimal pair was always presented with the same two objects but the relation of the object and label was counterbalanced within pairs. For Bin/Din and Din/Bin, Object A was always presented first and third, and Object B was always presented second and fourth. For Gem/Pem and Pem/Gem, Object C was always presented first and third, and Object D was always presented second and fourth. In addition to the familiarization and test videos, an audiovisual clip of a laughing baby (attention getter) was used between trials to orient infants' attention to the screen.

## **Procedure**

Prior to testing, all caregivers completed informed consent materials (Appendix H) as well as demographic and ASQ questionnaires. Upon the completion of initial paperwork, infants participated in an object-label learning task on the eye-tracker. A certificate of thanks and a copy of the consent form was given to the caregiver(s) at the end of the session. The IBQ and

the MCDI were sent home in a self-addressed and stamped envelope for caregivers to complete and return within one week.

The following description refers to the eye tracking procedure for both Experiments 1 and 2. Although the general procedure was identical for these two experiments, the important difference was in the nature of the female speaker's gaze during familiarization and test. For Experiment 1, her gaze was always toward the object on the screen (to her left or right), whereas for Experiment 2, her gaze was always away from the object on the screen (to her left or right). In both cases, the object is named while the female speaker is positioned forward, looking at the infant (Appendix F).

***Gaze Calibration.*** The infant was seated facing the eye tracker within a sound-attenuated room to listen to and watch the audiovisual presentations. Each infant's gaze was calibrated using an animated, audiovisual event presented at 5 fixation points (i.e., corners and center of the screen). Importantly, the calibration stimulus began in one location and did not advance to subsequent locations until a fixation occurred on that stimulus. Infants were considered successfully calibrated when their points of fixation were isolated within the physical center of the calibration stimuli. Calibration was repeated for fixation points that showed "splatter" outside of the target location. Such a calibration procedure gave the eye-tracker a way of verifying the location of fixations and provide an estimate of validity for fixations (i.e., how "confident" is the eye-tracker about a given fixation).

***Familiarization Trials.*** Each infant was exposed to two word-object pairings over four familiarization trials (i.e., LabelA+ObjectA on the left; LabelB+ObjectB on the left; LabelA+ObjectA on the right; LabelB+ObjectB on the right;). Prior to the onset of familiarization trials, infants' attention was focused on the screen by the attention getter for 2 s.

Next, the still image of the female speaker came on the screen. Familiarization trials began after the observer determined that the infant's attention was specifically focused within the eye region of the still image of the female speaker. The female speaker was then animated simultaneously with an object on one side of the screen. The first rotating object (e.g., Object A) was presented on the left of the screen. The female speaker greeted the infant (i.e., "Hi, baby!") in an infant-directed style and then shifted her gaze to one side of the screen, maintained that gaze briefly (approximately 0.5 s), and returned her gaze to the infant. She then said the object label (e.g., Label A). This sequence of alternating gaze between the side and the infant repeated so that the object label was spoken 8 times (27-30 s trials).

Following the termination of the first familiarization trial, the attention getter reappeared before the next familiarization trial began. The second object (e.g., Object B) was also presented on the left side of the screen. The female speaker greeted the infant and then turned her gaze and maintained her gaze briefly. She then looked back at the infant and said the object label (e.g., Label B). This sequence of alternating gaze between one side and the infant repeated so that the object label was spoken 8 times (27-30 s trials).

Pair A and pair B (i.e., ObjectA+LabelA and ObjectB+LabelB) were presented alternately with each object appearing on both sides of the screen. That is, across two familiarization trials (trials 1 and 3), ObjectA+LabelA was displayed on the left and right sides of the screen. Likewise, for the other two familiarization trials, ObjectB+LabelB was displayed on the left and right sides of the screen. After the 4 familiarization trials finished, testing began.

***Test Trials.*** There were a total of four test trials following familiarization. *Same* test trials were identical to those during familiarization (i.e., ObjectA+LabelA or ObjectB+LabelB), and *Switch* test trials were those in which the labels for the objects were transposed (i.e.,

ObjectA+LabelB or ObjectB+LabelA). Same and Switch trials were presented to all of the infants in the same order with Switch trials occurring first and last (i.e., Switch-Same-Same-Switch). Test trials lasted between 14 and 15 s, such that each object was labeled 4 times within a trial (as opposed to 8 times during familiarization). Again, the attention getter came on between trials to refocus or maintain the attention of the infant until the next trial began.

### **Data Preparation and Areas of Interest**

Tobii Studio rated the validity of individual fixations for each eye on a Likert scale from 0 to 4, with 0 as most valid and 4 as least valid. First, to be included in the dataset, any given infant needed to contribute at least 20% valid fixation data across the entire session. Second, any sampling frame for a fixation that did not have a validity rating of 0 for at least one eye was removed from the data. After removing invalid frames, any individual trial that did not retain at least 20% valid fixations was not included in subsequent analyses, even if the overall eye tracking validity was adequate.

Dynamic Areas of Interest (AOIs) were drawn on the videos to measure fixations on the object, and on the eyes and mouth of the speaker (see Figure 1). An AOI was also constructed in the non-object space (opposite of where the object actually appeared), identical in size to the Object AOI (i.e., these Non-object AOIs were positioned as if an object had appeared there). This allowed an analysis of scanning patterns to empty space in both experiments. AOIs were differentiated into segments based on the action of the speaker, including the initial “Hi, baby!” portion of the video (HB), when her gaze was forward toward the infant (Forward), and when her gaze was turned toward the object (Turned). These segments were generated such that there were four AOIs per segment (i.e., Object HB, Eyes HB, Mouth HB, Non-Object HB, Object

Forward, Eyes Forward, Mouth Forward, Non-Object Forward, Object Turned, Eyes Turned, Mouth Turned, and Non-Object Turned).

### **Primary Dependent Measures**

In the original switch studies (e.g., Werker, Cohen, Lloyd, Casasola, & Stager, 1998), evidence for infants' associative learning was manifest in their significant recovery of attention on switch compared to same test trials. That is, total looking increased when the object/label was transposed. This increase in attention is mostly the result of two important factors: (1) only objects are on the screen being viewed by the infant, allowing more non-target space to which infants can direct their attention when bored, and (2) the duration of test trials are infant-controlled such that looking away from the screen terminates the trial. The current design was slightly different in that we presented both a person and an object on the screen of the monitor at the same time, reducing non-target space, and trial duration was not under infant control. Moreover, invalid frames from the eye-tracking data were actually eliminated from the analyses which effectively reduced trial length by taking out non-attention time. For these reasons, it was unlikely that we would see significant increases in absolute fixation time during the critical test trial. Consequently, it was not possible to predict the same kind of increase in looking time as the sole measure of associative learning. However, differential looking patterns to the various AOIs of the visual information available during test trials was expected.

The first dependent measure of interest was percent fixation duration. Although all trials were of the same duration by design, the removal of fixation frames with low validity resulted in variable trial durations for analysis. Therefore, the percent fixation duration was calculated for each AOI of interest as a function of the remaining time in the trial. Further, percent fixation duration was calculated for relevant AOIs within individual segments. The second dependent

measure was fixation count (i.e., number of fixations in an AOI on a given trial and segment). After completing all relevant analyses with these two measures, it was clear that both percent fixation duration and fixation count were highly redundant (i.e., the results of analyses using these measures were virtually identical). Also, the relative lack of fixation to specific AOIs (e.g., eyes) resulted in a high number of zeros in the count measure, skewing means in a negative direction. Thus, analyses utilizing fixation count are not reported in the dissertation. The third dependent measure was latency to regard the Eyes during the initial still image of the woman.

HB, Forward, and Turned segments were analyzed separately. Further, because the eyes and mouth were less visible during Turned segments, analyses of these segments were limited to the Object AOI. Lastly, test trial analyses compared the first Switch trial and the first Same trial, due to infants' decreasing attention across test trials.

## **Experiment 1**

The purpose of Experiment 1 was to investigate whether ostensive cues (i.e., mutual gaze and infant-directed speech) and eye gaze-object regard contingency that are naturally present during referential learning contexts could augment infants' learning of object-label relations when using minimal pairs as labels. Following familiarization to two word-object pairs, infants' recognition of the relations was tested. It was hypothesized that differential looking patterns to targeted AOIs between switch and same test trials would indicate that the labels were learned, such that the trial types were processed differently.

### **Participants**

The participants in Experiment 1 were 13- to 15-month-olds ( $n = 20$ ; mean age = 14.5 mo;  $SD = 0.88$  mo; 9 female). Data collected from an additional 7 infants could not be used due to an inability to calibrate their eye position ( $n = 3$ ) or insufficient valid eye tracking data ( $n = 4$ ).

All participants were recruited from (and representative of) the surrounding geographic areas. Several recruiting strategies were employed: (1) sending a letter of invitation to parents from using the database maintained in the Department of Psychology's Developmental Research Suite, (2) listservs, (3) advertisements, and (4) placement of flyers in day care centers and local businesses. The infants in this sample were learning English as their primary language and Caucasian (18/20 infants). Infants were at least 37 weeks gestational age at birth, and all of the infants came from two-parent households. Of the 40 parents, 39 had at least some college experience, and 37 parents had obtained at least a Bachelor's degree. Fifteen of the 20 households reported annual incomes of \$50,000 or more.

## **Results for Experiment 1**

### **Looking Patterns During Familiarization**

In order to examine order effects, a 2 x 4 analysis of variance (ANOVA) with object+label pair (2: pair 1, pair 2) as a repeated factor and first word (4: Bin, Din, Gem, Pem) as a between-subjects factor on percent fixation duration to the Object in Forward segments was conducted. This analysis did not result in a significant main effect of pair ( $F(1,3)=2.04, p=.17$ ) or significant interaction of pair and first word ( $F(3,16)=1.22, p=.33$ ). Accordingly, the rest of the reported analyses below are collapsed across order of presentation.

Next, changes in infants' attention to the Object, Eyes, and Mouth<sup>2</sup> of the speaker were analyzed as they processed the audiovisual presentations during familiarization. To address this issue, repeated measures ANOVAs compared each infant's percent fixation duration as a function of AOI (3: Object, Eyes, Mouth) averaged across the four familiarization trials for the Hi Baby (HB) and Forward segments. An additional repeated measures ANOVA compare

---

<sup>2</sup> Non-object AOIs were not included in these analyses as no fixations were registered during any trial within this experiment.

infants' percent fixation duration as a function of AOI (3: Object, Eyes, Mouth) between the first and last familiarization trials for Forward segments. For Turned segments, paired samples *t*-tests compared percent fixation duration to the Object AOI only between Forward and Turned segments as well as between the first and last familiarization trials.

**HB Segments During Familiarization.** In order to examine scanning patterns when the speaker greeted the infant with “Hi baby!”, an ANOVA with AOI (3: Object, Eyes, Mouth) as the repeated factor was conducted. This analysis revealed a significant main effect ( $F(2,38)=205.11, p<.001, \eta_p^2=.92$ ). Percent fixation duration to the Mouth ( $M=.77, SD=.20$ ) was significantly larger than that to the Object ( $M=.06, SD=.08; t(19)=14.30, p<.001, d= 4.60$ ) and that to the Eyes ( $M=.01, SD=.03; t(19)=15.39, p<.001, d=5.35$ ). Also, percent fixation duration to the Object was significantly greater than that to the Eyes ( $t(19)=2.25, p=.04, d=1.06$ ).

**Forward Segments During Familiarization.** In order to examine scanning patterns when the speaker was looking at the infant and naming the object during familiarization, an ANOVA with AOI (3: Object, Eyes, Mouth) as the repeated factor was conducted. This analysis revealed a significant main effect ( $F(2,38)=112.53, p<.001, \eta_p^2=.86$ ). Percent fixation duration to the Mouth ( $M=.70, SD=.26$ ) was significantly larger than that to the Object ( $M=.05, SD=.06; t(19)=10.71, p<.001, d=3.45$ ) and to the Eyes ( $M=.02, SD=.04; t(19)=10.87, p<.001, d=3.67$ ). Percent fixation duration was not significantly different between the Object and the Eyes ( $t(19)=1.88, p=.08$ ).

In order to compare scanning patterns occurring at the start and end of familiarization, an ANOVA with trial (2: F1, F4) and AOI (3: Object, Eyes, Mouth) as repeated factors was conducted. This analysis resulted in significant main effects of trial ( $F(1,19)=12.98, p<.001, \eta^2<.001, \eta_p^2=.41$ ) and AOI ( $F(2,38)=102.93, p<.001, \eta^2=.81, \eta_p^2=.84$ ) and a significant trial x

AOI interaction ( $F(2,38)=6.70, p=.002, \eta^2=.01, \eta_p^2=.26$ ). Percent fixation duration to the Mouth was significantly larger in F1 ( $M=.76, SD=.27$ ) compared to F4 ( $M=.64, SD=.27; t(19)=3.15, p=.01, d=.44$ ). However, there were no significant differences between percent fixation to the Object ( $t(19)=1.64, p=.12$ ) or to the Eyes ( $t(19)=1.25, p=.23$ ). Within F1, percent fixation duration to the Mouth ( $M=.76, SD=.27$ ) was significantly greater than that to the Object ( $M=.04, SD=.09; t(19)=9.87, p<.001, d=3.51$ ) and to the Eyes ( $M=.03, SD=.05; t(19)=11.34, p<.001, d=3.73$ ), but no significant difference between object and eyes ( $t(19)=.63, p=.54$ ). Likewise, within F4, percent fixation duration to the Mouth ( $M=.64, SD=.27$ ) was significantly greater than to the Object ( $M=.09, SD=.12; t(19)=7.71, p<.001, d=2.92$ ) and to the Eyes ( $M=.01, SD=.03; t(19)=10.14, p<.001, d=3.64$ ). Additionally, percent fixation duration to the Object was significantly greater than to the Eyes ( $t(19)=2.57, p=.02, d=.87$ ).

**Turned Segments During Familiarization.** In order to examine scanning patterns when the speaker was looking at the object during familiarization, a paired samples *t*-test compared percent fixation duration to the object across Forward and Turned segments. Percent fixation duration to the Turned Object ( $M=.23, SD=.13$ ) was significantly greater than that to the Forward Object ( $M=.05, SD=.06; t(19)=7.57, p<.001, d=1.75$ ). This indicated that infants were more likely to look at the object after the female speaker had turned her eye-gaze toward it.

Also, a paired samples *t*-test compared percent fixation to the object between the first and last familiarization trials. Percent fixation duration to the Turned Object was significantly greater during the last familiarization trial ( $M=.29, SD=.17$ ) than the first familiarization trial ( $M=.14, SD=.11; t(19)=5.09, p<.001, d=1.00$ ) indicating increased attention to the object when it was referenced by gaze over familiarization time.

### Looking Patterns During Test<sup>3</sup>

Of primary interest was whether infants' looking patterns during the Switch test trial were different from their looking patterns during the Same test trial. To determine this, a repeated measures ANOVA compared each infant's percent fixation durations as a function of trial type (2: Switch, Same) and AOI (3: Object, Eyes, Mouth) for Forward segments. For Turned segments, a repeated measures ANOVA compared percent fixation duration as a function of the Object AOI only. Additionally, repeated measures ANOVAs were conducted to compare looking patterns on the last familiarization trial and each of the test trials as an indicator of perceived change (switch) and no-change (same).

**Forward Segments During Test.** In order to examine scanning patterns when the speaker was looking at the infant and naming the object during test, a 2 x 3 ANOVA with trial (2: Switch, Same) and AOI (3: Object, Eyes, Mouth) as repeated factors was conducted. This analysis resulted in significant main effects of trial ( $F(1,19)=11.92, p<.001, \eta^2<.01, \eta_p^2=.39$ ) and AOI ( $F(2,38)=115.47, p<.001, \eta^2=.83, \eta_p^2=.86$ ) and a significant trial x AOI interaction ( $F(2,38)=11.18, p=.01, \eta^2=.01, \eta_p^2=.37$ ). With regard to the two-way interaction, percent fixation duration to the Mouth was significantly larger during Switch ( $M=.76, SD=.22$ ) compared to Same ( $M=.62, SD=.25$ ) test trials ( $t(19)=3.67, p=.002, d=.58$ ). However, percent fixation durations to the Object during Switch ( $M=.06, SD=.10$ ) and during Same ( $M=.06, SD=.11$ ) were not significantly different ( $t(19)=.08, p=.94$ ). Percent fixation durations to the Eyes during Switch ( $M=.01, SD=.02$ ) and during Same ( $M=.03, SD=.06$ ) were not significantly different ( $t(19)=1.80, p=.09$ ). Within the switch trial, percent fixation duration to the Mouth was significantly greater than to the Object ( $t(19)=11.22, p<.001, d=3.96$ ) and to the Eyes

---

<sup>3</sup> The results of HB segment analyses were the same as those found in Familiarization (reported above) so are not reported for Test.

( $t(19)=14.95, p<.001, d=4.71$ ). Percent fixation duration was not significantly different between the Object and the Eyes ( $t(19)=2.01, p=.06$ ). Likewise, within the same trial, percent fixation duration to the Mouth was significantly greater than to the Object ( $t(19)=7.99, p<.001, d=2.88$ ) and to the Eyes ( $t(19)=9.45, p<.001, d=3.22$ ). Percent fixation duration was not significantly different between the Object and the Eyes ( $t(19)=.95, p=.35$ ; see Figure 2).

In order to examine scanning patterns reflecting infants' perception of a change in trial configuration from the end of familiarization to test, a 2 x 3 ANOVA with trial (2: F4, Switch) and AOI (3: Object, Eyes, Mouth) as repeated factors was conducted. This analysis resulted in a significant main effect of trial ( $F(1,19)=4.76, p=.04, \eta^2<.01, \eta_p^2=.20$ ) and AOI ( $F(2,38)=146.56, p<.001, \eta^2=.83, \eta_p^2=.89$ ) and a significant trial x AOI interaction ( $F(2,38)=3.94, p=.03, \eta^2=.01, \eta_p^2=.17$ ). With regard to the two-way interaction, percent fixation duration to the Mouth was significantly larger during Switch ( $M=.76, SD=.22$ ) compared to F4 ( $M=.64, SD=.27; t(19)=2.33, p=.03, d=.49$ ). However, percent fixation duration to the Object during Switch ( $M=.06, SD=.10$ ) and during F4 ( $M=.09, SD=.12$ ) were not significantly different ( $t(19)=.81, p=.43$ ). Also, percent fixation to the Eyes during Switch ( $M=.01, SD=.02$ ) and during F4 ( $M=.02, SD=.03$ ) were not significantly different ( $t(19)=1.06, p=.30$ ; see Figure 3).

In order to examine scanning patterns reflecting infants' perception of no-change in trial configuration from the end of familiarization to test, a 2 x 3 ANOVA with trial (2: F4, Same) and AOI (3: Object, Eyes, Mouth) as repeated factors was conducted. This analysis resulted in a significant main effect of AOI only ( $F(2,38)=92.02, p<.001, \eta^2=.78, \eta_p^2=.83$ ). Paired samples *t*-tests indicated that the percent fixation duration to the Mouth ( $M=.63, SD=.24$ ) was significantly larger than to the Object ( $M=.07, SD=.09; t(19)=9.19, p<.001, d=3.06$ ) and to the Eyes ( $M=.02,$

$SD=.04$ ;  $t(19)=10.73$ ,  $p<.001$ ,  $d=3.54$ ). Percent fixation duration to the Object was also significantly larger than to the Eyes ( $t(19)=2.02$ ,  $p=.04$ ,  $d=.71$ ).

**Turned Segments During Test.** Paired samples  $t$ -tests compared percent fixation duration, as only the Object AOI was analyzed. A paired samples  $t$ -test compared percent fixation duration to the Object between Switch and Same test trials. This resulted in a significant difference between trials ( $t(19)=2.08$ ,  $p=.05$ ,  $d=.35$ ), with larger percent fixation duration to the Object during Switch ( $M=.27$ ,  $SD=.18$ ) compared to during Same ( $M=.21$ ,  $SD=.19$ ).

A paired samples  $t$ -test compared percent fixation to the object between the last familiarization trial (F4) and test trials. Percent fixation duration to the Turned Object was not significantly different between F4 ( $M=.29$ ,  $SD=.17$ ) and Switch ( $M=.27$ ,  $SD=.18$ ;  $t(19)=.58$ ,  $p=.57$ ) but was significantly larger compared to Same ( $M=.21$ ,  $SD=.19$ ;  $t(19)=2.75$ ,  $p=.01$ ,  $d=.46$ ).

### **Individual Infant Looking Profiles**

In order to more carefully compare individual infants' looking patterns across test trials, difference scores were calculated by subtracting infants' percent fixation duration to the Forward Mouth and the Turned Object during same trial from those occurring during the switch trial. Thus, positive scores indicate greater attention to these AOIs on the Switch trial. Based on these different scores, infants' performances were assigned one of four looking pattern profiles: More Mouth<sup>4</sup>/More Object, More Mouth/Less Object, Less Mouth/More Object, and Less Mouth/Less Object. The distribution of infants across these profiles was significantly different from chance ( $\chi^2=8.8$ ,  $p=.03$ ), with the greatest number of infants in the More Mouth/More Object profile ( $n=10$ ) compared to More Mouth/Less Object ( $n=6$ ), Less Mouth/More Object ( $n=2$ ), and Less

---

<sup>4</sup> Positive classification was assigned for values greater than zero. Difference scores of zero were assigned a negative classification.

Mouth/Less Object ( $n=2$ ). Collapsing across object patterns, 16/20 infants showed more attention to the mouth of the speaker during the Switch trial.

### **Looking Pattern, Language Skill, and Temperament Correlations**

In addition to assessing looking patterns during familiarization and test, it was also of interest to examine relations between looking patterns on test trials and measures of infants' language skill and temperament (see Table 1 for descriptive statistics). To this aim, bivariate correlations were performed. As mentioned earlier, infants experienced a static, silent face prior to familiarization and only advanced from this image by regarding the woman's eyes. Latency to attend to the eyes of the woman's still face was negatively correlated with IBQ NEG scores ( $r=-0.53, p=.02$ ).

No significant correlations were found between percent fixation duration to AOIs during the switch test trial and any measures of language skill or temperament (see Table 2). However, there was a positive correlation between percent fixation duration on the Object (Turned segments) during Same trials and infants' receptive vocabulary ( $r=.45, p=.05$ ). As expected, all metrics of language skill were significantly correlated with one another. Receptive vocabulary was positively correlated with expressive vocabulary ( $r=.91, p<.001$ ) and the ASQ-C ( $r=.55, p=.01$ ). The ASQ-C and ASQ-PS were positively correlated ( $r=.49, p=.03$ ).

### **Discussion for Experiment 1**

The primary purpose of Experiment 1 was to examine whether ostensive cues could augment 14-month-old infants' associative learning of minimal pair object labels when eye gaze-object regard contingency is maintained. It was hypothesized that infants would attend differentially on Switch trials in spite of their familiarization with minimal pair labels because

the attention recruited by eye contact with an infant-directed dynamic speaker would facilitate infants' processing of object-label presentations, and learning of the word-object associations. Results clearly showed that infants attended more to the Forward Mouth and Turned Object (in terms of both percent fixation duration) when the object label was switched as compared to when the Object was labeled correctly. Moreover, this pattern was seen in a significantly high number of individual infants in addition to group level comparisons. Results also showed that attention to the Mouth dominated infants' looking patterns, with infants' attention to the Object and to the Eyes being relatively low and equal.

Although there are studies demonstrating 14-month-olds' successful mapping of minimal pair labels onto objects (Werker & Fennell, 2004; Fennell & Waxman, 2010; Fennell & Byers-Heinlein, 2014), those studies all utilized variations in linguistic cues. The current experiment, however, provides evidence that ostensive cues from a simulated social partner provide yet another way that information processing can be facilitated under cognitive challenge, by recruiting increased attention to the task through the social domain. What remains unclear is whether ostensive cues are sufficient to offset the cognitive load of the task if the contingency between a speaker looking to the object and then providing its label is compromised. That is, in natural referential learning contexts, infants use ostensive cues to direct their attention to relevant agents and actions, but also have learned to build expectations that the agent's behavior (e.g., head-turn + eye gaze) will be directed to the source that is being labeled. As such, ostensive cues alone should not be sufficient for infants to learn difficult object-label mappings if the agent-object contingency is unavailable. Experiment 2 was designed to address this question.

## Experiment 2

The purpose of Experiment 2 was to investigate whether ostensive cues (i.e., mutual gaze and IDS) are sufficient for facilitating infants' learning of minimal pair labels for objects, even when eye gaze-object regard contingency is unavailable. If a speaker does not visually regard the object to be labeled (e.g., by looking to the wrong side of the screen), can infants still learn the object-label relation? To this aim, in this experiment, the target of the female speaker's shifted gaze was incongruent with the target object's location. That is, if the object appeared on the left, she looked to the empty right side of the screen; if the object appeared on the right, she looked to the empty left side of the screen. Based on the results of Experiment 1, it was predicted that if infants were able to learn the association between objects and labels in spite of the violation of object-eye gaze contingency, they should increase their attention to the mouth of the speaker during the switch v. same test trial. However, if eye gaze-object regard contingency is an essential feature of referential learning, infants' looking patterns would not differ between switch and same test trials.

### Participants

The participants in Experiment 2 were 13- to 15-month-olds ( $n = 18$ ; mean age = 13.94 mo;  $SD = 0.73$ ; 8 female). Data collected from an additional 5 infants could not be used due to an inability to calibrate ( $n = 1$ ) or poor eye tracking data (i.e.,  $n = 4$ ). All participants were recruited from (and representative of) the surrounding geographic areas and were recruited with the same methods as in Experiment 1. Accordingly, the infants in this sample were learning English as their primary language<sup>5</sup> and a majority of infants were Caucasian (13/18 infants). Infants were at least 37 weeks gestational age at birth, and all of the infants came from two-

---

<sup>5</sup> Mandarin Chinese was the first language in the household for one infant in this sample (parents estimated 80% Mandarin Chinese, 20% English). Analyses without this subject in the sample were the same as the analyses with her in the sample, so this subject was retained in the dataset.

parent households. Of the 36 parents, 35 had at least some college experience, and 34 parents had obtained at least a Bachelor's degree. Fourteen of the 18 households reported annual incomes of \$50,000 or more.

## **Results for Experiment 2**

### **Looking Patterns During Familiarization**

In order to examine order effects, a 2 x 4 ANOVA with object+label pair (2: pair 1, pair 2) as a repeated factor and which word from the minimal pair was heard first (first word; 4: Bin, Din, Gem, Pem) as a between-subjects factor on percent fixation duration to the Object in Forward segments was conducted. This analysis resulted in significant main effect of pair ( $F(1,3)=5.31, p=.03, \eta^2=.23, \eta_p^2=.28$ ), but not a significant interaction of pair and first word ( $F(3,14)=1.45, p=.27$ ). Accordingly, the following analyses are collapsed across the first word.

Next, changes in infants' attention to the Object, Eyes, Mouth<sup>6</sup> of the speaker were analyzed as they processed the audiovisual presentations during familiarization. To address this issue, repeated measures ANOVAs compared each infant's percent fixation duration as a function of AOI (3: Object, Eyes, Mouth) averaged across the four familiarization trials for the Hi Baby (HB) and Forward segments. An additional repeated measures ANOVA compare infants' percent fixation duration as a function of AOI (3: Object, Eyes, Mouth) between the first and last familiarization trials for Forward segments. For Turned segments, paired samples *t*-tests compared percent fixation duration to the Object AOI only between Forward and Turned segments as well as between the first and last familiarization trials.

---

<sup>6</sup> Non-object AOIs were not included in these analyses as no fixations were registered during any trial within this experiment.

### **HB Segments During Familiarization.**

In order to examine scanning patterns when the speaker greeted the infant with “Hi baby!”, an ANOVA with AOI (3: Object, Eyes, Mouth) as the repeated factor was conducted. This analysis resulted in a significant main effect ( $F(2,38)=163.00, p<.001, \eta^2=.91, \eta_p^2=.91$ ). Percent fixation duration to the Mouth ( $M=.70, SD=.19$ ) was significantly larger than to the Object ( $M=.04, SD=.07; t(19)=12.44, p<.001, d=4.68$ ) and to the Eyes ( $M=.02, SD=.04; t(19)=13.98, p<.001, d=5.02$ ). Percent fixation duration to the Object was significantly greater than to the Eyes ( $t(19)=2.25, p=.04, d=.40$ ).

### **Forward Segments During Familiarization.**

In order to examine scanning patterns when the speaker was looking at the infant and naming the object during familiarization, an ANOVA with AOI (3: Object, Eyes, Mouth) as the repeated factor was conducted. This analysis resulted in a significant main effect ( $F(2,38)=113.05, p<.001, \eta^2=.87, \eta_p^2=.87$ ). Paired-sample  $t$ -tests indicated that the percent fixation duration to the Mouth ( $M=.66, SD=.21$ ) was significantly larger than to the Object ( $M=.04, SD=.04; t(17)=11.26, p<.001, d=4.00$ ) and to the Eyes ( $M=.04, SD=.05; t(17)=10.34, p<.001, d=4.07$ ). Percent fixation duration was not significantly different between the Object and the Eyes ( $t(19)=.00, p=1.00$ ).

In order to compare scanning patterns occurring at the start and end of familiarization, an ANOVA with trial (2: F1, F4) and AOI (3: Object, Eyes, Mouth) as the repeated factor was conducted. This analysis resulted in significant main effects of trial ( $F(1,17)=5.52, p=.03, \eta^2<.01, \eta_p^2=.25$ ) and AOI ( $F(2,34)=107.65, p<.001, \eta^2=.82, \eta_p^2=.87$ ) and a significant trial x AOI interaction ( $F(2,34)=5.20, p=.01, \eta^2=.01, \eta_p^2=.23$ ). Paired sample  $t$ -tests indicated that the percent fixation duration to the Mouth was significantly larger in F1 ( $M=.71, SD=.22$ ) as

compared to F4 ( $M=.60$ ,  $SD=.25$ ;  $t(17)=2.28$ ,  $p=.04$ ,  $d=.46$ ). In contrast, percent fixation duration to the Object was significantly larger in F4 ( $M=.06$ ,  $SD=.08$ ) as compared to F1 ( $M=.01$ ,  $SD=.03$ ;  $t(17)=2.63$ ,  $p=.02$ ,  $d=.83$ ). Between the two trials, there were no significant differences between percent fixation to the Eyes ( $t(19)=1.90$ ,  $p=.08$ ). Within F1, percent fixation duration to the Mouth was significantly greater than to the Object ( $t(17)=12.24$ ,  $p<.001$ ,  $d=4.42$ ) and to the Eyes ( $M=.06$ ,  $SD=.09$ ;  $t(17)=9.33$ ,  $p<.001$ ,  $d=3.83$ ). Percent fixation duration to the Object was also significantly greater than to the Eyes ( $t(17)=2.64$ ,  $p=.02$ ,  $d=.74$ ). Within F4, percent fixation duration to the Mouth was significantly greater than to the Object ( $t(17)=7.94$ ,  $p<.001$ ,  $d=2.93$ ) and to the Eyes ( $t(17)=8.61$ ,  $p<.001$ ,  $d=2.31$ ). Percent fixation duration was not significantly different between the Object and the Eyes ( $t(17)=1.88$ ,  $p=.08$ ).

#### **Turned Segments During Familiarization.**

In order to examine scanning patterns when the speaker was looking at the object during familiarization, a paired samples  $t$ -test compared percent fixation duration to the object across Forward and Turned segments. Percent fixation duration to the Turned Object ( $M=.09$ ,  $SD=.05$ ) was significantly greater than to the Forward Object ( $M=.04$ ,  $SD=.04$ ;  $t(17)=3.62$ ,  $p=.002$ ,  $d=1.14$ ). As in Experiment 1, this indicated that infants were more likely to look at the object after the female speaker had turned her eye-gaze toward it.

A paired samples  $t$ -test compared percent fixation to the object between the first and last familiarization trials. Percent fixation duration to the Turned Object was significantly greater during the last familiarization trial ( $M=.14$ ,  $SD=.12$ ) than to the first familiarization trial ( $M=.04$ ,  $SD=.04$ ;  $t(17)=3.33$ ,  $p=.003$ ,  $d=1.05$ ).

## Looking Patterns During Test<sup>7</sup>

Of primary interest was whether infants' looking patterns during the Switch test trial were different from their looking patterns during the Same test trial. To determine this, a repeated measures ANOVA compared each infant's percent fixation durations as a function of trial type (2: Switch, Same) and AOI (3: Object, Eyes, Mouth) for Forward segments. For Turned segments, a repeated measures ANOVA compared percent fixation durations as a function of the Object AOI only. Additionally, repeated measures ANOVAs were conducted to compare looking patterns on the last familiarization trial and each of the test trials, as a way of approximating analyses conducted for previous switch studies.

### Forward Segments During Test.

In order to examine scanning patterns when the speaker was looking at the infant and naming the object during test, a 2 x 3 ANOVA with trial (2: Switch, Same) and AOI (3: Object, Eyes, Mouth) as the repeated factor resulted in a significant main effect of AOI only ( $F(2,32)=57.15, p<.001, \eta^2=.73, \eta_p^2=.77$ ). Percent fixation duration to the Mouth ( $M=0.61, SD=.27$ ) was significantly larger than to the Object ( $M=.07, SD=.08; t(17)=6.81, p<.001, d=2.68$ ) and percent fixation duration to the Eyes ( $M=.02, SD=.03; t(17)=8.72, p<.001, d=3.05$ ). Percent fixation duration to the Object was significantly greater than percent fixation duration to the Eyes ( $t(17)=2.69, p=.02, d=.87$ ). Within the Switch test trial, percent fixation duration to the Mouth was significantly greater than that to the Object ( $t(17)=6.09, p<.001, d=2.48$ ) and that to the Eyes ( $t(17)=8.54, p<.001, d=2.96$ ). Percent fixation duration to the Object was also significantly greater than that to the Eyes ( $t(17)=2.56, p=.02, d=.84$ ). Within the Same test trial, percent fixation duration to the Mouth was significantly greater than that to the Object

---

<sup>7</sup> The results of HB segment analyses were the same as those found in Familiarization to are not reported for Test.

( $t(17)=6.40, p<.001, d=2.40$ ) and that to the Eyes ( $t(17)=7.66, p<.001, d=2.66$ ). Percent fixation duration was not significantly different between the Object and the Eyes ( $t(17)=1.64, p=.12$ , see Figure 4).

In order to examine scanning patterns reflecting infants' perception of a change in trial configuration from the end of familiarization to test, a 2 x 3 ANOVA with trial (2: F4, Switch) and AOI (3: Object, Eyes, Mouth) as the repeated factor was conducted. This analysis resulted in a significant main effect of AOI only ( $F(2,34)=64.31, p<.001, \eta^2=.76, \eta_p^2=.79$ ). Percent fixation duration to the Mouth ( $M=.61, SD=.25$ ) was significantly greater than that to the Object ( $M=.07, SD=.07; t(17)=2.89, p=.01, d=2.89$ ) and that to the Eyes ( $M=.02, SD=.04; t(17)=8.97, p<.001, d=3.25$ ). Percent fixation duration to the Object was significantly greater than that to the Eyes ( $t(17)=3.27, p=.01, d=.87$ ).

In order to examine scanning patterns reflecting infants' perception of no-change in trial configuration from the end of familiarization to test, a 2 x 3 ANOVA with trial (2: Last Familiarization, Same) and AOI (3: Object, Eyes, Mouth) as the repeated factor was conducted. This analysis resulted in a significant main effect of AOI only ( $F(2,34)=63.84, p<.001, \eta^2=.76, \eta_p^2=.79$ ). Percent fixation duration to the Mouth ( $M=.60, SD=.26$ ) was significantly greater than that to the Object ( $M=.06, SD=.07; t(17)=7.63, p<.001, d=2.82$ ) and that to the Eyes ( $M=.03, SD=.04; t(17)=8.60, p<.001, d=3.09$ ). Percent fixation duration to the Object was significantly greater than that to the Eyes ( $t(17)=2.13, p=.05, d=.63$ ).

### **Turned Segments During Test.**

Paired samples  $t$ -tests compared percent fixation duration, as only the Object AOI was analyzed. A paired samples  $t$ -test compared percent fixation durations to the Object between the Switch and Same test trials. This resulted in a significant difference between the trials ( $t(17)=.08$ ,

$p=.94$ ), with greater percent fixation duration to the Object in the Switch test trial ( $M=.19$ ,  $SD=.11$ ) than the percent fixation duration in the Same test trial ( $M=.18$ ,  $SD=.17$ ).

A paired samples  $t$ -test compared percent fixation to the object between the first and last familiarization trials. Percent fixation duration to the Turned Object during the last familiarization trial ( $M=.14$ ,  $SD=.12$ ) was not significantly different from that on the Switch ( $M=.19$ ,  $SD=.11$ ;  $t(17)=-1.66$ ,  $p=.12$ ) or Same ( $M=.18$ ,  $SD=.17$ ;  $t(17)=-.85$ ,  $p=.41$ ) test trials.

### **Individual Infant Looking Profiles**

As in Experiment 1, individual difference scores were calculated by subtracting infants' percent fixation duration to the Forward Mouth and the Turned Object during the same trial from those occurring during the switch trial. Thus, positive scores indicate greater attention to these AOIs on the Switch trial. Based on these difference scores, infants' performances were assigned one of four looking pattern profiles: More Mouth<sup>8</sup>/More Object, More Mouth/Less Object, Less Mouth/More Object, and Less Mouth/Less Object. Unlike this analysis in Experiment 1, the distribution of infants across these profiles was not significantly different from chance ( $\chi^2=1.1$ ,  $p=.77$ ), with relatively equal numbers of infants across the profiles: More Mouth/More Object ( $n=5$ ), More Mouth/Less Object ( $n=6$ ), Less Mouth/More Object ( $n=4$ ), and Less Mouth/Less Object ( $n=3$ ).

### **Looking Pattern, Language Skill, and Temperament Correlations**

In addition to assessing looking patterns during familiarization and test, it was also of interest to examine relations between looking patterns on test trials and measures of infants' language skill and temperament (see Table 1 for descriptive statistics). To this aim, bivariate

---

<sup>8</sup> Positive classification was assigned for values greater than zero. Difference scores of zero were assigned a negative classification.

correlations were performed. However, there were no significant correlations between looking patterns on the test trials and metrics of language skill or temperament (see Table 3).

### **Discussion for Experiment 2**

The primary purpose of Experiment 2 was to examine whether ostensive cues (i.e., mutual gaze and IDS) are sufficient for facilitating infants' learning of minimal pair labels for objects, even when eye gaze-object regard contingency is broken. That is, when eye gaze-object regard contingency is absent, can infants still learn in a referential context with the cognitive load is high? It was hypothesized that if eye gaze-object regard contingency is a necessary component of associative learning, infants would not show different looking patterns between Switch and Same test trials in Experiment 2.

Results showed that attention to the Mouth dominated infants' looking patterns, with infants' visual attention to the Object and to the Eyes being relatively low or equal. Most critically, in support of my hypothesis, there were no differences in looking patterns between the trial types on Forward or Turned segments. As differential attention to the Forward Mouth and Turned Object served as evidence of word learning in Experiment 1 and the number of individual infants demonstrating this pattern in Experiment 2 was not different from chance, these replicate previous demonstrations of 14-month-olds' difficulty with minimal pair labels in referential learning context. Accordingly, Experiment 2 provides evidence that although ostensive cues from a simulated social partner were available to recruit attentional resources, violation of the eye gaze-object regard contingency prevents successful mapping of minimal pair labels onto objects.

## General Discussion

The purpose of the experiments comprising this study was to investigate the potential impact of ostensive cues (i.e., mutual gaze and infant-directed speech) and eye gaze-object regard contingency from a simulated social partner on 14-month-old infants' object-label learning under increased cognitive load. Specifically, Experiment 1 examined whether mutual gaze and infant directed register would improve 14-month-old infants' associative learning of minimal pair object labels when the speaker's eye gaze-object regard contingency was maintained. The results of Experiment 1 showed that infants looked significantly more at the mouth of the speaker as well as the object during switch (i.e., when the object and label from familiarization had been transposed) compared to same test trials. No significant differences in looking were seen to the speaker's eyes during these trials. This helps to explain previous failures to find successful associative learning in 14-month-olds using minimal pairs in that the infants were presented with only objects and labels in the absence of any supportive ostensive cues (e.g., Stager & Werker, 1997). The positive findings from Experiment 1 are consistent with results from Frank, Vul, and Saxe (2012) and partially corroborate Lewkowicz and Hansen-Tift (2012) who showed English-learning 12-month-olds videos of a woman speaking either English or Spanish in an infant-directed manner. When comparing looking patterns, Lewkowicz and Hansen-Tift found that the infants presented with English attended equally to the eyes and mouth of the speaker, whereas the infants presented with Spanish attended more to the speaker's mouth than to her eyes (see also Pons, Bosch, & Lewkowicz, 2015). Because these infants were monolingual English-learning, processing of audiovisual Spanish generated higher cognitive demands than those required when processing English. Under such a load, attention was allocated to the mouth to facilitate audiovisual processing of the speech. Similarly, the results

from Experiment 1 show increased attention to the mouth of the speaker when she directs her gaze to a familiar object but then looks back at the infant and mislabels the object.

Infants' increased attention to the mouth and object during the Switch trial in Experiment 1 indicated that transposing the object label was detected and infants' expectations for the label associated with that object were violated. When infants' expectations for the word-object association were violated, they looked to features of the presentation that helped them disambiguate or understand the change, namely the speaker's mouth when she was looking forward and labeling the object and the object itself when the speaker regarded it. Such increased reliance on areas of information in dynamic speakers that help processing was seen in 12-month-old bilingual infants (Pons, Bosch, & Lewkowicz, 2015). These infants appeared to increase their regard of a speaker's mouth while listening to non-native speech compared to a cohort of monolingual infants, presumably because the bilingual infants have learned to rely more heavily on audiovisual speech information.

In addition to the increase in looking to mouth and object during test, infants in Experiment 1 also showed an overall bias toward the mouth region of the speaker throughout familiarization, although attention to the mouth slightly decreased over familiarization trials whereas attention to the objects slightly increased. This attention preference was seen in a recent study by Heck, Panneton, and Mills-Smith (2015) in which 12-month-olds showed significantly more scanning of the mouth of a dynamic speaker regardless of three different kinds of emotion expressions. Likewise, Tenenbaum, Shah, Sobel, Malle and Morgan (2013) found that 12-month-olds allocated significantly more attention to a speaker's mouth when she was speaking compared to smiling. Thus, it is likely that the high amount of attention to the speaker's mouth

per se is responsible for more effective mapping of labels onto objects, particularly when the labels are highly similar.

Given that the ostensive nature of information in Experiment 1 facilitated infants' associative learning of difficult word-object pairs, Experiment 2 further examined the importance of the speaker's eye gaze-object contingency for infants' associative learning. Is heightened attention to the co-occurrence of objects and labels via speaker characteristics sufficient to support good learning even if the goal-directed nature of the speaker's actions is no longer clear? Both the emergentist coalition model (ECM: Hollich et al, 2000) and the resource limitation hypothesis (RLH: Werker, Fennell, Corcoran, & Stager, 2002) discussed earlier support the view that additional attentional resources are required when the cognitive load of word learning is high. The differentiation of domains of attention in the ECM calls to question whether recruiting attention from one domain (e.g., social) can be helpful when the resources of another domain (e.g. linguistic) are challenged. Further, infants clearly benefit from detecting and utilizing environmental contingencies, so Experiment 2 explored the need for a specific agent-action contingency (eye gaze-object regard) for successful referential learning. In order to examine the importance of the speaker's contingent gaze to the object, Experiment 2 violated this goal-directed action by having the speaker turn away from (in the opposite direction of) the object she then labeled. Otherwise, infants' attention was elicited in the same way as in Experiment 1.

In contrast to Experiment 1, the results of Experiment 2 showed that infants did not look significantly more at the mouth of the speaker or the object during switch (i.e., when the object and label from familiarization had been transposed) compared to same test trials. Thus, the only clear evidence for referential learning appears to be from Experiment 1. Much like the individual difference approach taken by Fais et al (2012), analyses of individual profiles of

infants explored the patterns of looking more to the mouth and more to the object during switch compared to same trials. Chi-square analyses examining the distribution of infants across four possible profiles showed a significant number of infants with More Mouth/More Object looking (10/20) in Experiment 1 compared to Experiment 2 (5/18).

Of note is the remarkably low amount of fixation to the eyes of the speaker when she was turned forward, looking at the infant, and naming the object on the screen (both during familiarization and test). Given the presumed importance of eye gaze for initiating and maintaining infants' attention in interactive contexts, it was expected that infants would have referenced the eyes of the speaker more frequently than they did in either experiment. Importantly, it should be acknowledged that both the object and the facial features of the speaker were available for continuous viewing in both experiments if peripheral vision is taken into account while the speaker was facing forward. The analyses of eye-tracking patterns is based on foveal fixations that occur within targeted areas of interest but does not capture what visual information is available in the periphery. Thus, it is possible that the infants in both experiments were aware of the speaker's eye gaze toward them, but focused on the mouth more because of the difficulty of the object-label relation. It will be necessary to conduct future studies examining these patterns under a variety of contexts, including referential learning of *easy* word-object relations (e.g., boog v. neem).

Although infants in the current experiments showed little focused attention on the eyes of the speaker, the entire labeling sequence was only initiated after infants specifically directed their gaze to her eyes (i.e., before familiarization began). Evidence has been found that eye gaze, in the absence of any other ostensive cue, directs the attention of 18-month-old infants but not infants of younger ages (Moore & Corkum, 1998). Further, studies examining cortical responses

to ostensive cues indicate that directed gaze from a social partner and infant-directed speech both activate networks of brain regions associated with attention (Parise & Csibra, 2013), but *both* are required for activating regions of cortex associated with processing audiovisual communication from a social partner (Lloyd-Fox, Széplaki-Köllöd, Yin, & Csibra, 2015). In the current experiments, unlike most previous Switch experiments, audiovisual disambiguation of the minimal pair labels was possible because of the availability of the dynamic speaker and the mutual gaze and infant-directed speech she provided. Based on the findings of Parise and Csibra (2013) and Lloyd-Fox et al (2015), infants' attention to the audiovisual presentations should have been maximally primed in both experiments based on identical availability of mutual gaze and infant-directed speech. Accordingly, the overall lack of evidence for learning minimal pair object labels in Experiment 2 is most likely due to the violation of eye gaze-object regard contingency.

In Experiment 1 (and the real world), eye gaze-object regard contingency is available such that speakers' ostensive cues and the labels they provide correspond specifically with the target of their attention. Breaking this contingency is not what is expected based on real-world experience and does not promote the necessary information processing for cognitively challenging word learning. Studies have explored manipulations of similar contingencies generally, and also in the context of word learning. For example, in a study by Beier and Spelke (2012), 9- and 10-month-old infants' viewed a man and a woman who greeted one another either while looking toward one another or with their gaze averted from one another. Ten-month-olds, but not 9-month-olds were sensitive to the difference between gaze in that they looked significantly longer at the couple when their gaze was averted (something unexpected). This finding indicates that when the behavior of social partners' does not correspond to their intended

goal (in this case, greeting one another), by late in the first year, infants' will recognize the violation of contingency between intended action and actual behavior.

In a study by Woodward (2003), contingency between eye gaze and object regard was specifically investigated. Infants were habituated to a person with two potential objects to regard. During habituation, the person always looked to the same object (i.e., always disregarded the other), establishing an eye gaze-object regard contingency. Twelve-month-olds paid increased attention when the person looked to the previously disregarded object, violating the expected contingency (Woodward, 2003). Similarly, Moore and Povinelli (2007) conducted a study with 12-month-olds, in which, during a period of familiarization, an adult's head turn to a toy was followed by the activation of that toy. One of two possible testing conditions was administered to each infant. During the Head Turn condition, the toys were removed, but the adult's head turns continued. During the Toy condition, toys were activated despite the adult keeping their face and gaze forward. In the Head Turn condition, 12-month-old infants stopped following the shifting gaze of the adult. That is, when there was no longer an eye gaze-object contingency, 12-month-olds disregarded the cues of the adult. During the Toy condition, 12-month-old infants continued looking to activated toys even though the adult was no longer attending (Moore & Vocinelli, 2007). Taken together, these results indicate that exposure to eye gaze-object regard contingency promotes infants' continued attention to previously regarded targets. In a slightly older sample of children participating in a word learning task (i.e., 24-month-olds), Baron-Cohen, Baldwin, and Crowson (1997) showed that a break in eye gaze-object contingency prevented toddlers from learning object labels. The experimenter either shared gaze on a target object with the toddlers while providing the label *or* averted their gaze from the target object but continued to provide the label (Baron-Cohen, Baldwin, & Crowson

(1997). Toddlers only learned the object label when they shared gaze with the adult on the object.

Overall, the current study replicates Baron-Cohen, Baldwin, and Crowson's (1997) finding that that eye gaze-object regard contingency is necessary for successful mapping of labels onto objects. That is, only infants in Experiment 1 were able to learn the object-label pairs whereas those in Experiment 2 showed less robust evidence of learning. In addition, infants in Experiment 2 attended to objects in a manner similar to those in Moore and Povinelli's Toy condition. That is, despite the break in eye gaze-object regard contingency, infants still attended to the object. Interestingly, despite this object regard, the majority of infants in Experiment 2 did not recognize the Switch based on increased attention to the speaker's mouth. As the violation of eye gaze-object regard contingency did not disrupt attention to the referent, the cognitive load of the task falls to the label. Alternatively, the cognitive load of the task shifts from the labels to the incongruence between goal and behavior. Ostensive cues abated the load generated by the label for a majority of infants in Experiment 1 but not Experiment 2. Accordingly, eye gaze-object regard contingency and the information it provides about the goals of the partner contribute to the ostensive message. As previously described, ostensive cues help social partners understand when information is meant for them. When eye gaze-object regard contingency is compromised, the goal of the ostensive cue provider and their intention to share information becomes ambiguous. In this way, infants in Experiment 2, although primed by the ostensive cues in the same manner as infants in Experiment 1, may not have received the same message about the woman's intention to label the object and/or allocated more attention toward figuring out the goal.

In comparison to Fais et al. (2012), infants in Experiment 1 successfully mapped minimal

pair labels on to objects, despite the pre-recorded nature of the presentations. Fais et al. only found minimal pair success in their live experiment, and only for those infants who exhibited a high degree of mutual eye regard with the experimenter. Importantly, infants in both Experiments 1 and 2 of the current study were required to engage in mutual gaze (with the still image of the woman) before familiarization began. Forcing mutual gaze, however, did not seem to be sufficient in maintaining referential learning across the manipulations. In the study by Fais et al., the experimenter's infant-directed speech was limited to her production of the word labels. In the current study, however, infant-directed speech occurred at the start of each trial with the "Hi, baby" greeting, and eye regard during these segments was low. Increased availability of infant-directed speech, in addition to infant-directed gaze, may have activated the ostensive attention network sufficiently to remove the need to make specific eye contact (fixations) while supporting infants' mapping of minimal pair labels onto their object referents.

To address this issue, a future study should replicate the methods of Experiment 1 but without the still image prior to familiarization. If the "Hi, baby" greeting was also removed, the study would replicate the recorded experiment by Fais et al. (2012). Alternatively, Experiment 2 could be replicated with a live experimenter with a protocol somewhat like that from Fais et al. (2012). Such an experiment would identify the importance of eye gaze-object regard contingency in a more naturalistic context, as a live social partner would no longer visually regard the referent of the label provided.

The question remains, if the woman never looked at the object in Experiment 2, why did infants look at her? Infants may have continued to look at her in an attempt to sort out the difference between the target indicated by her ostensive cues and the available target object. In contrast, although she never visually regarded the object, she was consistent to do so. That is,

the probability that the woman in Experiment 2 would look at the object was zero but was 100% in Experiment 1. In both experiments, although the target of the turned gaze differed, the woman's movements and regard were completely predictable. Future studies could manipulate the predictability of her gaze target to determine how consistent a social partner needs to be in their eye gaze-object regard contingency before the cognitive load of the word learning task overpowers the benefits of the ostensive cues.

### **Limitations and Conclusions**

In comparison to other switch studies, the trials in this study were not infant-controlled. That is, the duration of any given trial was fixed and not a function of an infant's attention to the screen. Particularly for Experiment 2, infants learning of the minimal pair labels may have benefitted from a longer period of familiarization. Pilot work for Experiment 1 attempted to employ a greater number of familiarization trials but this resulted in increased infant attrition (i.e., more infants failing to complete the entire procedure). If an infant-controlled habituation paradigm had been utilized, infants could have processed information at their own pace, with all infants achieving similar levels of habituation. In addition, differences between the number of trials required to reach habituation could also be used as a metric to quantify cognitive load differences between the two experiments.

Nonetheless, the results of this study highlight the benefits of ostensive cue utilization on behalf of infants during word+object learning, especially when the referential context is challenging. However, the power of ostensive information to activate and maintain infants' attention is not necessarily sufficient for successful associative learning. Additionally, the intention of a social partner needs to be transparent and consistent with the information that the partner is trying to convey to the recipient. Although cues from linguistic and social domains are

necessary for typical word learning, contingency is also critical and may serve as a component of ostension.

## References

- Baldwin, D. A. (1995). Understanding the link between joint attention and language. In C. Moore and P. J. Dunham (Eds.), *Joint Attention: Its origins and role in development* (pp. 131-158). Hillsdale, NJ: Erlbaum.
- Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M. and Tidball, G. (1996). Infants' Reliance on a Social Criterion for Establishing Word-Object Relations. *Child Development, 67*, 3135–3153. doi: 10.1111/j.1467-8624.1996.tb01906.x
- Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with Autism use the speaker's direction of gaze strategy to crack the code of language? *Child Development, 68*(1), 48-57.
- Beier, J. S., & Spelke, E. S. (2012). Infants' developing understanding of social gaze. *Child Development, 83*(2), 486-496. doi: 10.1111/j.1467-8624-2011.01702.x
- Bhat, A.N., Galloway, J.C. and Landa, R.J. (2010). Social and non-social visual attention patterns and associative learning in infants at risk for autism. *Journal of Child Psychology and Psychiatry, 51*, 989–997. doi: 10.1111/j.1469-7610.2010.02262.x
- Bricker, D., & Squires, J. (1999). Ages and Stages Questionnaires: A parent-completed child-monitoring system. Baltimore, MD: Paul H. Brookes Publishing Co.
- Csibra, G. (2010). Recognizing communicative intentions in Infancy. *Mind & Language, 25*, 141–168. doi: 10.1111/j.1468-0017.2009.01384.x
- Curtin, S. (2009). Twelve-month-olds learn novel word-object pairings differing only in stress pattern. *Journal of Child Language, 36*, 1157-1165.

- DeCasper, A. J., & Carstens, A. A. (1981). Contingencies of stimulation: Effects on learning and emotion in neonates, *Infant Behavior and Development*, *4*, 19-35, doi: 10.1016/S0163-6383(81)80004-5
- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, *208*, 1174-1176.
- Dunham, P. H., & Moore, C., (1995). Current themes in research on joint attention. In C. Moore and P. J. Dunham (Eds.), *Joint Attention: Its origins and role in development* (pp. 15-28). Hillsdale, NJ: Erlbaum.
- Fais, L., Werker, J. F., Cass, B., Leibowich, J., Vilela Barbosa, A., & Vatikiotis-Bateson, E. (2012). Here's looking at you, baby: What gaze and movement reveal about minimal pair word-object association at 14 months. *Journal of Laboratory Phonology*, *3*, 91-124.
- Fennell, C. T. and Waxman, S. R. (2010). What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development*, *81*, 1376–1383.  
doi: 10.1111/j.1467-8624.2010.01479.x
- Fennell, C. T., & Byers-Heinlein, K. (2014). You sound like Mommy: Bilingual and monolingual infants learn words best from speakers typical of their language environments. *Journal of Infant Behavior and Development*. *38*(4), 309-316. doi: 10.1177/0165025414530631
- Fenson, L., Bates, E., Dale, P. S., Hartung, J. P., Pethick, S. J., Reilly, J. S., ... & Thal, D. J. (1996). *The MacArthur communicative development inventories*. Singular publishing Group, Incorporated.

- Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form version of the MacArthur communicative development inventories. *Applied Psycholinguistics, 21*, 95-116.
- Frank, F. C., Vul, E., & Saxe, R. (2012). Measuring the development of social attention using free-viewing. *Infancy, 17*, 355-375. doi: 10.1111/j.1532-7078.2011.00086.x
- Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant temperament via the revised infant behavior questionnaire. *Infant Behavior and Development, 26*, 64–86.  
doi:10.1016/S0163-6383(02)00169-8
- Gogate, L. J. (2010). Learning of syllable-object relations by preverbal infants: The role of temporal synchrony and syllable distinctiveness. *Journal of Experimental Child Psychology, 105*, 178-197. doi: 10.1016/j.jecp.2009.10.007
- Goldstein, M. H., King, A. P., & West, W. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *PNAS, 100*, 8030-8035. doi: 10.1073/pnas.1332441100
- Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language, 21*, 125-155.
- Heck, A. R., Panneton, R. K., & Mills-Smith, L. (2015). Bimodal dynamics of faces and voices influence older infants' perception of emotion. *Infancy*. doi: 10.1111/inf.12121
- Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., Hennon, E., Rocroi, C., & Bloom, L. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development, (i-135)*. Maiden, MA: Blackwell Publishers.

- Hood, B. M., Willen, J. D., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, *9*, 131-134. doi: 10.1111/1467-9280.00024
- 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*, 335-344. doi:10.1162/jocn.1991.3.4.335
- Kenward, B. (2010). 10-Month-Olds Visually Anticipate an Outcome Contingent on Their Own Action. *Infancy*, *15*, 337-361. doi: 10.1111/j.1532-7078.2009.00018.x
- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocation attention to visual sequences that are neither too simple nor too complex. *PLoS ONE*, *7*(5), e36399.
- Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100*, 9096-9101.
- Lloyd-Fox, S., Széplaki-Köllöd, B., Yin, J., & Csibra, G. (2015). Are you talking to me? Neural activations in 6-month-old infants in response to being addressed during natural interactions. *Cortex*, *70*, 35-48. doi: 10.1016/j.cortex.2015.02.005
- Locke, J. L. (2001). First communion: The emergence of vocal relationships. *Social Development*, *10*, 294-308. doi: 10.1111/1467-9507.00167
- Moore, C., & Corkum, V. (1998). Infant gaze following based on eye direction. *British Journal of Developmental Psychology*, *16*, 495-503.
- Moore, C., & Povinelli, D. J. (2007). Differences in how 12- and 24-month-olds interpret the gaze of adults. *Infancy*, *11*(3), 215-231. doi: 10.1080/15250000701310348

- Mundy, P., Fox, N., & Card, J. (2003). EEG coherence, joint attention and language development in the second year. *Developmental Science*, 6, 48-54. doi: 10.1111/1467-7687.00253
- Naigles, L. R. (2002). Form is easy, meaning is hard: Resolving a paradox in early child language. *Cognition*, 86(2), 157-199. doi: 10.1016/S0010-0277(02)00177-4
- Parise, E., & Csibra, G. (2013). Neural responses to multimodal ostensive signals in 5-month-old infants. *PloS one*, 8(8), e72360. doi: 10.1371/journal.pone.0072360
- Paulus, M. & Fikkert, P. (2014). Conflicting social cues: Fourteen- and 24- month-old infants' reliance on gaze and pointing cues in word learning. *Journal of Cognition and Development*, 15, 43-59, doi: 10.1080/15248372.2012.698435
- Pons, F., Bosch, L., & Lewkowicz, D. J. (2015). Bilingualism modulates infants' selective attention to the mouth of a talking face. *Psychological Science*, 26(4), 490-498. doi:10.1177/0956797614568320
- Putnam, S. P., Helbig, A. L., Gartstein, M. A., Rothbart, M. K., & Leerkes, E. (2014). Development and assessment of short and very short forms of the infant behavior questionnaire-revised. *Journal of Personality Assessment*, 96, 445-458. doi: 10.1080/00223891.2013.841171
- Rovee-Collier, C. (1984). The ontogeny of learning and memory in human infants. In R. Kail and N. E. Spear (Eds.), *Comparative perspectives on the development of memory* (pp. 103-134). Hillsdale, NJ: Erlbaum.
- Salley, B. J., & Dixon, W. E. (2007). Temperamental and joint attentional predictors of language development. *Merrill-Palmer Quarterly*, 53, 131-154.

- Salley, B., Panneton, R. K., & Colombo, J. (2013). Separable Attentional Predictors of Language Outcome. *Infancy, 18*, 462-489. doi: 10.1111/j.1532-7078.2012.00138.x
- Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology, 18*, 668-671. DOI: 10.1016/j.cub.2008.03.059.
- Squires, J., Twombly, E., Bricker, D., & Potter, L. (2009). Ages and stages questionnaires: User's guide. City, ST: Paul H. Bookes Publishing
- Stager, C. L. & Werker, J. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388*(6640), 381-382.
- Swingle, D., & Aslin, R. A. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Phonological Sciences, 13*(5), 480-484.
- Tenenbaum, E. J., Shah, R. J., Sobel, D. M., Malle, B. F., & Morgan, J. L. (2013). Increased focus on the mouth among infants in the first year of life: A longitudinal eye-tracking study. *Infancy, 18*(4), 534-553. doi: 10.1111/j.1532-7078.2012.00135.x
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore and P. J. Dunham (Eds.), *Joint Attention: Its origins and role in development* (pp. 103-130). Hillsdale, NJ: Erlbaum.
- Tomasello, M. & Farrar, J. (1986). Joint attention and early language. *Child Development, 57*(6), 1454-1463.
- Wang, Q., Bolhuis, J., Rothkopf, C. A., Kolling, T., Knopf, M., & Triesch, J. (2012). Infants in control: Anticipation of action outcomes in a gaze-contingent paradigm. *PLoS ONE, 7*, e30884. doi:10.1371/journal.pone.0030884
- Watson, J. S., & Ramey, C. T. (1972). Reactions to response-contingent stimulation in early infancy. *Merrill-Palmer Quarterly of Behavior and Development, 18*, 219-227.

- Werker, J. F., Cohen, L., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology, 34*, 1289-1309.
- Werker, J. F., & Fennell, C. T. (2004). Listening to sounds versus listening to words: Early steps in word learning. In D. G. Hall & S. R. Waxman (Eds). *Weaving a lexicon*, 79-109. Cambridge, MA: MIT Press.
- Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy, 3*(1), 1-30.
- Werker, J. F., & Tees, R. C. (2002). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 25*, 121-133. doi: 10.1016/S0163-6383(84)80022-3
- Woodward, A. L., Infants' developing understanding of the link between looker and object. *Developmental Science, 6*(3), 297-311.
- Wu, R., & Kirkham, N. Z. (2010). No two cues are alike: Depth of learning during infancy is dependent on what orients attention. *Journal of Experimental Child Psychology, 107*, 118-136. doi: 10.1016/j.jecp.2010.04.014

Table 1. Means, standard deviations and ranges for language skill and temperament metrics for infants in both experiments.

Measure	<i>M</i>	<i>SD</i>	Range
Experiment 1			
MCDI-RV <sup>a</sup>	40.85	23.90	14-120
MCDI-EV <sup>a</sup>	8.90	10.79	0-49
ASQ-C <sup>a</sup>	49	9	30-60
ASQ-PS <sup>a</sup>	45.50	10.95	20-60
IBQ PAS <sup>b</sup>	5.15	.50	4.17-5.75
IBQ NEG <sup>b</sup>	4.31	.84	2.67-6.44
IBQ REG <sup>b</sup>	5.08	.48	4.08-5.91
Experiment 2			
MCDI-RV	40.67	20.77	11-89
MCDI-EV	9.06	8.50	0-34
ASQ-C	45.85	11.30	25-60
ASQ-PS	48.35	9.25	35-60
IBQ PAS	5.44	.53	4.50-6.23
IBQ NEG	3.95	.77	2.92-5.64
IBQ REG	5.03	.67	3.83-6.27

*Note.* Experiment 1: <sup>a</sup>*n*=20, <sup>b</sup>*n*=19. Experiment 2: *n*=18.

Table 2. Correlations Between Percent Fixation Duration on Test Trials, Language Skill, and Temperament (Experiment 1)

	Switch				Same				
	Latency	Forward		Turned	Forward		Turned		
		Object	Eyes	Mouth	Object	Object	Eyes	Mouth	Object
MCDI-RV	.36	-.17	-.07	-.05	.08	.03	.24	-.02	.15
MCDI-EV	.15	.31	-.18	.18	.23	.24	-.13	.26	.43
ASQ-C	-.05	.39	-.25	.14	.17	.33	.04	.21	.39
ASQ-PS	-.05	.36	-.03	.31	.26	.31	-.20	.20	.14
IBQ PAS	-.07	.25	.16	.23	.01	.16	-.23	.28	.06
IBQ NEG	<b>-.72*</b>	.06	-.32	-.21	-.08	-.07	.06	-.21	-.25
IBQ REG	.22	.16	.23	-.32	-.14	.19	.01	-.06	.09

Note. \* $p < .01$

Table 3. Correlations Between Percent Fixation Duration on Test Trials, Language Skill, and Temperament (Experiment 2)

	Switch				Same				
	Latency	Forward		Turned	Object	Forward		Turned	
		Object	Eyes	Mouth		Object	Eyes	Mouth	Object
MCDI-RV	.30	-.06	-.20	.16	.05	.28	-.02	.17	.43
MCDI-EV	.20	.01	-.26	-.04	-.08	.06	.15	.11	.25
ASQ-C	.45	.05	-.09	.07	.23	.42	-.22	.10	.42
ASQ-PS	.36	.16	.08	.05	.25	.35	-.39	.15	.27
IBQ PAS	-.11	.06	.03	.11	-.32	-.13	.04	.12	.17
IBQ NEG	-.20	.01	.04	.17	-.06	-.12	-.38	.34	.21
IBQ REG	.07	-.34	.08	.44	-.18	-.12	.05	.26	-.02

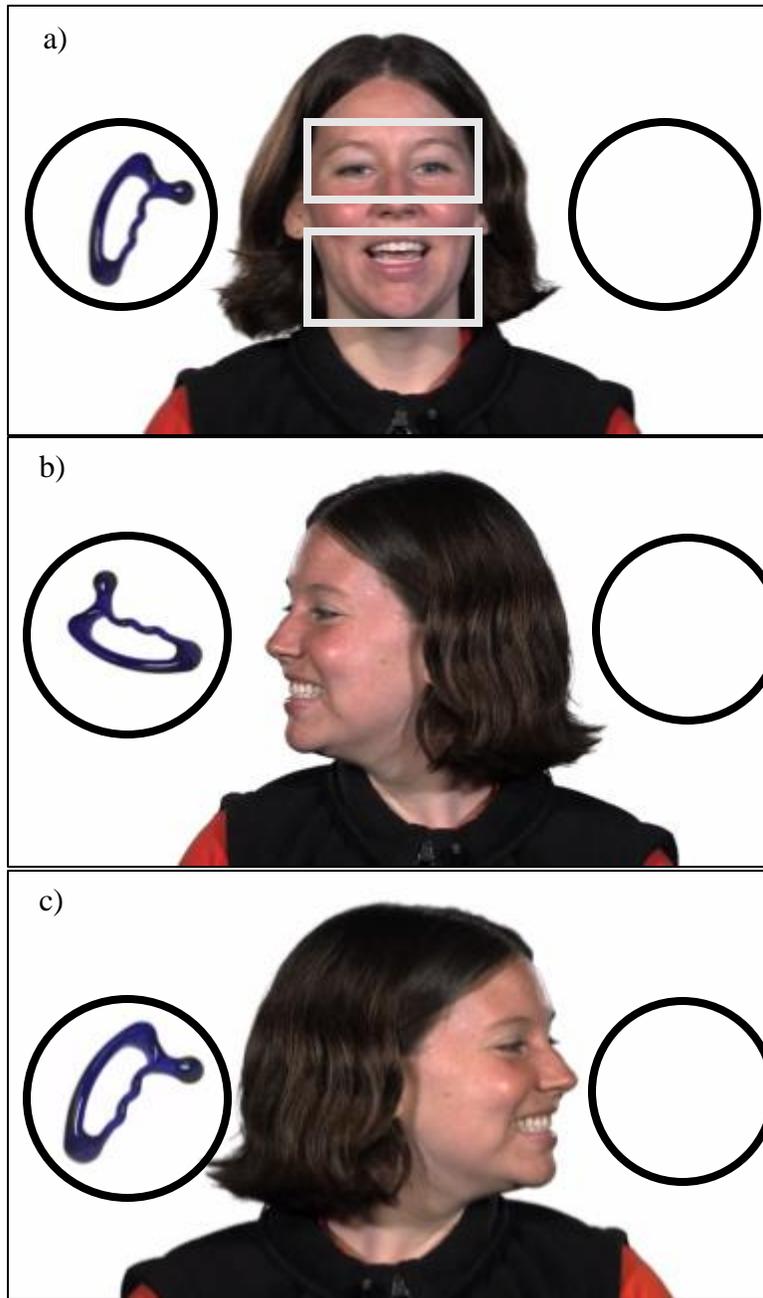


Figure 1 – Dynamic AOIs for a) Hi Baby and Forward segments (both experiments), b) Experiment 1 Turned segments, and c) Experiment 2 Turned segments. Object, eye, mouth, and non-object AOIs were drawn for Hi Baby and Forward segments, whereas object and non-object AOIs were drawn for Turned Segments.

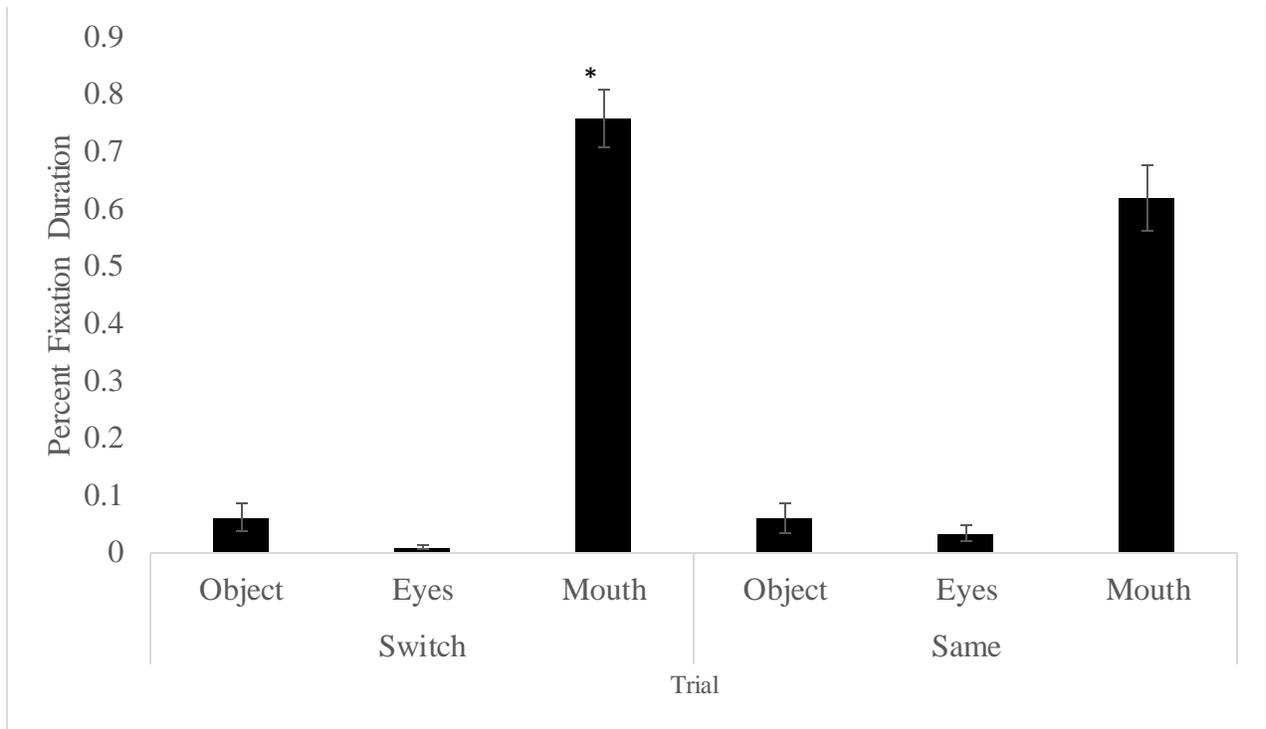


Figure 2. Average percent fixation duration on three AOIs (object, mouth, eyes) occurring during the switch and same test trials during Experiment 1 (standard errors included). There was significantly more looking to the mouth AOI during switch compared to same trials ( $*p < .05$ ).

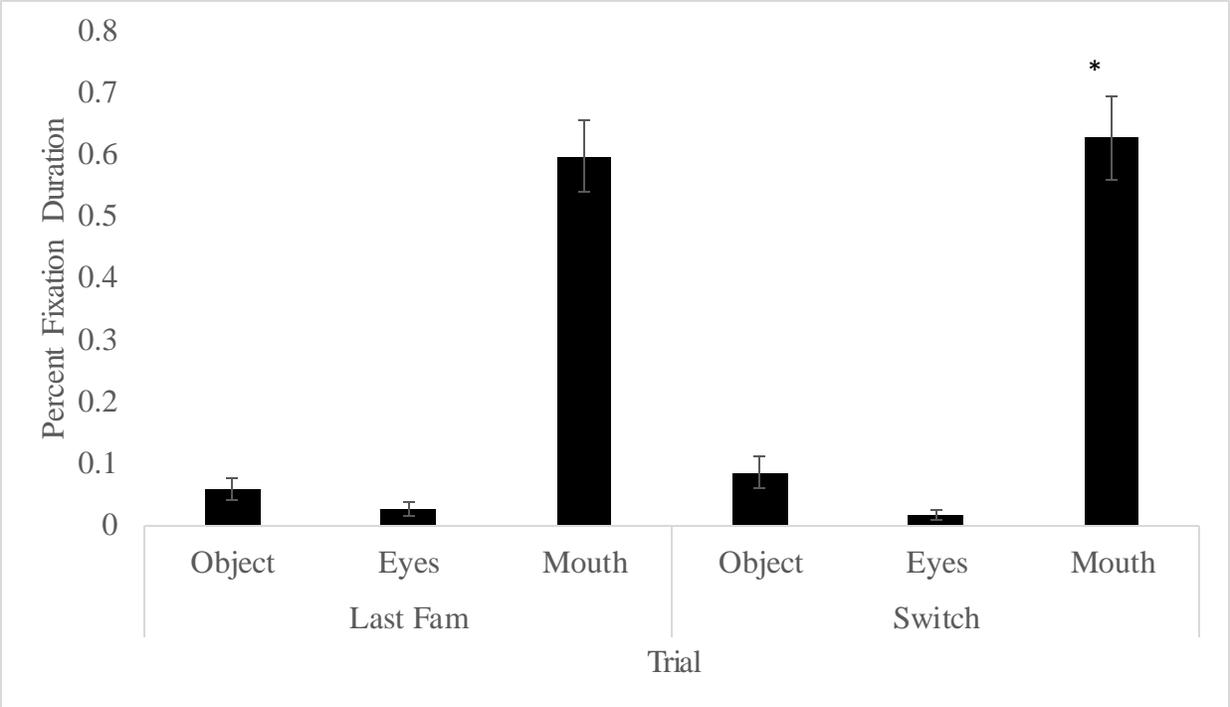


Figure 3. Average percent fixation duration on three AOIs (object, mouth, eyes) occurring during the last familiarization trial and the switch trial in Experiment 1 (standard errors are also included). There was significantly more looking to the mouth AOI during switch compared to final familiarization trials ( $*p < .05$ ).

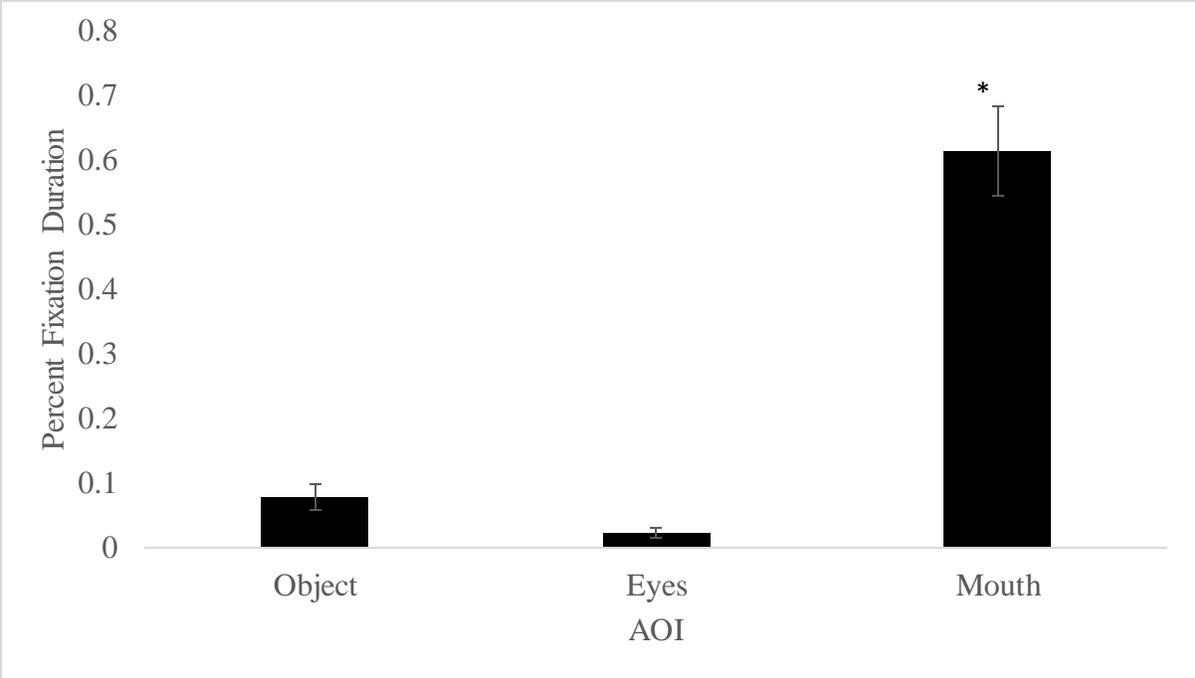


Figure 4. Average percent fixation duration on three AOIs (object, mouth, eyes) occurring during test (collapsed across switch and same trials) in Experiment 2 (standard errors are also included). There was significantly more looking to the mouth AOI ( $*p < .05$ ).

Appendix A

**i-LEAP Laboratory**  
**Family Information Sheet**  
**(All information is strictly confidential)**



Infant's Birth Month & Year: \_\_\_\_\_ Mother's Age: \_\_\_\_\_ Father's Age: \_\_\_\_\_

Mother's Occupation: \_\_\_\_\_ Father's Occupation: \_\_\_\_\_

Mother's Education:    High School    Partial College    College    Master's    Ph.D.

Father's Education:    High School    Partial College    College    Master's    Ph.D.

Annual Family Income:    \$10,000-\$20,000    \$20,000-\$35,000    \$35,000-\$50,000  
   \$50,000-\$65,000    \$65,000-\$80,000    \$80,000-\$95,000    > \$95,000

Marital Status:    Married    Separated    Divorced    Unmarried/Single    Widowed

Mother's Race:    White/Caucasian    African American    Hispanic    Asian    Native American  
Other \_\_\_\_\_

Father's Race:    White/Caucasian    African American    Hispanic    Asian    Native American  
Other \_\_\_\_\_

Was your infant: Full Term (38-42 weeks)    Premature (≤ 37 weeks)    Postmature (>42 weeks)

Infant's Birthweight: \_\_\_\_\_ lbs \_\_\_\_\_ ozs

Has your infant had any medical problems? Yes    No    Please List: \_\_\_\_\_

Please list the birth month/year and gender of any older children:

\_\_\_\_\_ M F \_\_\_\_\_ M F

\_\_\_\_\_ M F \_\_\_\_\_ M F

What is the primary language spoken in your home? \_\_\_\_\_

Please list any other languages that are spoken in your home: \_\_\_\_\_

Has any child in the family been suspected of a developmental delay/diagnosis? Yes    No

If yes, please describe: \_\_\_\_\_

Have you been diagnosed with depression since your infant was born? Yes    No

Do you currently take any medication related to clinical depression? Yes    No

Appendix B  
**Infant Behavior Questionnaire-Revised Very Short Form**

Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant temperament via the revised infant behavior questionnaire. *Infant Behavior and Development*, 26, 64–86.  
doi:10.1016/S0163-6383(02)00169-8

Appendix C  
**Ages and Stages Questionnaire**

Bricker, D., & Squires, J. (1999). *Ages and Stages Questionnaires: A parent-completed child-monitoring system*. Baltimore, MD: Paul H. Brookes Publishing Co.

Appendix D  
**MacArthur Communicative Development Inventory-Short Form**

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form version of the MacArthur communicative development inventories. *Applied Psycholinguistics*, 21, 95-116.

## Appendix E

### Maternal Sensitivity Task

Maternal Sensitivity Task 1

#### Maternal Sensitivity Task

*Adapted from Fish, Belsky, & Stifter (1991); Shapiro & Mangelsdorf (1994); Calkins, Hungerford, & Dedmon (2004); MacDonald & Park (1984); with input from Cindy Smith and her lab's coding schemes and changes from Robin Parmeton and her graduate students (Spring 2015).*

Epoch: 30 sec

#### Facilitates Attention

*To what extent and how skillfully did mother facilitate the infant's response to the object through such behavior as positioning the infant so that he/she can see the object and is comfortable; drawing the infant's attention to the object verbally ("See the keys?") or by pointing at or tapping it.*

1. Very low; little (if any) facilitation, or behavior that doesn't help the infant focus on the target.
2. Low to moderate. There may be a few brief instances of facilitative behavior based on infant's behavior, but mother is not really directing infant's attention.
3. Fairly high, most of the mother's facilitative behavior is appropriate, there are a few brief instances of mother not helping infant focus.
4. Very high; mother exhibits appropriate and facilitative behaviors and seems especially turned in to her infant (e.g., refocuses the infant's attention when he/she loses interest).

#### Demonstrates Contingency

*To what extent did the mother respond directly to the infant's behavior; issuing either verbal or physical responses to something the infant did in a timely manner (Infant looks at rattles, and mother says "You like those rattles, don't you?"; Infant knocks down blocks and laughs, mother says "Did you like knocking those blocks down?")*

1. Very low; mother does not respond to what the infant is doing, either in her own behavior and/or in her vocalizations.
2. Low to moderate; mother occasionally responds to what the infant is doing, but in a poorly timed way (e.g., interrupts the infant's activity by responding too quickly or "misses" the infant's activity by not responding quickly enough).
3. Fairly high; mother is often responding to infant's activity but either timing is not sensitive or she is inconsistent in whether she acts or speaks to infant based on what infant is doing.
4. Very high; mother consistently responds to infant's behavior in a timely manner. Very tuned into what the infant is currently doing.

#### Intrusiveness

*To what extent does the mother display intrusive, over-controlling behavior? Does the mother address her own agenda and ignore or override the infant's preferences, such that the child's behavior is redirected. Does the mother fail to moderate her behavior after infant shows disinterest or aversion to her? Does she offer a barrage of toys without gauging infant's*

*interest? Take away toys infant is still observing? Intrusively grab the infant's face to direct his/her attention?*

1. Very low; mother shows little to no instances of intrusive or over-controlling behaviors.
2. Low to moderate; mother may show a few instances of intrusive or over-controlling behavior but more often than not, the mother is not intrusive or over-controlling.
3. Fairly high; mother mixes intrusive and over-controlling behavior with short periods of non-intrusive or non-controlling behavior.
4. Very high; mother is consistently intrusive or over-controlling; constantly redirects the child's behavior or attention, without regard for the current focus of the infant's attention. Mother engages in activity (whether physical, verbal, or both) to a level that seems to make the infant uncomfortable.

**Maternal Positive Affect**

*To what extent does the mother convey positive emotion to the infant, either in her mannerisms, her voice, or both? Does she consistently engage the infant with a positive tone in her voice? Does she smile, laugh, and nod in a positive way when responding to the infant's behavior?*

1. Very low; Little to no positive affect; mother expresses no positive emotion when communicating with child either in; mother's emotional expression was neutral or negative
2. Low to moderate; intensity positive; slight or very brief smile.
3. Fairly high; clear smile or prolonged slight smiles; uses prolonged positive tone.
4. Very high; intense smile or laugh, or smiling for more prolonged period; uses positive tone throughout episode.

*NOTE---* for the above variables, ratings are summed across epochs and then divided by the number of epochs to get a mean value.

Appendix F  
Diagram of Experiments



Experiment 1

Experiment 2



Appendix G  
Objects Used in Word-Object Learning Task



# Appendix H Informed Consent

## VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

### Informed Consent for Participants in Research Projects Involving Human Subject

**Title:**

The Role of Contingent Eye Gaze on 14-Month-Olds' Word-Object Associative Learning for Minimal Pairs

If for any reason, your infant cries or falls asleep during either task, testing will be discontinued.

**Principle Investigators:**

Robin Panneton, Ph.D.  
Laura Mills-Smith, M.S.

**III. Risks**

There are no apparent risks to your infant or to yourself for participation in this study. Sound levels for all auditory stimuli will be verified prior to the testing of each adult and infant. All sound tracks are kept at @ 68 dB SPL at the head of the listener. This is the sound level of normal conversational speech.

**I. Purpose of this Research/Project**

The purpose of this project is to determine the influence that eye gaze from a social partner plays on language learning during late infancy. Specifically, we are interested in how infants' social regard may influence their ability to learn word-object relations.

**IV. Benefits**

Although there are no direct benefits to the participants in this study, all parents will receive a summary report of the results of this project (a general analysis of the patterns of looking and language learning across all of the infants). Parents will also receive a certificate of appreciation and the results of the study will contribute to a broader body of research on infant social attention and language learning.

**II. Procedures**

Following the completion of questionnaires, you and your infant will be tested for approximately 30 minutes. For the eyetracking task, the baby will be placed in a highchair sitting next to you. The infant will face a monitor of an eye-tracking system. All movies will be presented on this screen. First, the infant will see a simple screen with moving dots so that we can calibrate their eye positions. Following calibration, the infant will see a series of movie clips of a woman speaking and looking at objects. The woman on the screen will name two objects several times, and then we will test the infant by presenting them with one of the labels and a screen full of objects. We will measure how accurately the infant looks at the object being named.

**V. Extent of Anonymity and Confidentiality**

All of the information gathered in this study will be kept confidential and the results will not be released without parental consent. However, the results of this project may be used for scientific and/or educational purposes, presented at scientific meetings, and/or published in a scientific journal.

The second task simply requires you and your infant play with toys for 15-20 minutes in a typical fashion. This play will be recorded with a video camera to determine your child's interactive style.

**VI. Compensation**

There is \$10 compensation to be earned from participation in this project.

**VII. Freedom to Withdraw**

You have the right to terminate your or your infant's involvement at any point in time and for any reason should you choose to do so. You will be fully compensated if you choose to discontinue participation.

**VIII. Participant's Responsibilities**

I voluntarily agree to have my infant participate in this study. I voluntarily agree to participate in this study.

**IX. Participant's Permission**

I have been given an opportunity to ask further questions about this procedure and I understand that I have the right to end this session for any reason if I so choose. This project has been approved by the Human Subjects Committee of the Department of Psychology and the Institutional Review Board of Virginia Tech. If I have any questions regarding this research and its conduct, I should contact one of the persons named below. Given these procedures and conditions, I give my permission to Dr. Panneton, Ms. Mills-Smith, and their co-workers to test my son/daughter.

I hereby acknowledge the above and give my voluntary consent for my infant to participate in this study.

**Infant's Name:** \_\_\_\_\_

**Signature of Parent:** \_\_\_\_\_

**Date:** \_\_\_\_\_

I hereby acknowledge the above and give my voluntary consent to participate in this study.

**Parent Printed Name:** \_\_\_\_\_

**Signature of Parent:** \_\_\_\_\_

**Date:** \_\_\_\_\_

I would like to be contacted by phone regarding future studies:    YES        NO

Dr. Robin Panneton, Principle Investigator	231-5938
Laura Mills-Smith, Graduate Student and Co-Investigator	231-3972
Dr. David Harrison, Chair, Human Subjects Committee	231-4422
David M. Moore, DVM, Assistant Vice Provost for Research Compliance	231-4991