Here’s looking at you, baby: What gaze and movement reveal about minimal pair word-object association at 14 months

LAUREL FAIS*, JANET F. WERKER*, BRONWYN CASS*, JULIA LEIBOWICH*, ADRIANO VILELA BARBOSA†, and ERIC VATIKIOTIS-BATESON†

*Department of Psychology, University of British Columbia
†Cognitive Systems Program and Department of Linguistics, University of British Columbia

Abstract

The ability, or lack thereof, of 14-month-old infants to associate novel, minimal pair word-forms with novel objects in a variety of experimental settings has been a crucial research window into how infants go about the quintessential linguistic task of learning words. Here we ask whether the presence of a human interactor in the experimental setting facilitates minimal pair word-object association at this age. In addition to standard looking time measures to test this question, we also introduce the use of measures of infant movement derived by the application of an efficient algorithm that measures motion from 2D video. Infant gaze patterns across the experimental session identified two groups of infants, those engaging in more and those engaging in less mutual gaze with the Experimenter; both groups demonstrated success in the task by both looking time and movement measures. Infants did not succeed in the task by either measure when a videotaped Experimenter presented the labels. We suggest that infants at this age are in transition from being good “information consumers” to becoming good “information seekers,” and that the presence of the live Experimenter plays a crucial role in making it possible for infants to demonstrate their nascent word learning abilities. Further, we explore insights into the looking time results provided by the movement measures as well as novel contributions to our understanding of language acquisition afforded by the examination of infant movement.

1. Background

When infants are interested in something, they look at it. Researchers in the area of infant language acquisition have taken advantage of this fact for decades, measuring how long infants look at a visual stimulus co-occurring with an auditory stimulus in order to infer information about infants’ understanding of the auditory...
stimulus. Though widely used, looking time measures are indirect and need to be interpreted with a responsible level of caution (Aslin 2007; Cohen 2004). In this work, we examine the relationship between looking time measures and movement measures in two experiments in infant language acquisition. We demonstrate that measures of infant movement can corroborate looking time results as well as provide new ways to understand how infants process language stimuli.

One of the intriguing findings in infant language development research based on looking time measures concerns the ability of 14-month-old infants to associate minimal pair word forms and nonsense objects in a word-object association task. Very young infants are sometimes characterized as “universal listeners,” able to discriminate sounds from a wide variety of the world’s languages, but by the end of their first year of life, they are narrowing their perception of speech sounds to those sounds that appear only in their own language environment (Kuhl et al. 2006; Pegg and Werker 1997; Werker and Tees 1984). At 14 months, infants draw on their speech sound discrimination abilities to successfully associate two very different word-forms (e.g., lif [lɪf] and neem [nim]) to nonsense objects (Stager and Werker 1997). However, they are unable to link two very similar word-forms (bih [bɪ] and dih [dɪ]) to novel objects, despite the fact that they can discriminate the sounds of those two word-forms (Pater et al. 2004; Stager and Werker 1997). This puzzling result has been explored in a host of subsequent studies aiming to determine just what factors contribute to success in this task (see Werker and Fennell [2008] for a review). In the current work, we use both looking time and movement measures to examine the effect of a live interactor in the experimental context on 14-month-old infants’ ability to link minimal pair word-forms and novel objects. Our results both broaden our understanding of the phenomenon and demonstrate the novel contributions of movement measures in infant research.

1.1. Movement

Movement has not seen much use as a dependent variable in infant development research. Some work has investigated the relation between motion, especially approaching/withdrawal movements, and affect or preference; for example, Weinberg and Tronick (1994) used the Infant Regulatory Scoring System (Tronick and Weinberg 1990) to correlate infant gestural and postural behavior in mother-infant interactions with affective states (“social engagement,” “object engagement,” “passive withdrawal,” and “active protest”). Zentner and Kagan (1998) used infant “motoric activity” and measures of “turning away” to assess infant preference for consonant or dissonant music. Head turning movement that facilitates gaze toward a reinforcer in the head turn preference or observer-based psychoacoustic procedure has also been used to measure infant response, especially to acoustic auditory stimuli (Olsho et al. 1987; Werker, Shi et al. 1998). In all these cases, instances of pre-defined types of movement were counted, but their kinematics were not measured. Little if any work has been done using measures of motion itself as a means
to understand cognitive processes such as categorization or language acquisition. There have been good reasons for this: until recently, movement has been extraordinarily labor-intensive to measure and to analyze, severely limiting the number and types of movements that tend to get measured. This is potentially problematic because the infants themselves are under no such constraints; they appear to move any and all parts of their bodies that they are capable of moving.

One indication that infant movement can yield insights into infant processing of auditory stimuli comes from a large study examining the movement responses of infants to either language or music stimuli (Fais et al. 2010). In this study, no expense was spared in an attempt to comprehensively hand-code every detectable movement (with the exception of facial expression) of 20 infants 5- to 6-months old as they participated in laboratory studies in which 10 infants listened to language stimuli, and 10 to music. Using these hand-coded data, it was possible to identify which body parts moved reliably differently to language and to music, and which did not. Infants move their heads and torsos more overall in the language condition than the music condition. This includes shifts of gaze, which should be no surprise, since looking time is a venerable measure of infant interest and implicates gaze shift. But in addition, these movement measures revealed greater head and torso movement accompanying voiced vocalization directed away from the source of the stimuli (audio speakers located behind a curtain at the front of the room). Fais and colleagues propose that this pattern of movement indicates searching behavior by the infants in the language, but not the music, condition. Whether or not their interpretation is correct, it is undeniably the case that infants moved differently in these two conditions, and that that movement was evidenced not only in their gaze shifts, but also in their vocalizations, and head and torso movement.

This result is interesting, but it was achieved at the cost of hundreds of hours spent hand-coding a few minutes of video data per infant. Although revealing, this approach has the great disadvantage of being labor-, time-, and funding-intensive. By contrast, the work we report on here demonstrates the value of using an extremely efficient optical flow algorithm to derive 2D measures of movement automatically from even relatively low-quality video. Both looking time and motion measures were made for 16 14-month-old infants who participated in a word-object association task. The two sets of results are remarkably consistent. Further, the movement measures extend our understanding of infant word-object association beyond the insights that looking time measures alone provided. We suggest that, just as infant looking is not random, neither is infant movement; infant motion is structured and integrally linked to cognitive functioning in these experimental tasks, and an examination of motion can yield new insights into that functioning.

1.2. Word-object association in 14-month-old infants

Stager and Werker demonstrated 14-month-old infants’ ability to associate lif and neem with nonsense objects, and inability to associate the minimal pair bih and dih
with nonsense objects, using a habituation/dishabituation procedure known as the “Switch” task (Werker, Cohen et al. 1998). In this paradigm, the looking time of the infant to objects shown on a monitor acts as the dependent measure. An infant seated on her caregiver’s lap is presented with tokens of two pairings of a novel object and a novel word-form (presented through speakers) until her looking time to the objects decreases to a criterial level, indicating that she has habituated to the pairings. In one of the two trials of the test phase, the infant sees a pairing identical to one of those already presented; this is called the Same trial. In the second, she sees a pairing that is switched, i.e., a pairing of a familiar word-form and habituated object that had never occurred together, the Switch trial. Greater looking time to this switched pairing signals the infant’s awareness of a difference in the pairing of word-forms and objects, and thus, recognition of the word-object pairings presented during the habituation phase.

A simpler, one-object, version of the task pairs just one word-form, say, *bih*, with one object during the habituation phase. At test, the infant is exposed to a Same trial identical to those occurring in the habituation phase, and a Switch trial that couples the object with a new word-form, in this case, *dih*. Eight-month-old infants succeed at this simpler task with these minimal pair words, but 14-month-old infants fail even this less complex version (Stager and Werker 1997). Why this disparity?

It turns out that 8-month-old infants, unlike 14-month-olds, fail to associate the different sounding *lif* and *neem* with two objects in the two-object task. This suggests that the success of the 8-month-old infants in the one-object *bih/dih* task rests not on an ability to associate a word-form with an object, but simply on the ability to discriminate the minimal pair *bih* and *dih*. By 17 months, however, infants show robust success in associating *bih* and *dih* to two novel objects (Werker, Fennell, Corcoran, and Stager 2002). This configuration of results gave rise first to the resource limitation hypothesis (Fennell and Werker 2003; Werker et al. 2002), and later to a developmental framework for Processing Rich Information from Multidimensional Interactive Representations, PRIMIR (Werker and Curtin 2005; Curtin and Werker 2007), a more integrative framework. It was suggested that at 14 months, infants perceive the phonetic particulars that distinguish word-forms, but do not yet know that that information is any more important in this task than indexical cues such as intonation, gender of the voice, etc. Hence, the failure of 14-month-old infants is explained by their inability to preferentially attend to contrastive phonetic information over information such as intonation, loudness, etc., while at the same time processing the association between word-form and object. By 17–20 months infants have learned that it is the phonetic differences that guide word learning (Dietrich et al. 2007), and are able to highlight that information as they process the Switch task, thus succeeding in pairing two minimal novel word-forms and two novel objects.

This view receives support in a variety of subsequent studies with 14-month-olds. The logic of these studies suggests that if some aspect of the task demand is
eased, or if the task highlights the phonetic contrast, there is a greater likelihood that infants aged 14 months can succeed. Infants succeed if they are familiar with the words (e.g., \textit{ball} and \textit{doll}; Fennell and Werker [2003]; Swingley and Aslin [2002]) or with the object (Fennell 2011), or if given a two-choice test phase (Yoshida et al. 2009); in each of these cases the task demands are lessened. Infants also succeed when habituated to variable phonetic information from multiple speakers (Rost and McMurray 2009) or to minimal pairs embedded in disambiguating phonetic contexts (Thiessen 2007).

Fennell, Waxman and colleagues reasoned that, at 14 months, infants might highlight the relevant phonetic detail and succeed in the Switch task if it were made clear that the experimental context was one of word learning. Indeed, infants succeeded under two referential conditions: when the nonsense word-forms were embedded in short naming phrases, e.g., \textit{Look, it's the bin!} (Fennell 2006; Fennell and Waxman 2010), and when a referential context was established prior to habituation by presenting infants with familiar word-object pairings, e.g., a picture of a shoe and the word \textit{shoe} (Fennell and Waxman 2010; Fennel et al. 2007).

2. The current study

Word learning “in the wild” entails not only understanding that word-forms label objects, but also that those word-object relationships are shared by a community of speakers (Clark 1991; Diesendruck 2005; Henderson and Graham 2005). Indeed, in a typical word learning situation, an infant hears word-forms produced by a human interactor who is speaking with the intent to communicate information to the infant, often using gaze and other gestural cues to pick out a relevant object in her environment as she embeds labels in an informative, referential context. We reasoned that another way to support minimal pair word-object association would be to include a human interactor in the experimental situation.

In Experiment 1, we replaced the recorded auditory stimuli typically used in the Switch task with live auditory stimuli delivered by an Experimenter seated in the study room with the infant and caregiver. The presence of this second salient visual stimulus allows us to examine how infants pattern their looking to the object and to the Experimenter.

To determine how just the presence of an adult interactor might affect infant performance, we attempted to eliminate other sources of information that could support success. We restricted the Experimenter to delivering single word-forms, in infant-directed speech (just as in previous recorded stimuli), and to keeping her hands folded in her lap. She did, however, in the interests of naturalness, alternate gazing pleasantly at the infant with turning to look at the object displayed on the monitor. We know that gaze following in 10- to 11-month-old infants is associated with faster rates of vocabulary acquisition (Brooks and Meltzoff 2008), and that when an Experimenter labels an object in the line of regard of the infant, that infant
is better able to make novel word-object associations (Baldwin 1991; Baldwin et al. 1996; Booth et al. 2008; Dunham et al. 1993). Further, coupled with a richly social, multimodal setting, infant gaze-switching from mother to object is associated with greater success in forming word-object relations in 6- to 8-month-old infants (Gogate et al. 2006; Matatyaho and Gogate 2008). The presence of the Experimenter in our study allows us to examine how gaze affects the association between novel, minimal pair word-forms and novel objects at 14 months.

There are a number of reasons why gaze could be an effective cue. Researchers like Baldwin, Tomasello and others (e.g., Baldwin et al. 1996; Tomasello et al. 1996) maintain that the naming intention inherent in gaze facilitates infant word comprehension; Csibra and Gergely (2009; cf. Csibra 2010) claim that infants are born prepared to make use of this information. On the other hand, infant performance might be boosted by the visual language cues on the face of the Experimenter (MacLeod and Summerfield 1987; Teinonen et al. 2008). We attempt to tease apart the social effects of gaze from the processing effects of visual cues in Experiment 2. We videotaped the same Experimenter who was present in the study room during Experiment 1 such that she appeared to alternate her gaze between looking at the infant and at the object projected on the monitor to the left of the video of the Experimenter. This recording constituted the audio-visual source for the word-forms in Experiment 2. Kuhl and colleagues (Kuhl et al. 2003) found that 9-month-old infants could learn a non-native phonetic contrast if taught by a person, but not if taught by a video of that person. This finding, as well as the social interpretation of the effects of gaze, predict that the infants in our two different word-object associative learning contexts will perform differently: those experiencing the Experimenter live will show enhanced word-object association over those experiencing the videotaped presentation. A visual language processing explanation, on the other hand, predicts that infant performance will be similarly augmented by the presence of visual speech cues in both the live and videotaped presentations.

In the typical Switch task, success is measured by looking time to the objects, which constitute the most salient visual images in the infant’s field of vision. The presence of an Experimenter adds a second, quite attractive focus of visual attention, which complicates the interpretation of looking at or away from the object as the sole measure of infant success in the task. In previous work, in which the object was the only salient visual focus, it was reasonable to interpret looking at the object as interest, and looking away as disinterest. In the presence of this second, extremely salient point of visual attention, i.e., the Experimenter, looking away from the object does not necessarily imply disinterest in the object; it simply means that the infant is less interested in the object than he is in the Experimenter. Thus it was crucial that we also measure how long infants looked at the Experimenter, and we therefore explored looking time to both object and Experimenter in our analyses.

In addition to these methodological considerations, there are compelling theoretical reasons to examine patterns of infant looking to the Experimenter as well.
Koenig and Echols (2003) examined 16-month-old infants’ looking to a familiar object and to four different sources for its label (which was either true or false): a human gazing at the object, an audio speaker, a hidden audio speaker with a silent human, and a human gazing away from the object. When the object was labeled incorrectly, the infants looked significantly longer at the human labeler, but not to any of the other three sources. Koenig and Echols suggest that infants at this age understand humans to be sources of accurate information about the world, and register surprise by looking at the human labeler, rather than at the object, when this expectation is thwarted by a false labeling event. The Switch trial acts as a false labeling event for the newly learned nonsense word/novel object pairings in our task; thus, we might expect infants to register surprise not by looking longer at the object as in the standard Switch task, but by looking longer at the Experimenter.

Though the looking behavior of the Experimenter in both experiments “invited” infants to share looking at the object, and provided them with the opportunity for mutual gaze with her, infants may or may not actually have taken advantage of that “invitation.” For that reason, we calculated three specific measures that sum infant looking across pretest, posttest, and test trials in order to characterize infant looking behavior across the entire experiment session:

- the amount of time the infant spent looking at the Experimenter overall, regardless of the Experimenter’s line of regard (overall looking time to Experimenter)
- the amount of time the infant spent looking at the object at the same time that the Experimenter was looking at the object (shared looking to object)
- the amount of time the infant spent looking at the Experimenter at the same time that the Experimenter was looking at the infant (mutual gaze with Experimenter)

We examined how these behaviors affected looking time to both the object and the Experimenter in the test trials.

In addition to looking time, we also measured infant movement during test trials. We used an optical flow algorithm to derive movement measures automatically, efficiently and objectively, from the 2D videos taken of the infants as they participated in the two studies. We investigated how these movement measures interacted with infant looking patterns to the Experimenter and to the object, and compared the results derived from the two types of measures.

3. Experiment 1

3.1. Methods

3.1.1. Participants. Sixteen 14-month-old infants (8 female, 8 male; range: 14 months, 7 days to 15 months, 6 days; mean: 14 months, 19 days), hearing at least
80% English at home participated in Experiment 1. Nineteen additional infants participated but were excluded from analysis due to fussiness (7), experimental error (3), less than one second looking time to object at test (4), parental interference (3), and lack of habituation within set parameters (2). All participants were healthy, full-term infants, recruited from a local women’s hospital; university Behavioral Research Ethics Board guidelines for the ethical treatment of human subjects were observed.

3.1.2. Stimuli. Visual stimuli for the objects consisted of two colorful novel shapes, formed from plasticine and photographed against a black background. The shapes were presented singly, on a 32 inch flatscreen television monitor, and were animated to move slowly back and forth across the width of the frame.

Auditory stimuli were produced by an Experimenter seated on the infant’s right, between the infant and the monitor, at a 90 degree angle from both (Figure 1), wearing the same black sweater and dark pants for each session. Replicating previous Switch task experiments, she delivered the labels in a friendly, infant-directed tone. She was instructed to divide her time between looking at the infant and looking at the monitor in a pleasant and natural way; this meant that, in some cases she met the infant’s gaze and “invited” shared looks to the monitor by turning her head and gazing at the object. At other times, if the infant did not happen to look at her in an appropriate time frame, she turned to look at the object without having established gaze with the infant. The dimly lit room did not contain any other objects as salient as the moving object on the screen and the Experimenter. Given the interest that infants exhibit in people and faces (Frank et al. 2009), the

Figure 1. The arrangement of the study room for Experiment 1; the infant is seated on the caregiver’s lap (right) in front of the monitor showing the objects, and the Experimenter is seated to the infant’s right.

©Julia Leibowich 2009
Experimenter thus constituted a second visual stimulus as well as providing the source for the auditory stimulus.

3.1.3. Procedure. The testing was conducted in a small, sound-attenuated room. The infant sat on his caregiver’s lap in front of the television monitor, which he viewed through an opening in a dark curtain that completely masked the experimental equipment. Caregivers wore headphones delivering entertaining music so that they could not hear the Experimenter. The experiment was controlled using Habit (Cohen et al. 2002); the audio output from Habit signaled the Experimenter as to which word-form she should produce for that trial. The Experimenter wore a small ear bud in her right ear (the side away from the infant so that the infant could not see it) that enabled her to hear the signal tones.

Each testing session began with a 15 sec pretest consisting of a visual stimulus of a colorful, moving waterwheel coupled with the Experimenter saying the word-form *neem*. This pretest trial allowed infants to become accustomed to the room, to the Experimenter, and to the fact that pictures of objects would be presented on the monitor. After the pretest, the habituation phase of the experiment began. Once infants fixated on a bright, animated attention-getter (during which presentation the Experimenter sat silently with a pleasant expression on her face), they were presented with one of two nonsense objects, moving back and forth across the screen on a black background, and heard the Experimenter say either *bin* or *din*. During the habituation phase, infants experienced a semi-randomized series of consistent word-object pairings: *bin* paired with either a colorful rounded object or an object resembling a molecule, and *din* paired with the other object, for a maximum of 24 trials.

A digital video camera whose lens poked out of the curtain below the monitor (see Figure 1) recorded the infant’s face and allowed a second, distal experimenter to view the infant’s behavior and measure looking time to the object on the monitor via the Habit software. This measure was chosen for determining habituation in order to maintain consistency with previous work. The Experimenter was recorded via a second video camera, also positioned behind the curtain, but placed so that it captured a frontal view of the Experimenter as she gazed directly in front of her, i.e., at a 90 degree angle from both the participant and the monitor. In this way, the Experimenter’s gaze to both the infant on her left and the object on the monitor to her right could be clearly tracked.

Once a criterial decrease in looking time was reached (pre-defined as a block of four habituation trials in which the infant looked 65% or less of the longest amount of looking time registered in any previous four-block window), the test phase began. Habit delivered the visual test stimuli to the monitor as well as the signal tones for the appropriate auditory test stimuli to the Experimenter. The test phase consisted of two test trials. One test trial repeated one of the sets of visual-auditory habituation stimuli (the Same test trial); the second presented a mis-pairing of habituated visual and auditory stimuli (the Switch trial). In the Switch trial, either
the same word-form as that in the Same trial was repeated, paired with the other object, or the same object as the one presented in the Same trial was presented, paired with the other word-form. Thus, in the Switch trial, a novel pairing of habituated word-forms and objects was presented. We counterbalanced across gender whether the word-form or object was switched, the order of the Same and Switch trials, and the pairings of word-form and object. After the test trials, there was a posttest phase, during which participants were presented with the same waterwheel and neem stimulus they had experienced in the pretest.

3.1.4. Data collection: Looking time. Both infant looking and Experimenter looking during the pretest, test trials, and posttest were coded from the videotapes, using software that allowed frame-by-frame accuracy (approximately 30 frames per sec). The videos of the participant and the Experimenter were synced (using FinalCut Pro) on the onset of the auditory stimulus. Two dependent measures were taken: the amount of time the infant looked at the object during test (looking time to object), and the amount of time the infant looked at the Experimenter during test (looking time to Experimenter). In addition, we summed measures from the pre- and posttests as well as from the test trials to characterize infant looking in three different ways: the overall amount of time the infant looked at the Experimenter (overall looking to the Experimenter), the amount of time that the infant looked at the object while with Experimenter was looking at the object (shared looking to object), and the amount of time that the infant looked at the Experimenter while the Experimenter was looking at the infant (mutual gaze with Experimenter).

3.1.5. Data collection: Movement. We analyzed the motion of each of the 16 infants using a procedure that derives 2D motion from video for regions of interest (ROI) that can be defined within the video frame (Barbosa et al. 2008). Horizontal and vertical motion is derived from pixel intensities using Horn and Schunk’s (1981) optical flow analysis algorithm. In essence, the algorithm derives, for every pair of consecutive frames in the video sequence, a matrix containing both the horizontal and vertical velocities of the transition of each pixel from the first frame to the second frame in the pair. The 2D pixel motion can also be expressed as a vector in polar coordinates consisting of an angle and a magnitude. For a given ROI (in our case, the entire frame), pixel magnitudes can be summed yielding an overall magnitude of motion (Figure 2). For this study, we generated a summed motion magnitude time series for both test trials of all of the infants, from the videos taken of the infants during the experiment.

3.2. Results

3.2.1. Dependent measure: Looking time to the object. A 2 (test trial: Same, Switch) × 2 (order: Same-first, Switch-first) × 2 (gender: female, male) repeated...
measures ANOVA revealed no main effect and no interactions. Infants did not, as a group, look significantly differently to the object during the Same and Switch trials. Neither the order of the test trial presentation nor the gender of the participant had a significant effect on participants’ looking time to the object on the monitor; therefore we do not include these factors in the analyses reported below.

Because patterns of gaze behavior have been shown to affect infants’ ability to associate word-form and object, we looked more closely at how infants distributed their looking between the object and the Experimenter during the experimental session. For each participant, we determined the overall time spent looking to the Experimenter, the time spent in shared looking at the object while the Experimenter also looked at the object, and the time spent in mutual gaze with the Experimenter, from the pre- and posttests and the test trials. Using the median score for each of these three measures, we divided the participants into two groups for each measure: the infants whose average looking time measures were above the median were considered to have demonstrated a relatively “high” amount of the behavior, the infants below, a “low” amount. In three repeated measures ANOVAs, we examined whether infants showing different amounts of these three looking behaviors across the experimental session also showed different patterns of looking to the object during test trials.

For overall looking time to the Experimenter, a 2 (test trial: Same, Switch) × 2 (overall looking to Experimenter: high, low) ANOVA revealed no main effect and no interaction. That is, whether infants looked for a relatively longer or shorter amount of time at the Experimenter over the course of the session did not modulate their ability to demonstrate success in associating the novel words and novel objects.
Likewise, for shared looking to the object with the Experimenter, a 2 (test trial: Same, Switch) × 2 (shared looking to object: high, low) ANOVA revealed no main effect and no interaction. The amount of time infants spent looking at the object with the Experimenter, i.e., at the same time that the Experimenter was looking at it, did not differentiate their performance at test.

For mutual gaze with the Experimenter, a 2 (test trial: Same, Switch) × 2 (mutual gaze with Experimenter: high, low) ANOVA revealed no main effect, but did show an interaction (test × mutual gaze: $F(1, 14) = 4.52, p = .052, \eta^2_p = .24$). Subsequent t-tests revealed that infants who participated in relatively more mutual gaze with the Experimenter across the experimental session, but not infants who had relatively less, looked significantly longer to the object in the Switch trial than they did in the Same trial ($t(7) = -3.48, p = .01, \eta^2_p = .63$). That is, infants who participated in relatively more mutual gaze with the Experimenter demonstrated success in associating two novel, minimal pair word-forms with two novel objects (as measured by their looking time to the object stimuli), while those who participated in less mutual gaze with the Experimenter did not (Figure 2).

3.2.2. Dependent measure: Looking time to the Experimenter: The Experimenter herself provided a salient visual stimulus for the infant participants as well as a focus for social reference (Koenig and Echols 2003). Thus, it was important to measure infant looking to the Experimenter during the test phase as well. As we did for looking time to the object, we examined infant looking time to the Experimenter with respect to the amount of time infants spent looking at the object with
the Experimenter, and the amount of time they spent in mutual gaze with the Experimenter. Because our independent measure in this case was looking time to the Experimenter, it was not appropriate to examine infants having greater and lesser amounts of overall looking time to the Experimenter as we did above for the analyses of looking time to the object, so we have not included that analysis here.

A 2 (test trial: Same, Switch) × 2 (order: Same-first, Switch-first) × 2 (gender: female, male) repeated measures ANOVA revealed no main effect and no interactions. Infants did not, as a group, look significantly differently to the Experimenter during the Same and Switch trials. Because neither the order of the test trial presentation nor the gender of the participant had a significant effect on participants’ looking time to the Experimenter, we do not include these factors in the analyses reported below.

In two repeated measures ANOVAs, we examined the effects of shared looking with the Experimenter to the object, and of mutual gaze with the Experimenter on infant looking time to the Experimenter during test trials. Note that while the criterion for characterizing infants as having either high or low amounts of mutual gaze with the Experimenter constitutes a subset of their looking time to the Experimenter, the behavior of interest during test trials is the relative difference in looking time from one trial to the other. This difference is unaffected by whether the infant has an overall higher or lower amount of mutual gaze with the Experimenter, as measured not only during test trials, but also during pre- and posttest trials.

For shared looking to the object, a 2 (test trial: Same, Switch) × 2 (shared looking to the object: high, low) ANOVA revealed no main effect and no interaction. This indicates that the amount of time infants spent looking at the object at the same time as the Experimenter looked at the object did not differentiate their performance at test.

For mutual gaze with the Experimenter, a 2 (test trial: Same, Switch) × 2 (mutual gaze with Experimenter: high, low) ANOVA revealed no main effect, but did show an interaction (test × mutual gaze: \(F(1,14) = 10.86, p < .01, \eta^2_p = .437\)). Subsequent t-tests revealed that infants who participated in relatively more mutual gaze with the Experimenter overall, looked reliably longer to the Experimenter during the Same trial than the Switch trial, while infants who had relatively less mutual gaze with the Experimenter overall, looked significantly longer to the Experimenter in the Switch trial than they did in the Same trial (high mutual gaze: \(t(7) = 2.25, p = .059, \eta^2_p = .420\); low mutual gaze: \(t(7) = -2.70, p = .03, \eta^2_p = .511\); Figure 4). That is, infants who participated in relatively more mutual gaze with the Experimenter overall showed success in associating two novel, minimal pair word-forms with two novel objects by looking less at the Experimenter during the Switch trial. Those who participated in less mutual gaze with the Experimenter overall showed success in associating the novel word-forms and objects by looking more at the Experimenter during the Switch trial.
3.2.3. Dependent measure: Movement magnitude. Measures of movement magnitude, summed for the entire video frame, were made for each infant for each test trial. While it would have been ideal to define a ROI that contained only the infant—in order to exclude extraneous movement from the mother, for example—infants moved around within the frame too much for that to be feasible. Accurate object-tracking has not yet been incorporated into our analysis of optical flow data, and thus ROIs are static and no match for squirming infants. However, measuring motion magnitude for the entire frame did have the advantage that it kept the measures consistent across infants. The fact that a full-frame ROI may also have introduced noise renders any consistent findings that much more robust.

A 2 (test trial: Same, Switch) × 2 (order: Same-first, Switch-first) × 2 (gender: female, male) repeated measures ANOVA revealed no main effect and no interactions. Infants did not, as a group, move significantly differently during the Same and Switch trials. Because neither the order of the test trial presentation nor the gender of the participant had a significant effect on participants’ looking time to the object on the monitor, we do not include these factors in the analyses reported below. The lack of main effect of Test Trial on overall motion replicates the looking time results that showed no effect of Test Trial on looking at the object or the Experimenter for the infants as a group.

For overall looking time to the Experimenter, a 2 (test trial: Same, Switch) × 2 (overall looking to Experimenter: high, low) ANOVA revealed no main effect and no interaction. That is, whether infants looked for a relatively longer or shorter
amount of time at the Experimenter over the course of the session did not modulate their amount of movement during test.

Likewise, for shared looking to the object with the Experimenter, a 2 (test trial: Same, Switch) × 2 (shared looking to object: high, low) ANOVA revealed no main effect and no interaction. The amount of time infants spent looking at the object with the Experimenter did not differentiate their movement at test.

For mutual gaze with the Experimenter, a 2 (test trial: Same, Switch) × 2 (mutual gaze with Experimenter: high, low) ANOVA revealed no main effect, but did show a weak interaction (test × mutual gaze: $F(1,14) = 3.41, p = .086, \eta_p^2 = .189$; subsequent t-tests showed no significant effect for either group alone). This result is consistent with our looking time results; the trend was toward reliable movement differences for the same, specifically defined, grouping of infants that yielded reliable differences in looking time to the Experimenter (Figure 5).

### 3.3. Discussion: Experiment 1

#### 3.3.1. Looking time

If only measures of looking time to the object are taken into account, these results are consistent with results from previous work using recorded auditory stimuli; when measured by looking time to the object, the presence of a live source for the auditory stimulus did not enable 14-month-old infants, as a group, to succeed in associating novel, minimal pair word-forms with nonsense objects. Even the group of infants who spent relatively more time looking at the

![Figure 5. Movement magnitude measures during Same and Switch test trials for infants above the median for mutual gaze with the Experimenter (left) and infants below the median (right). Bars indicate standard deviation.](image-url)
Experimenter or looking at the object along with the Experimenter (much as in the “gaze-following” conditions of Baldwin 1991; Baldwin et al. 1996; Booth et al. 2008; Dunham et al. 1993), did not succeed. However, those infants who had participated in a greater amount of mutual gaze with the Experimenter over the course of the experiment did display successful word-object association.

The presence of the live Experimenter allowed us to identify the successful group of infants participating in higher amounts of mutual gaze with the Experimenter. But it also allowed us to examine a second measure of looking behavior during test trials, namely, looking to the Experimenter. Using this measure, again, the presence of the Experimenter alone did not support across-the-board success. But as with looking time to the object, the criterion of mutual gaze revealed a striking result; both the groups of infants participating in high and low overall mutual gaze with the Experimenter, in fact, demonstrated success in the task, but did so in opposite ways. Those in the high mutual gaze group looked significantly longer at the Experimenter during the Same trial than during the Switch trial; those in the low mutual gaze group looked longer at the Experimenter during the Switch trial than the Same trial.

When these two sets of results are taken together, a reasonable conclusion is that all infants exhibited success in associating minimal pair word-forms and novel objects; they just did it in different ways. The group of infants in the high mutual gaze group looked longer at the object during the Switch trial than the Same trial, and longer at the Experimenter during the Same trial than the Switch trial. That is, it seems that these infants tended to pay attention to the Experimenter until they encountered the unexpected Switch trial, at which point their attention was drawn back to the object, so much so that they looked significantly longer at the object, and significantly less at the Experimenter, during the Switch trial than during the Same trial. For this group of infants, the two measures may not have been completely independent; longer looking at the object was accompanied by shorter looking at the Experimenter. These two measures for the group of infants in the low mutual gaze group, on the other hand, were independent; these infants showed the same level of interest in the object whether it was presented in an expected or unexpected pairing, and regardless of changes in their looking to the Experimenter. When they experienced a switched pairing, their looking time to the Experimenter increased significantly, that is, they referenced the Experimenter significantly more when faced with an apparent discrepancy in a previously habituated word-object pairing. This behavior matches that found by Koenig and Echols (2003) for 16-month-old infants hearing an Experimenter falsely label a familiar object; these infants also showed increased looking time to the Experimenter when faced with an unexpected word-object pairing.

Crucially, then, both groups of infants, those having more and those having less mutual gaze with the Experimenter, demonstrated success in the task. The addition of the live Experimenter was critical to this demonstration in two ways. First, it allowed the identification of two groups of infants, based on mutual gaze, who
appear to have approached the task in two different ways. Second, the addition of a live Experimenter provided a new measure of looking behavior at test, i.e., looking to the Experimenter, that revealed success in both groups of infants.

3.3.2. Movement. Movement measures, conceptually and practically quite different from looking time measures, yielded identical configurations of results across all groupings analyzed in this study. As with looking time, movement measures showed no significant differences for the infants as an undifferentiated group, when infants were grouped according to whether they looked for a longer or shorter time overall at the Experimenter, and when they were grouped according to whether they spent a longer or shorter amount of time looking at the object along with the Experimenter. Movement measures did, however, suggest a difference between the group of infants who participated in mutual gaze with the Experimenter for more time, and the group with less time, just as looking time measures revealed.

An examination of results shows that infants tend to exhibit less motion when they spend more time looking at the Experimenter (Figures 4 and 5). In fact, we found a negative correlation between looking time and amount of motion; higher amounts of looking at the Experimenter correspond to lower amounts of motion ($r = -0.597, p = 0.015$). Human faces are especially effective at engaging the attention of infants (Frank et al. 2009), and our results support what makes intuitive sense; namely, that infants move less when they pay more attention to the Experimenter.

While congruent with looking time results, infant motion was not merely a byproduct of looking patterns. We found no correlation between the number of looks infants made to the object, to the Experimenter, or to both combined, and the magnitude of the movement infants exhibited during the test phase of the experiment. Thus, the amount of movement inherent in looking to and away from the object or the Experimenter did not match the global movement patterns in these data.

4. Experiment 2

How does the presence of a live Experimenter benefit the infant in this task? We begin to address this question in Experiment 2, in which a videotape of the Experimenter replaced the Experimenter herself. We know that young infants respond in some ways to a videotaped person, including following gaze, similarly to the way in which they respond to the actual person (Hains and Muir 1996; Murray and Trevarthen 1985; Senju and Csibra 2008), and that one-year-old infants can associate familiar words and objects when the words are presented by videotaped labelers (Gliga and Csibra 2009). But video does not allow for contingent interaction. Thus, if live, social interaction crucially supports infant success in
Experiment 1, then participants in Experiment 2 should not perform as well in its absence. On the other hand, it may be that infants succeed in the presence of an Experimenter by using the clear visual cues that differentiate the production of [b] (obvious opening and closing of the lips) from that of [d] (little involvement of the lips other than having them slightly open). These cues are equally available from a live or videotaped Experimenter, and thus, infants in Experiment 2 should perform similarly to those in Experiment 1.

4.1. Methods

4.1.1. Participants. Sixteen, 14-month-old infants (8 female, 6 male; range: 14 months, 1 day to 15 months, 5 days; mean: 14 months, 22 days), hearing at least 80% English in the home participated in Experiment 2. Twenty-nine additional infants participated but were excluded from analysis due to fussiness (8), experimental error (1), less than one second looking time to object at test (13), parental interference (5), lack of habituation within set parameters (1), and software malfunction (1). All participants were healthy, full-term infants, recruited from a local women’s hospital; university Behavioral Research Ethics Board guidelines for the ethical treatment of human subjects were observed.

4.1.2. Stimuli. Visual stimuli for the objects were the same as for Experiment 1, but were presented on the left side (as viewed by the infant) of a 52 inch monitor, at the same scale as in Experiment 1.

Auditory stimuli were produced by a video of the same Experimenter who participated in Experiment 1, wearing the same black sweater and producing the single word labels in a similar 15 sec long “melody” of infant-directed speech. She practiced the melody before taping so that the bin and din stimuli could be reasonably matched for pitch, tempo and amplitude.

The life-sized visual image of the Experimenter was presented on the right of the monitor (as viewed by the infant), simultaneously with the presentation of the objects. She performed the same number and timing of head turns for the bin and for the din stimuli (designed to replicate the average number and timing of head turns in the live sessions), such that she alternated between appearing to look at the infant, and turning her head, ostensibly to look at the object that appeared to her right on the large monitor.

4.1.3. Procedure. With the exception of the change from live to videotaped presentation of the auditory stimuli, the procedure was virtually the same as Experiment 1. As in Experiment 1, the attention-getting stimulus was presented in the same area of the monitor as the object, but there was no video of the Experimenter during that time. Habit (Cohen et al. 2002) was used to control the delivery of the videos of object-plus-Experimenter.
4.1.4. Data collection. Data collection was done in the same way as in Experiment 1, with the exception that the stimulus videotapes of the Experimenter were synced with those of each infant to determine shared looking to the object and mutual gaze between the infant and the (video of the) Experimenter. Infant looking to the videotaped Experimenter was determined from the lateral motion of the infant’s eyes using the same offline coding software and method used for Experiment 1; the large size of the screen made this eye movement easily discernible. As in Experiment 1, we analyzed both looking time to the object and looking time to the Experimenter.

Movement data were collected in exactly the same way as in Experiment 1.

4.2. Results

4.2.1. Dependent measure: Looking time to the object. A $2 \times 2 \times 2$ repeated measures ANOVA revealed no main effect and no interactions. Infants did not, as a group, look significantly differently to the object during the Same and Switch trials. Because neither the order of the test trial presentation nor the gender of the participant had a significant effect on participants’ looking time to the object, we do not include these factors in the analyses reported below.

In three repeated measures ANOVAs, we examined the effects of amount of overall looking to the Experimenter, shared looking to the object and mutual gaze with the Experimenter on infant looking time to the object during test trials. A $2 \times 2$ ANOVA revealed no main effect and no interaction. That is, whether infants looked for relatively longer or shorter amounts of time at the videotaped Experimenter over the course of the session did not differentiate their performance in this task.

A $2 \times 2$ ANOVA revealed no main effect and no interaction. Thus, the amount of time infants spent looking at the object at the same time that the videotaped Experimenter appeared to do so did not modulate their performance at test.

Similarly, a $2 \times 2$ ANOVA revealed no main effect, and no interaction. Participating in either more or less mutual gaze with the videotaped Experimenter also did not result in differential performance at test.

4.2.2. Dependent measure: Looking time to the Experimenter. A $2 \times 2 \times 2$ repeated measures ANOVA revealed no main effect and no interactions. Infants did not, as a group, look significantly differently to the Experimenter during the Same and Switch trials. Because neither the order of the test trial presentation nor the gender of the participant had a significant effect on participants’ looking time
to the Experimenter, we do not include these factors in the analyses reported below.

In two repeated measures ANOVAs, we examined the effect of shared looking to the object and of overall mutual gaze with the Experimenter on infant looking time to the Experimenter during test trials. A 2 (test trial: Same, Switch) × 2 (total time in shared looking to object: high, low) ANOVA revealed no main effect and no interaction. This indicates that the amount of time infants spent looking at the object at the same time as the videotaped Experimenter appeared to do so did not differentiate their performance at test.

Similarly, a 2 (test trial: Same, Switch) × 2 (mutual gaze with Experimenter: high, low) ANOVA revealed no main effect, and no interaction. Thus, differing levels of mutual gaze with the videotaped Experimenter did not modulate infant looking behavior during test.

4.2.3. Dependent measure: Movement magnitude. A 2 (test trial: Same, Switch) × 2 (order: Same-first, Switch-first) × 2 (gender: female, male) repeated measures ANOVA revealed no main effect and no interaction with order. There was a weak interaction between test and gender (test × gender: $F(1,12) = 4.005$, $p = .069$, $\eta^2_p = .25$). Subsequent paired t-tests revealed no main effect of test for females, but a trend for males (test trial: $t(1,7) = -2.058$, $p = .079$, $\eta^2_p = .377$) to move less during the Same trial than during the Switch trial (Figure 6). Because order had no effect on the results, we do not include it in the analyses below.

For overall looking time to the Experimenter, a 2 (test trial: Same, Switch) × 2 (overall looking to Experimenter: high, low) × 2 (gender: male, female) ANOVA revealed no main effect and no interactions. Whether infants looked for a relatively
longer or shorter amount of time at the Experimenter over the course of the session did not modulate their amount of movement during test.

Likewise, for shared looking to the object with the Experimenter, a 2 (test trial: Same, Switch) × 2 (shared looking to object: high, low) × 2 (gender: male, female) ANOVA revealed no main effect and a weak interaction between test and gender, as above ($F(1,12) = 4.028$, $p = .068$, $\eta^2_p = .25$). Whether infants looked for a relatively longer or shorter amount of time at the object along with the Experimenter did not modulate their amount of movement during test.

For mutual gaze with the Experimenter, a 2 (test trial: Same, Switch) × 2 (mutual gaze with Experimenter: high, low) × 2 (gender: male, female) ANOVA revealed no main effect and no interactions. Whether infants looked for a relatively longer or shorter amount of time at the Experimenter while the Experimenter looked at them did not modulate their amount of movement during test.

As in Experiment 1, measures of movement are not correlated with looks either to the object, to the Experimenter, or to both combined at test. Unlike in Experiment 1, however, there was also no correlation between the amount of movement infants exhibited at test and the amount of time they spent looking at the (videotaped) Experimenter.

4.3. Discussion: Experiment 2

Against the background of the results in Experiment 1, the fact that 14-month-old infants failed to demonstrate minimal pair word-novel object associations by every looking time measure examined in Experiment 2 is telling. It suggests that the videotaped Experimenter did not provide adequate scaffolding for the infants to succeed. Similarly, the infants listening to Koenig and Echols’ (2003) false labeling events looked longer at the labeling source only when that source was a human gazing at the object, not when an audio speaker with or without an accompanying silent human, or when a human facing away from the object provided the labels. Further, the addition of visual cues to infants’ differentiation of [b] and [d] is likely not the source of their success in Experiment 1; those visual cues were just as present, after all, in the video stimuli in Experiment 2. The primary difference between Experiment 1 and Experiment 2 was whether the infants were viewing a live or videotaped Experimenter.

While looking time measures failed to reveal success for any group of infants, motion measures showed that males tended to move less during the Same trial than during the Switch trial. Because there was no effect of order, and the order of Same and Switch trials was counterbalanced across gender, this could not have been a simple matter of fatigue or fussiness. These results indicate that movement measures are capable of revealing differences in behavior that looking time measures do not capture; in this case, movement is a significant metric for male, but not female, behavior in the task. The lower variability (SD) for male movement, and the greater stability for the variability in female movement bear out this observation.
L. Fais et al.

(for males: SD Same = 22.3, SD Switch = 37.0; for females: SD Same = 79.8, Switch = 79.8).

5. A comparison of Experiments 1 and 2

5.1. Results

It is informative to compare infant behavior in the two experiments. Infants spent a longer overall amount of time looking at the videotaped Experimenter of Experiment 2 than at the live Experimenter in Experiment 1 ($t(30) = -3.42, p = .002, r^2 = .28$), as measured in the offline pre- and posttest and test trial coding. Further, infants participated in more mutual gaze with the videotaped Experimenter of Experiment 2 ($t(30) = -3.86, p = .001, r^2 = .33$). On the other hand, infants spent a longer amount of time in shared looking to the object with the live Experimenter in Experiment 1 than with the videotaped Experimenter in Experiment 2 ($t(30) = 6.32, p < .001; r^2 = .57$, Figure 7). However, two, 2 × 2 repeated measures ANOVAs showed no differences in the amounts of time during test trials that infants in the two experiments looked at the object or the Experimenter.

Likewise, a 2 × 2 ANOVA showed no differences in the amount of movement exhibited by infants in the test trials of the two experiments. Thus, overall movement levels were comparable for the two experiments.

<table>
<thead>
<tr>
<th>Looking time (sec)</th>
<th>Live Experimenter (Exp.1)</th>
<th>Video Experimenter (Exp. 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall looking at Exp</td>
<td>$p = .002$</td>
<td></td>
</tr>
<tr>
<td>Mutual gaze with Exp</td>
<td>$p = .001$</td>
<td></td>
</tr>
<tr>
<td>Shared looking to object with Exp</td>
<td>$p &lt; .001$</td>
<td></td>
</tr>
</tbody>
</table>

![Figure 7](image-url)  
Looking times for overall looking to Experimenter (left), mutual gaze with Experimenter (center), and shared looking to object with Experimenter (right) for both Experiments 1 and 2. Bars indicate standard deviation.
5.2. Discussion

It is interesting that infants spent a greater amount of time looking at the videotaped Experimenter in Experiment 2 than the live Experimenter in Experiment 1. Given the interest that infants show in faces (Frank et al. 2009), the participants would have found the videotaped, talking Experimenter to be more highly engaging than the silent object that appeared on the same screen, resulting in relatively long looking times to the videotaped Experimenter (see Robinson et al. 2005, for results supporting this suggestion for both 8- and 14-month-old infants). Longer overall looking at the videotaped Experimenter had as a natural consequence a greater possibility for participating in mutual gaze with the videotaped Experimenter in Experiment 2 than the live Experimenter in Experiment 1, especially since the videotaped Experimenter’s shifts to look at the object were predictable. On the other hand, infants in Experiment 1 spent more time in shared looking to the object with the live Experimenter than they did in Experiment 2 with the videotaped Experimenter. We suggest that infants regarded the live, but not the videotaped, Experimenter as a compelling partner in an interaction that involved looking at the object on the monitor, resulting in a significantly higher amount of shared looking to the object in Experiment 1. Notice that even the predictability inherent in the head movements of the videotaped Experimenter (looking toward the infant and then looking toward the object) did not support greater amounts of shared looking to the object in Experiment 2.

These differences in overall looking patterns might have been coupled with greater or lesser interest in the object or Experimenter during the test phases of the two experiments. However, this was not the case; infants spent equivalent amounts of time looking at the object and the Experimenter in the test trials of the two experiments, and moved equivalent amounts as well. This similarity gives us even greater confidence in comparing the results of the two experiments, since it implies that the different looking and movement patterns found in the test trials are not simply the result of differences in general levels of interest in the stimuli.

6. Experiment 3

We conducted Experiment 3 as a check on the consistency of the presentation of test stimuli by the Experimenter in Experiment 1. While the use of videotapes assured us of complete uniformity of stimuli presentation in Experiment 2, the Experimenter in Experiment 1, despite best efforts, might have unknowingly cued the infants as to the “special” status of the stimuli in the Switch trials. It is highly unlikely that cuing was present and that infants still showed a more robust recovery of looking time to the object when mutual gaze was taken into account than when overall gaze was considered. Nevertheless, it remained a possibility that needed to be examined; thus, we investigated adults’ perceptions of the test trials.
in Experiment 1 to explore whether they could perceive cues to the status (Same or Switch) of the test trials.

6.1. Methods

6.1.1. Participants. Sixteen participants (8 male, 8 female) were tested, age range 18 yr, 4 months, 24 days to 22 yr, 8 months, 12 days; mean age 20 yr, 1 months, 6 days. They were recruited through the participant pool of the university Department of Psychology, and were compensated with academic credit. University Behavioral Research Ethics Board guidelines for the ethical treatment of human subjects were observed.

6.1.2. Stimuli. The videos taken of the Experimenter’s face during all test trials from Experiment 1 were excised into separate video clips, for a total of 32 (one Same trial and one Switch trial for each of the 16 participants), 15 sec videos. Each clip of the Experimenter’s face, then, was paired with the video of the object stimulus that was presented for that trial such that the object appeared on left side of the screen and the video of the Experimenter’s face as she named the object appeared on the right side (just as they were viewed by the infant participants).

6.1.3. Procedure. Each adult participant was tested in a sound-attenuated room. They were told that their goal was to determine if the object shown in each video clip had been named “correctly” by the Experimenter in the video. The set-up of the original experiment was described, but not its purpose. The procedure was programmed in PsyScope (Cohen et al. 1993); the clips were displayed on a laptop screen, one-by-one, to the participants, who pressed “1” if they thought the label was correct and “0” if not. The participants were reminded of the meaning for each press (“1” or “0”) before the presentation of each clip, and were debriefed as to the purpose of the study after completion.

6.1.4. Data collection. We tallied how many times each participant was correct about whether or not the Experimenter had given the object its “correct” name.

6.2. Results

If the Experimenter had produced cues as to whether the object was being named with its “correct,” i.e., habituated label, and if the adult participants were able to access these cues, the number of correct responses by the participants should have been greater than chance. Participants made correct judgments 53.3% of the time; their answers did not differ from chance.
6.3. Discussion

We cannot guarantee that infants and adults are sensitive to the same sorts of cues, and thus that this result demonstrates conclusively that there were no salient cues present in the test stimuli that the infants were responding to. However, this result, coupled with the fact that the total amount of looking time to the Experimenter was not the factor that revealed success in the task, helps rule out the possibility that infants were responding to task-irrelevant aspects of the test stimuli presentation.

7. General discussion

This set of experiments broadens our understanding of the ability of 14-month-old infants to associate novel, minimal pair word-forms with novel objects, and has allowed us to explore the use of measures of movement in infant language acquisition tasks. Both looking time and movement measures indicate that infants’ success in this task is supported by the presence of a live interactor. All infants demonstrated success in the task, mediated by their level of mutual gaze with the Experimenter. Interestingly, simply looking longer at the live Experimenter did not support infants’ demonstration of success, nor did any of the looking behaviors we examined in the case of the videotaped Experimenter. Further, the benefit accrued from interacting with a live Experimenter was realized even though infants actually spent a longer amount of time looking at and participating in mutual gaze with the videotaped Experimenter. These results demonstrate that infant word-object association can be crucially enabled by a live, social source. Mani and Plunkett (2008) showed that 14-month-old infants could use minimal differences in vowel sounds in novel words to recognize habituated word-object pairings; however, in their work, infants learned these associations through richly interactive play, and habituation to the labels delivered in sentential contexts. Our work shows that this level of social embedding is not necessary, at least not for consonants; the 14-month-old infants in our study learned the association simply with the support of the possibility for referencing a live interactor.

Interestingly, movement measures indicate that males may be succeeding in the task in the presence of the videotaped Experimenter as well. Because this result does not match results from looking time measures, and because overall infant movement in the presence of the videotaped Experimenter does not correlate with looking time to the Experimenter (as it did for the live Experimenter), we speculate that movement measures in this case are capturing behavior stemming not from attention, but perhaps from agitation at the violation of the pairing. Clearly, this is an area in need of continued investigation. We are currently examining movement in infants participating in one of the early minimal pair Switch studies (Pater et al. 2004) for enlightening comparison to the results found in the current study.
It is tempting to interpret our results as showing that greater amounts of mutual gaze with a live Experimenter lead to success in the minimal pair word-object association Switch task for 14-month-old infants. This conclusion is only true when looking time to the object is taken into account. Infants who participated in lower amounts of mutual gaze showed success as well, by looking longer at the Experimenter during the Switch test trial. Thus, we cannot conclude that some sufficient amount of mutual gaze is the critical factor for success across the board; rather, the measure of mutual gaze is the parameter by which we identified two groups of infants who both demonstrate success in the presence of a live, but not videotaped, Experimenter: the group participating in a relatively high amount of mutual gaze showed success in the criterial manner for the Switch task (greater looking time to the object during the Switch trial) as well as in their looking behavior toward the Experimenter (significantly less looking to the Experimenter during Switch), and the group having relatively less mutual gaze demonstrated success by virtue of their increased looking at the Experimenter during the Switch trial. Further, movement measures corroborate these two different patterns; across all infants, looking at the Experimenter during test was inversely correlated with movement. Thus, the high mutual gaze infants who showed less looking to the Experimenter during the Switch trial, also showed greater movement during the Switch trial. Conversely, the low mutual gaze infants who showed greater looking to the Experimenter during the Switch trial, showed less movement during the Switch trial.

The different patterns of looking behavior for these two groups of infants may contain clues as to how each approached the task, and how, ultimately, these different behaviors both led to the demonstration of success. We suggest that 14-month-old infants, just as they have difficulty using fine phonetic information in word learning without help from familiarity or referential context, may also be at the early stages of their ability to utilize social information in word learning contexts.

We know that infants do not learn equally well from interacting and non-interacting humans in labeling tasks (e.g., Baldwin 1991; Baldwin et al. 1996; Koenig and Echols 2003), consistent with our finding that the availability of an interacting labeler supports word-object association. Csibra and Gergely (2006; 2009) theorize that a host of early preferences and sensitivities conspire to make infants receptive to communicative signals from con-specifics; they propose that preferences for human speech over non-speech (Vouloumanos and Werker 2004), for faces (Johnson and Morton 1991) and for infant-directed speech (Cooper and Aslin 1990), as well as sensitivity to eye contact (Batki et al. 2000) and contingency of vocalization (Murray and Trevarthen 1985) all contribute to infants’ readiness to learn from interactive contexts. Baldwin and Moses (1996) suggest that over the second year of life, infants develop gradually from being “good information consumers,” who are able to utilize the cues provided by an information source in the form of eye contact, gaze direction, posture, infant-directed intona-
tion, etc., in understanding information, to being good “information seekers” (p. 1931), who are able to actively direct their attention to relevant and useful aspects of human sources of information. This development may be paralleled by the shift from “passive joint engagement,” during which an infant attends to an object concurrently with an interactor, but pays little attention to the interactor, to “coordinated joint interaction,” in which the infant also references the interactor, documented by Bakeman and Adamson (1984) from 6 to 18 months of age. They observed a particularly sharp increase in coordinated joint interaction between the ages of 12 and 15 months, encompassing the age of the infants in this study. The correlation of movement and looking to the Experimenter supports the idea that the infants in Experiment 1 were good “information consumers;” their movement patterns differed when they were looking at an object or to the Experimenter. The greater stillness when looking at the Experimenter suggests that the infants were paying attention to the live information source in the task in a way that differed from their looking at the object.

We propose that the group of infants in our study who participated in greater amounts of mutual gaze with the Experimenter were acting on their preparedness for receiving communicative information, “practicing” using eye contact to reinforce their take-up of the information in the infant-directed vocalizations of the Experimenter during “coordinated joint interaction.” But they had not yet become good “information seekers;” their interest in the human as a source of information extended only to their association of humans with labeling events. Thus, their surprise at the Switch trial was surprise at the difference in the pairing of label and object, which they exhibited by their looking longer to the object (and thus, incidentally, by looking less at the Experimenter).

By contrast, the group of infants who participated in relatively less mutual gaze with the Experimenter were, in fact, good “seekers of information.” We suggest that their lower level of mutual gaze across the experiment is an indication of their greater efficiency in taking up information from the Experimenter. In the Switch trial, their looking to the object was unaffected. However, like the slightly older infants in Koenig and Echols’ study (2003), these infants significantly referenced the Experimenter, seeking information about the unexpected pairing from a communicatively accessible source, i.e., the human interactor.

7.1. Alternative explanations

7.1.1. Visual language processing. It could be argued that what allowed infants to succeed in the context of the live, but not the videotaped, Experimenter was the availability of three-dimensional, visual language information on the face of the live Experimenter. No work has been done comparing the relative availability of visual language information to infants in two vs. three dimensions; however, we do know from studies of infants’ visual language processing abilities utilizing two-
dimensional videos that infants are able to make use of visual language cues at very young ages, for example, to discriminate languages at 4 and 6 months (Weikum et al. 2007), to match a lip shape with a heard vowel sound at 4 (Kuhl and Meltzoff 1982) and even 2 months of age (Patterson and Werker 2003), and to match heard running speech to the appropriate face (Spelke and Cortelyou 1981). In light of these early abilities of young infants, demonstrated using two-dimensional video, it is unlikely that 14-month-old infants are only able to utilize visual language information from the three-dimensional, live Experimenter in Experiment 1, and not from the videotaped Experimenter in Experiment 2.

7.1.2. **Stimuli variability.** Rost and McMurray (2009) claimed that the variability of the acoustic information they provided to the 14-month-old infants in their study enabled those infants to form the more robust understanding of the [b]-[d] distinction necessary to succeed in the task. Certainly the infants in Experiment 1 received more variable auditory – and visual – stimuli than those in Experiment 2, by virtue of the fact that each token heard throughout the infant’s session was different from every other token in Experiment 1, while only the tokens within each trial differed from one another in Experiment 2. It might be argued that it was simply the range of variability of the stimuli presented in Experiment 1 that supported success in our task as well, even though the intra-speaker differences of one speaker are likely far less variable than the inter-speaker differences exhibited by the 18 speakers in the Rost and McMurray study. While this suggestion cannot be ruled out, it does not provide an explanation for the different patterning of looking behavior evidenced by the high- and low-mutual gaze groups, nor for the differences in how each group evidenced success in the task.

7.1.3. **Attention.** Kuhl and colleagues found that infants learning from a live source had a higher observer rating for attention than those learning from audio-visual or audio-only input (Kuhl et al. 2003), and credited the social cues available from a live interactor with motivating more successful language learning. Measuring “attention” by the amount of time infants spent looking at or participating in mutual gaze with the Experimenter, infants in our studies actually paid more attention to the videotaped Experimenter in Experiment 2 than to the live Experimenter in Experiment 1 (Figure 4). This would predict that infants in Experiment 2 would succeed, and those in Experiment 1 would not; in fact, the opposite was the case.

On the other hand, the movement data suggest that attention, as determined by greater stillness when looking at the live, but not the videotaped Experimenter, might still be an important factor. Salley et al. (2010) showed that 14-month-old infants’ response to joint attention, but not their overall attention to a videotaped face, was a significant predictor of their performance on a two-object Switch task using novel, non-minimal pair word-forms. This finding is consistent with our results, and suggests that mutual gaze with the live Experimenter was particularly effective in focusing infant attention, leading to success in the task (Samuelson and
Smith 1998). However, recall that even infants who spent less time participating in mutual gaze with the live Experimenter also evidenced success in the task. This suggests that attention alone cannot fully explain our results.

8. The novel contribution of movement measures

Decades of research using looking time has informed our interpretation of those results (see for example, Aslin 2007; Cohen 2004), but we have just begun to explore the intricate relationships among looking, movement, and attention, and understanding the insights that movement provides into infant language acquisition requires extensive further study. But this work allows us to take a few first steps by linking global infant movement with level of attention-to-a-person, and demonstrating that such movement is not simply the outcome of shifting direction of gaze. The fact that infant movement patterns so closely mirror looking time results in such specific ways is additional confirmation of the usefulness and meaningfulness of measures of movement. In addition, we have also seen in the case of the males in the videotaped Experimenter study that movement measures can be revealing of behaviors that looking time measures are not sensitive to.

We have begun to explore other ways in which movement measures provide additional new insights into infant response to stimuli. Looking time measures (especially those that take into account averaged values for a small set of test trials) are essentially one-dimensional; they are interpreted as measuring infant attention to some (changed) feature of the experiment context. Movement measures, on the other hand, have the potential to yield values in three spatial dimensions. Optical flow from video incorporates two in this study: the horizontal and the vertical (use of optical flow to assess looming would constitute a third). The analysis of these dimensions allows us to examine infant orientation toward or away from the Experimenter on the infant’s right (horizontal, or x-axis motion), and infant postural response, e.g., slumping (vertical, or y-axis motion).

We examined vertical and horizontal infant movement in all three of the groupings described above. We found different patterns of horizontal movement depending upon how much infants looked at the object on the monitor at the same time that the Experimenter did. A $2 \times 2$ ANOVA for test trial (Same, Switch) and shared looking at object (more, less) yielded no main effect, but showed a trend toward a significant interaction ($F(1,14) = 3.736, p = .074, \eta^2_p = .21$). Subsequent independent samples t-tests revealed that all infants engaged in similar horizontal motion during the Same test trial, moving slightly away from the Experimenter. However, during the Switch trial, the two groups behaved differently: those who had engaged in more shared looking at the object moved horizontally away from the Experimenter; those who engaged in less, moved toward the Experimenter ($t(14) = 4.276, p = .03, r^2 = .29$; Figure 8). The former group may have been “checking” the object, just as those infants with greater mutual gaze with the Experimenter looked...
longer at the object during the Switch trial; the latter group may have been referencing the Experimenter, just as we suggested above for the infants with low mutual gaze with the Experimenter. Directional movement information, e.g., whether infants are withdrawing from or approaching the Experimenter, thus provides a new dimension to our understanding of infant response in experimental settings.

9. Conclusion

Fourteen-month-old infants crucially benefited from the presence of a live Experimenter in demonstrating success in associating novel, minimal pair word-forms and novel objects, a task at which they are unable to exhibit success when tested in an unaided experimental setting (i.e., as in Stager and Werker [1997]). The fact that a measure of mutual gaze as well as of movement revealed success for both those infants participating in more and those participating in less mutual less gaze with the Experimenter argues that the opportunity to engage in reciprocal attention with a labeling interactor is important for infants in a word learning context. These results are in sharp contrast to the lack of comparable results for any other looking time or movement measures.

This work also demonstrates for the first time the application of a video-based algorithmic analysis yielding continuous measures of the timing, magnitude and
direction of movement to research into infant language acquisition. The fact that measures of motion can be extracted completely non-invasively, and automatically, simply from the video routinely recorded during laboratory studies of infants, addresses two crucial problems faced by researchers who want to investigate movement but have been, until now, limited to using hand coding methods: the extraordinary labor- and thus time-intensiveness of the enterprise, and the introduction of subjective error during hand-coding. A comparison of the optical flow analysis with that of looking time confirmed the validity of these movement measures as indicators of meaningful patterns of infant response to language stimuli. Further, movement measures suggested explanations for looking time behaviors, as well as uncovering meaningful behaviors that were not captured by looking time measures. This outcome is encouraging for the use of motion as a meaningful dependent measure in infant language research. The addition of movement to our repertoire of measures in developmental research can reveal fresh insights into the cognitive processing of infants in language learning tasks, and enrich the dimensionality of our understanding of infant language acquisition.

Correspondence e-mail address: lfais@psych.ubc.ca

References


