Pupillometry showcases the benefits of motionese for infants’ processing of dynamic human action

Jessica E. Kosie1,2 & Dare A. Baldwin2

1 Department of Psychology, Princeton University

2 Department of Psychology, University of Oregon

Author note

Correspondence concerning this article should be addressed to Jessica E. Kosie, Department of Psychology, Princeton University, Princeton, NJ 08540 USA. E-mail: [jkosie@princeton.edu](mailto:jkosie@princeton.edu)

Abstract

Infants rapidly acquire skill in making sense of dynamic human action: they not only recognize many actions, but predict, respond to, remember, and learn from them. Such action-processing fluency depends on finding structure within unfolding activity – in particular, boundaries at which units of action begin and end. However, action unfolds rapidly and is just as quickly gone. How do infants find boundaries in the complex, fleeting stream of motion that they actually observe? Caregivers’ infant-directed action demonstrations might assist with this challenging task. In interactions with infants, caregivers modify their motion in a variety of ways that engage infants’ overall attention (i.e., “motionese;” Brand, Baldwin, & Ashburn, 2002). It seems plausible that these modifications promote infants’ detection of internal structure within action. We harnessed pupillometry to investigate this possibility, comparing pupil dilation patterns when infants viewed a given activity stream in motionese versus adult-directed formats. Infants’ pupil size (an indication of attention or cognitive engagement) increased in response to action boundaries, but only for motionese demonstrations. Thus, in addition to engaging overall attention, motionese facilitates infants’ detection of action’s internal structure. These findings showcase the benefits of both motionese for infants’ action processing, and pupillometry for shining light on key developmental phenomena.

*Keywords:* action segmentation, action processing, event processing, motionese, pupillometry, infancy

*Word count:* X

Pupillometry showcases the benefits of motionese for infants’ processing of dynamic human action

# Introduction

Human activity generates a motion stream that is both complex and rapidly unfolding. Making sense of this dynamically streaming sensory information is a challenging cognitive enterprise; actions must be discerned “on the fly” as information flows past. The ability to find structure within unfolding activity (i.e., where individual units of action begin and end) is a key skill that is linked to fluency across domains including learning (Bailey, Kurby, Giovanetti, & Zacks, 2013), memory (Sonne, Kingo, and Krøjgaard, 2016, 2017; Sargent et al., 2013; Flores, Bailey, Eisenberg, & Zacks, 2017), social understanding (Zalla, Labruyére, & Georgieff, 2013), and language acquisition (Levine, Buchsbaum, Hirsh-Pasek, & Golinkoff, 2018). Early in life, infants seem to have acquired the ability to find structure in at least some kinds of activity sequences (see Levine et al., 2018 for a review). Less is known about how infants rise to the challenge of finding this structure as they first encounter novel action and watch it rapidly unfold over time.

It is important to consider, however, that infants are often not alone as they face this challenge. For example, in interactions with infants, caregivers modify their behaviors in a variety of ways that engage infants’ attention and facilitate learning (e.g., Brand, Baldwin, & Ashburn, 2002; Fernald, 1985; Csibra & Gergely, 2009). It seems likely that caregivers’ modifications to infant-directed action (e.g., “motionese;” Brand et al., 2002) could serve specifically to help infants find structure, such as action boundaries, as activity unfolds. As yet this hypothesis remains untested, in part because methods with which to measure infants’ online processing of streaming visual information have rarely been implemented in the action domain. However, the recent development of a novel, open-source, inexpensive, and infant-friendly system for measuring infants’ pupillary response to cognitive stimuli (the Stimulus-Induced Pupil Response (SIPR) system; Patent Pending; Bala, Keller, Whitchurch, Baldwin, & Takahashi, 2016) provides a methodology with which to explore infants’ online processing of visual information. In the present research, we used the SIPR system to explore the extent to which motionese influences infants’ ability to detect action boundaries as activity unfolds across time.

In what follows, we summarize what is currently known about infants’ processing of dynamically unfolding activity as well as initial evidence that caregivers scaffold infants’ processing of human action. Finally, we describe a body of evidence indicating that pupillometry offers potential insight into infants’ processing of dynamically unfolding activity.

## Action processing in infancy

A growing body of literature suggests that action segmentation processes are operative early in life (see Levine et al., 2018 for a recent review). In particular, infants display sensitivity to boundaries in a variety of everyday intentional action sequences (Baldwin, Baird, Saylor, & Clark, 2001; Hespos & Saylor & Grossman, 2009; Saylor, Baldwin, Baird & LaBounty, 2007; Hespos, Grossman, & Saylor, 2010). For example, in seminal work on action segmentation in infancy, Baldwin and colleagues (2001) familiarized 10- and 11-month-old infants to a video depicting an actor engaging in a series of everyday activities. At test, infants were shown the same videos with pauses at action boundaries (i.e., the initiation and completion of intentional action units – like the moment at which one grasps an object to pick it up) or at non-boundary junctures. Infants looked longer to test videos that depicted pauses at non-boundary junctures, suggesting that they readily detect structure in unfolding intentional action, parse human behavior with respect to this structure, and are surprised when this structure is violated.

Recently, Sonne, Kingo, and Krøjgaard (2016) demonstrated that older infants’ memory is influenced by the availability of information about action boundaries, extending findings from previous studies with adults (e.g., Swallow, Zacks, Abrams, 2009; Radvansky & Zacks, 2017; Gold, Flores, & Zacks, 2017). In their research, 16- to 20-month-old infants who had been presented with dynamic activity featuring occlusions at boundaries had weaker memory for the activity, two weeks later, than infants for whom non-boundary regions had been occluded. In an extension of this work, Sonne, Kingo, and Krøjgaard (2017) additionally demonstrated that, at a delay of ten minutes after viewing, 21-month-old infants more accurately remembered specific objects presented at action boundaries than those presented at non-boundary junctures. Apparently infants, like adults, selectively attend to boundaries within unfolding activity. An open question, however, entails just how infants begin to find action boundaries within dynamically unfolding activity sequences.

Statistical learning is one mechanism that seems likely to facilitate infants’ detection of such structure in action. Several lines of evidence suggest that infants use knowledge of predictability structure to guide processing of unfolding activity (e.g., Ambrosini et al., 2013; Kanakogi & Itakura, 2011; Monroy, Gerson, & Hunnius, 2017, NAGAI PREDICTION ERROR REVIEW?). For example, Monroy and colleagues (2017) found that infants not only detected statistical regularities in novel, continuous action sequences, but used these regularities to guide anticipatory gaze on later occasions on viewing the same sequences. As well, infants can capitalize on statistical regularities embedded within novel activity sequences to demarcate units within the unfolding stream (Baldwin, 2012; Roseberry, Richie, Hirsh-Pasek, Golinkoff, & Shipley, 2011; Stahl, Roseberry, Hirsh-Pasek, Romberg, & Golinkoff, 2014). For example, 7- to 9-month-old infants viewed videos of hand movements (Roseberry et al., 2011) or an animated agent performing action sequences (Stahl et al., 2014). As in previous work with adults (e.g., Baldwin, Andersson, Saffran, & Meyer, 2008; Hard, Meyer, & Baldwin, 2018; Swallow & Zacks, XXXX), infants readily discovered units of action within the streaming activity, relying on low-predictability transitions within the stream to define boundaries between action units.

In sum, infants seem to be sensitive to the internal structure of at least some kinds of everyday activity, and their enhanced memory for activity occurring at boundaries suggests that their processing privileges these regions. Statistical learning is one likely mechanism that enables infants to discover the structure of action over time. However, it is unclear as yet how much or what kind of repeated exposure is necessary before the statistics of a novel activity sequence can be learned and used to guide infants’ subsequent processing. In infants’ day-to-day experience, some contexts might serve to enhance these statistics, promoting infants’ identification of attention-worthy regions of activity (i.e., action boundaries), thereby supporting infants’ rapid acquisition of action processing skill. One particular context that might be especially influential in this regard occurs when caregivers specifically attempt to demonstrate novel activities to infants.

## “Motionese” might scaffold infants’ detection of structure in action.

When demonstrating novel action to infants, caregivers modify their behavior in a variety of systematic ways that seem well suited to promoting infants’ processing of the dynamic activity stream. Recent research investigating this phenomenon provides initial confirmation that infants benefit from such “motionese” demonstrations. It remains unclear, however, whether motionese specifically scaffolds infants’ detection of structure within action, although this seems highly plausible.

A first study documenting motionese found that, when demonstrating novel objects to 6- to 13-month-old infants, caregivers exhibited increased interactiveness, proximity to their infant interactive partner, enthusiasm, range of motion, repetition, and simplicity in their actions (Brand et al., 2002). These modifications capture infants’ attention, in that infants prefer to watch action demonstrations in a motionese format over action characteristic of demonstrations directed toward adults (Brand & Shallcross, 2008). Toddlers are also more likely to imitate actions demonstrated using motionese (Baldwin, Myhr, & Brand, in preparation; Williamson & Brand, 2014), and use of motionese increases 8- to 10-month-old infants’ subsequent object exploration, which can have downstream benefits for overall learning (Koterba & Iverson, 2009).

The motionese modifications just summarized parallel modifications in language directed to infants, commonly called “motherese” (Snow & Ferguson, 1977), and are likely part of a suite of infant-directed modifications jointly constituting a natural pedagogy phenomenon that has received extensive investigation in the developmental literature (Sage & Baldwin, 2010; Csibra & Gergely, 2009). Benefits of motherese in speech include facilitating infants’ attention (Fernald, 1985; ManyBabies, 2020), with subsidiary benefits such as enhancing infants’ processing of the acoustic and segmental properties of speech (Kuhl, 2004), and promoting structure detection within streams of fluent speech (Thiessen, Hill, & Saffran, 2005; Kemler-Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989).

Given such findings regarding motherese, it seems highly plausible that motionese analogously promotes infants’ detection of structure within activity. In fact, there is existing evidence that certain features of motionese could serve to specifically direct infants’ attention to action boundaries. For example, during object demonstrations to their 7- to 12-month-old infant, mothers’ infant-directed gaze is systematically aligned with boundary junctures (Brand, Hollenbeck, & Kominsky, 2013). Features of mothers’ infant-directed speech during action demonstrations is often aligned with action boundaries as well. For example, the onset and offset of mothers’ action-describing speech tends to be aligned with boundaries occurring at the initiation or completion of an action unit (Meyer, Hard, Brand, McGarvey, & Baldwin, 2013; Hirsh-Pasek & Golinkoff, 1996), and infants tend to group such packaged action into coherent “chunks” (Brand & Tapscott, 2007). At action boundaries, mothers also tend to speak with rising or falling intonation, perhaps signaling the completion of an action unit (Rohlfing, Fritsch, Wrede, & Jungmann, 2006). Features such as repetition (Brand, McGee, Kominsky, Briggs, Gruneisen, & Orbach, 2009) and turn taking (Brand, Shallcross, Sabatos, & Massie, 2007) in infant-directed demonstrations coincide systematically with action boundaries, and may additionally serve to facilitate infants’ attention to the segmental structure of unfolding activity.

All in all, current evidence strongly suggests that motionese may assist infants in detecting action boundaries within continuously flowing activity, which would facilitate learning. For adults, finding structure in activity has downstream benefits for comprehending action (Newtson & Engquist, 1976; Zacks, Tversky, & Iyer, 2001), remembering what has occurred (Sonne et al., 2016, 2017; Swallow et al., 2009), and performing actions oneself (Bailey et al., 2013). Infants’ skill at detecting action boundaries would also promote social understanding (Baldwin & Baird, 2001; Zalla et al., 2013) and language learning (Baldwin, 2005; Levine et al., 2018). [MOVE PAST 2 SENTENCES ELSEWHERE?] As yet, however, the possibility that motionese scaffolds infants’ detection of boundaries within streaming activity has not been put to direct test, largely because most methods used to investigate infant perception and cognition are not well-suited to the task. In particular, existing techniques for investigating infants’ action processing have typically been limited to first exposing infants to action sequences and then, at later test, measuring infants’ recognition/discrimination with respect to the stimuli that they previously viewed (e.g., Woodward, 1998; Baldwin et al., 2001; Stahl et al., 2014). Fortunately, a relatively new technique – measuring ongoing involuntary changes in pupil diameter concomitant with cognitive engagement – offers a novel approach to exploring issues related to infants’ processing of unfolding action. This technique thus offers a novel window on ways in which motionese may scaffold such processing.

## Pupillometry as a promising solution

Pupil dilation response (hereafter PDR) occurs spontaneously with changes in luminance (Loewenfeld, 1993) as well as in response to a variety of cognitive stimuli (Goldwater, 1972; Sirois & Brisson, 2014; Laeng, Sirois, & Gredebäck, 2012). Among other things, changes in pupil dilation are thought to reflect the attentional demands imposed by a cognitive task (Beatty & Lucero-Wagoner, 2000; Goldinger & Papesh, 2012). In general, available findings provide strong confirmation of Kahneman’s (1973) suggestion that pupil diameter provides an online indication of the “intensity of attention” being allocated by an observer.

Of particular relevance to the present research, one recent study confirms that adults display systematic pupil dilation in relation to the internal structure of activity sequences. Specifically, Tanaka, Kosie, & Baldwin (in preparation) presented adults with a series of short clips of sport activities, each containing one coarse-level action boundary (e.g., when the athlete completed their primary goal, such as striking a tennis ball with a racket during a serve). As predicted, systematic changes in pupil diameter occurred in relation to action boundaries. PDR tended to systematically increase immediately prior to action boundaries, peak at or shortly after boundaries, and return toward baseline over an extended period thereafter. This pattern of response indicates that the PDR methodology offers a window on viewers’ detection of segmental structure within dynamic activity as processing is underway.

Infancy researchers are also increasingly taking advantage of pupillometry to investigate infant cognition; infants’ PDR profiles have shown a range of similarities to those of adults’ (e.g., Addyman, Rocha, & Mareschal, 2014; Jackson & Sirois, 2009; Sirois & Jackson, 2012; Gredebäck & Melinder, 2010; Hepach & Westermann, 2016).

Based on this collection of findings, we opted to employ pupillometry to test for boundary-related PDRs as infants viewed novel activity sequences in motionese versus adult-directed formats. We recruited infants in the 9-12 month age range, because prior research has documented both that caregivers produce motionese to infants of this age, and infants in the 9-12 month range are also known to be sensitive to structure in at least some everyday activity streams. We addressed three main questions; these questions and attendant hypotheses and predictions follow: (1) To what extent does infants’ previously-observed preference for “motionese” over adult-directed action replicate via pupillometry (as opposed to standard looking-time measures as utilized in prior research)? (2) To what degree do infants spontaneously display systematic pupil-dilation response to action boundaries within streaming activity? and (3) To what extent does motionese specifically scaffold infants’ detection of action boundaries within continuous activity sequences?

*To what extent does pupillometry reveal infants’ previously-documented preference for motionese over adult-directed action?* In pupillometry research, increased arousal manifests in increases in tonic (or sustained) PDR (Kahneman & Beatty, 1966; Laeng et al., 2012). Such sustained increase in PDR to high-arousal social stimuli has been observed across a variety of infant and preschooler research studies (e.g., Hepach, Vaish, & Tomasello, 2012 & 2015; Martineau, Hernandez, Hiebel, Roché, Metzger, & Bonnet-Brilhault, 2011; Geangu, Hauf, Bhardwaj, & Bentz, 2011; Nuske, Vivanti, Hudry, & Dissanayake, 2014 ; Nuske, Vivanti, & Dissanayake, 2015). We therefore predicted that the previously documented looking-time preference for motionese (Brand & Shallcross, 2008) would be reflected in an enhanced tonic PDR to motionese action sequences relative to that observed in relation to the adult-directed action sequences. As a direct replication of previous research, we additionally measured infants’ looking time to motionese and adult-directed activity (i.e., how long infants looked at the videos). Again, we expected to replicate prior research, predicting that infants would look longer to motionese than adult-directed action sequences.

*To what degree do infants spontaneously display a PDR in relation to action boundaries?* As described earlier, a substantial body of prior evidence indicates that, even in the absence of caregiver scaffolding, infants selectively attend to action boundaries, at least in some kinds of simple, familiar activity sequences (Baldwin et al., 2001; Saylor et al., 2007; Hespos et al., 2009). Interestingly, a study by Jackson and Sirois (2009) provided incidental evidence highlighting the plausibility of this prediction. They measured pupil diameter as infants viewed a train repeatedly entering and exiting a tunnel; infants’ pupil dilation profiles displayed clear signs of a systematic PDR to the juncture at which the train exited the tunnel (that could not be explained by a change in luminance alone), which seems likely to coincide with what adults would judge to be an action boundary. Although investigating infants’ PDR to action boundaries was not the focus of their research, their results nevertheless seem to provide suggestive evidence that infants exhibit a PDR in response to action boundaries, thereby increasing our confidence in predicting that infants would display a systematic PDR to action boundaries in human activity streams.

*To what extent does motionese enhance infants’ detection of action boundaries within continuous activity?* Speech modifications that are characteristic of motherese have been shown to enhance infants’ ability to extract structure from dynamic streams of auditory stimuli (Thiessen et al., 2005; Kemler-Nelson et al., 1989). On analogy with motherese, motionese also may highlight action boundaries (e.g., Brand et al., 2013; Meyer et al., 2013; Rohlfing et al., 2006; Brand et al., 2009; Brand et al., 2007). We thus expected to find a similar facilitative effect of motionese on infants’ processing of action. That is, we predicted that, while infants would display an enhanced PDR to action boundaries even in the non-motionese condition, there would be a synergistic effect in that an increase in infants’ PDR to boundaries would be larger when actions were demonstrated via motionese relative to adult-directed action. To summarize, we predicted a) an overall tonic effect of motionese, such that tonic PDR would be larger for actions demonstrated using motionese over an adult-directed format, b) phasic responses to action boundaries across both motionese and adult-directed demonstrations, but c) larger phasic responses to action boundaries when actions were demonstrated in a motionese than an adult-directed format.

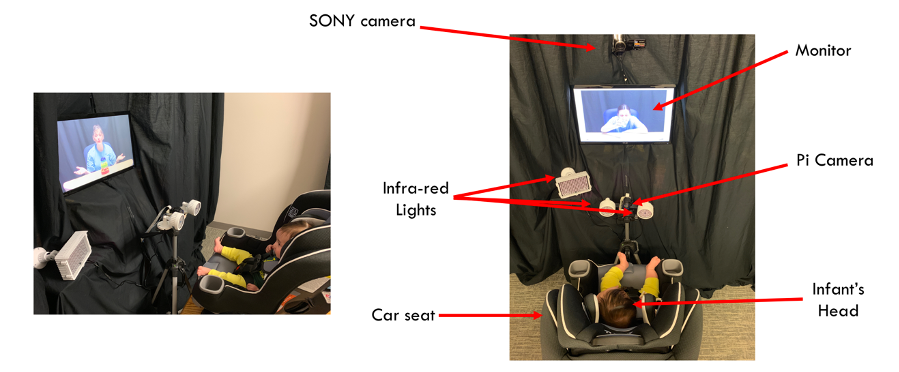
# Methods

## Participants

Thirty-six infants ranging from 9 to 12 months (XX females; Mean = XX days; SD = XX days) and their caregivers were recruited to participate from a database of community members with infants. One infant was immediately excluded due to serious medical issues at birth. Race/ethnicity and maternal education level were representative of the local community; all participants (XX; XX%) identified as White, XX participants (XX%) additionally identified as another ethnicity, and XX% reported some level of graduate training.

## Apparatus

Infants were seated securely in a car seat approximately 82cm from a black floor-to-ceiling curtain, in front of which was a 58cm wide-screen monitor that presented stimuli at a size of 1920 x 1080 pixels. Infants’ movements in the car seat was not restricted, but padding near the head decreased movement.

Pupil dilation was digitally recorded via a Raspbery Pi NoIR camera (infrared camera) placed approximately 38cm from the infant’s eyes, just out of reach. Video from the camera was recorded to a Raspberry Pi single-board computer at a rate of 30 frames per second. Two small infrared lights were placed on either side of the Pi camera and a third, larger, infrared light was placed immediately to the left of the Pi camera. These lights helped to illuminate the infant’s face and make the pupils more readily detectible on the resulting video recording. A second SONY video camera was placed above the monitor and zoomed in to gain a close view of the infant’s face. The video file to which this camera recorded was synchronized with the video being played to the infant, resulting in a recording of the infant’s face that also depicted what the infant was seeing. This second video was used for hand coding infants’ looking throughout the pupillometry session. Figure 1 depicts the experimental setup.  ## Stimuli To investigate the extent to which caregivers’ modifications to infant-directed action (i.e., “motionese”) influenced infants’ processing of dynamic activity sequences, we first created a corpus of videos of infant- and adult-directed action amenable to exploring motionese-related effects. To create the video corpus, we filmed 53 caregivers interacting with their infants (who ranged from 9 to 18 months (29 females; Mean = 403 days; SD = 82.2 days)) with a set of toys. Our methods for eliciting infant- and adult-directed action largely paralleled those of Brand and colleagues (2002): caregivers were asked to demonstrate five different set of objects to their own infant and then to another adult (in this case, a research assistant). We selected objects that were likely to maximize caregivers’ use of motionese in demonstrations to their infants. Thus, in line with the majority of previous research on motionese (e.g., Brand et al., 2002; Koterba & Iverson, 2009; Williamson & Brand, 2014; Brand & Shallcross, 2008), we opted to use novel, rather than familiar, objects. For each object, we identified three or four actions that could be done with the object and provided these suggestions to caregivers.

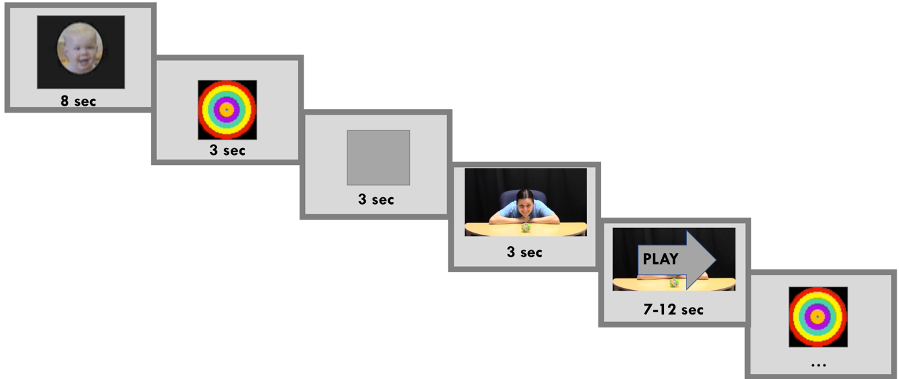
Because PDR is known to be impacted by luminance (e.g., Loewenfeld, 1993), we minimized luminance differences across videos, including (a) installing blackout film on windows, (b) seating all caregivers directly under the overhead light fixture, and (c) disabling the video camera’s automatic exposure mode. We additionally took steps to control for visual differences between caregivers: all caregivers wore the same light blue t-shirt, pulled their hair back away from their face, and had just one object of interest on the table at a time.

From this set of newly collected digital videos (infant- and adult-directed action for each caregiver), we chose twelve clips (with six unique objects and six unique caregivers) to be used as stimuli for the pupillometry study. Our goal was to select videos in which infant- and adult-directed action were clearly distinct, while at the same time controlling for extraneous factors that might account for differences in infants’ attention to adult and infant-directed demonstrations. For example, we opted to have the same actor depicted in both the infant- and adult-directed clips involving interaction with the same object. Additionally, because one of our goals was to optimize the chance that each infant in the pupillometry study would attend to multiple presentations of each clip, we opted to keep the clips fairly short (i.e., seven to twelve seconds in length). While we did allow length of the clips to vary across objects, each pair of infant- and adult-directed clips involving a given object were equated in length. Clips were selected and trimmed from full videos such that one major action boundary (e.g., the moment at which a caregiver finished placing the final item onto a tower) occurred at approximately the same location within a pair of adult and infant-directed demonstrations. See Figure 2 for a detailed summary of the twelve video clips. 

The twelve video clips chosen for the pupillometry task were then coded for a number of features. First, we demonstrated that there were no significant differences in luminance across the videos that might influence infants’ pupil dilation to infant- versus adult-directed action. Next, we verified, with a naïve group of coders, the location of the one major action boundary in each of the video clips. Additionally, a group of trained undergraduate research assistants coded all videos for the extent to which motionese characteristics were used. The results of this coding validated that the selected infant-directed videos were indeed representative of motionese and that the adult-directed videos were rated relatively lower in motionese features. Finally, we coded videos of infants’ spontaneous looking to, and play with, the objects during the infant interaction task and found no differences in infants’ baseline interest across any of the included objects. Caregivers’ questionnaire responses further supported this finding; infants were relatively unfamiliar with all of the objects. Overall, the resulting set of twelve infant- and adult-directed action video clips seemed appropriate for exploring the effects of motionese on infants’ processing of unfolding activity using the pupillometry methodology. NOTE-do i even need all this detail without reporting data–NOTE

## Design

To avoid effects of familiarity with actors and objects, each infant saw only one video from each pair; that is, they saw either the infant- and adult-directed version of each action on a given object. As a result, each infant viewed six unique videos, three infant-directed and three adult-directed. A set of six videos (three infant-directed, three adult-directed) constituted one “block.” Infants viewed up to six total blocks (all six blocks were comprised of the same six videos); thus each infant viewed up to 36 total videos (six repetitions of the same six videos). Because we could not fully counter-balance the videos given our expected sample size, we opted to randomly choose two groups of videos (three infant-directed and three adult-directed in each) and assign an equal number of infants to each group. For each infant, the order of the selected videos was pseudo-randomly assigned such that they never saw more than two infant- or adult-directed videos in a row.

Stimulus presentation was programmed in PsychoPy (Peirce, 2007). As depicted in Figure 3, all blocks started with a brief video of a laughing baby as an attention-getter to help infants orient to the monitor, which played for eight seconds. After the laughing baby attention-getter, a set of moving concentric circles played for three seconds as infants heard a chime sound. The laughing baby and chimes stimuli were acquired via publicly shared materials from the ManyBabies study of infant-directed speech preference (ManyBabies Consortium, 2020). While the laughing baby attention-getter was only played at the start of each block, the circle and chimes attention-getter was played before each video. For three seconds at the beginning of each video, infants were presented with a grey screen signaling the start of the trial. A secondary goal of the grey screen was to match the luminance of the first frame of the video. [FOOTNOTE: Due to an inadvertent change in luminance of the grey screen during stimulus creation, this goal was not satisfied. However, because we were not specifically interested in infants’ PDR to content immediately following the grey screen, this issue was not problematic for interpretation of our results (and is thus explained in further detail in supplementary materials).] After the grey screen, infants were presented with a three-second still image depicting the first frame of the action sequence. The still image was included to allow infants’ pupils to adapt to both the luminance and the characteristics (e.g., featured actor and object) of the visual scene that would be viewed in the upcoming video. After the three-second still, the action sequence began to play silently at a standard rate of 30 frames per second. Upon completion of a trial, infants again heard the chimes while viewing the concentric circles, and then the next trial started with a grey screen followed by a still frame. Once infants had completed their six unique trials the laughing baby played again, starting the next block. This repeated for a total of six blocks or until infants became too fussy to continue.  ## Procedure Caregivers were seated shoulder-to-shoulder with infants, but facing away from the monitor. This setup allowed infants to see caregivers should they look over, but avoided the possibility that infants would be influenced by any caregiver reaction to the stimuli. Caregivers were asked to remain facing away from the monitor and not to interact with infants. We requested that, if infants started to fuss, caregivers simply put their hand on infants. However, if at any point they wanted to take a break or stop the experiment, caregivers should feel free to let us know and we would stop immediately. Once caregivers and infants were seated, the experimenter adjusted the focus of the Raspberry Pi NoIR camera to ensure a clear picture of infants’ pupil. She then went behind the curtain and began the pupillometry session. If infants completed the entire set of six blocks, this part of the session lasted approximately 12 minutes.

## Inclusion Criteria

As mentioned previously, each block of videos contained six unique trials, each trial depicting a unique motionese or adult-directed action sequence. Infants had the opportunity to view six total blocks of videos, each containing the same six trials (i.e., activity sequences). Trials were considered unusable if infants were fussy (as coded by the experimenter) and/or not looking at the screen for at least 50% of the trial. The number of unusable trials was approximately equal for motionese and adult-directed activity. An entire block (i.e., one presentation of the six videos) was dropped from analysis if an infant’s data were unusable on more than 50% of trials within the block. All infants in the current study contributed at least one block of data, thus none were completely excluded from analyses. In total, XX trials across XX infants were included in the pupillometry analyses. The median number of trials contributed by each infant was XX (out of 36 total possible trials).

## Data acquisition

We recorded a separate video for each infant via the Raspberry Pi NoIR camera and Raspberry Pi computer. Each video was run through a Matlab (Matlab, 2019) pupil detection program designed to advance frame-by-frame through the video recorded by the Raspberry Pi camera, find circles, and measure their diameter. Additional detail regarding this process is available in Supplementary Materials on the OSF.

Following previous research using this pupillometry system (Bala et al., 2016) as well as that of other experts in pupillometry research (e.g., Unsworth & Robinson, 2015; Miller, Gross, & Unsworth, 2019), we opted not to interpolate missing values and thus preserved the original data to the extent possible. We z-scored pupil size measurements for each participant, including all relevant frames for each participant (i.e., data from the grey screen, still frame, and video across all blocks and trials but ignoring responses to the attention-getting stimuli) and used these same z-scored data across all analyses. Specifically, we calculated the mean and standard deviation of pupil size for each participant (across all blocks and trials), subtracted the individual’s mean from their pupil diameter at each frame, and divided this value by that individual’s standard deviation. Z-scoring was done for the following reasons: (1) the Matlab program records pupil diameter in pixel size, which is dependent on features of the Pi video (e.g., the degree of zoom on the infants pupil) thus z-scoring made the pupil size measurements more comparable across participants, (2) z-scoring both pupil diameter and luminance makes these measurements more interpretable and more easily comparable as well, and (3) z-scoring controls for individual baseline pupil size differences across subjects, while (4) preserving within-participant pupil diameter differences across motionese and adult-directed action.

After z-scoring, raw pupil values were filtered to eliminate random fluctuations in the data. While there are multiple possible filters that can be used to smooth pupillometry data (see Hepach & Westermann, 2016 for a review in infancy research), we opted to use a hanning filter with a standard window size of 11 frames. The hanning filter uses a moving average, which is one of the common ways of filtering data in pupillometry research and is among those suggested by the creators of R packages for analyzing pupillometry data (e.g., Hepach & Westermann, 2016; Geller et al., 2019). Additionally, the hanning filter can handle missing data, enabling us to perform pupillometry analysis without first interpolating missing values due to blinks or “look aways.” Finally, a visual comparison of filtered and unfiltered data suggested that the hanning filter would appropriately preserve effects of interest while removing extreme values. The hanning filter uses a weighted moving average by generating a normal distribution of weights centered on the frame being filtered and encompassing the surrounding 10 frames (when the window is set at 11, which is the standard, recommended window in pupillometry research). Because of this distribution of weights, the frame of interest has the largest influence on the filtered pupil value, the frames on either side have the next largest influence, and the amount of influence decreases until the distribution covers 11 total frames. Frames outside this window do not contribute to the estimate of pupil size. The z-scored, filtered data are referred to simply as “pupil size” for the remainder of this manuscript.

After filtering the data, we generated baseline-corrected values for each participant on each trial. Our measurement of baseline was the average pupil diameter in the one-second region before onset of the video, calculated separately on each trial. This baseline was chosen for two reasons. First, on viewing infants’ pupil diameter to the grey screen, still frame, and start of the video, it appeared that there was a large luminance effect when the grey screen changed to the still frame. This luminance effect appeared to take about two seconds (of the three-second still frame) to begin to recover. Thus, the final one second of the still frame seemed to be the most appropriate baseline measurement. Secondly, while there is no real consensus regarding how to choose a baseline value, the one second before stimulus onset baseline has been used in a number of infant pupillometry studies (e.g., Geangu et al., 2011; Hepach & Westermann, 2013; Morita et al., 2012, Nuske et al., 2015). In line with methods used by Tanaka and colleagues (in preparation), we opted to control for baseline in analyses via covariation rather than correct for baseline via subtracting or dividing pupil sizes by baseline pupil size. Either subtracting or dividing pupil size by a baseline value would have the disadvantage that, if infants’ pupils were large at baseline, the degree of possible change as they viewed the videos would be attenuated and stimulus-related effects would be obscured.

# Results

Our goals in the current analyses were to (1) examine the effect of motionese on infants’ looking and pupil diameter, (2) explore the extent to which infants’ pupil diameter is indicative of action segmentation, and (3) investigate the influence of motionese on infants’ action segmentation. We used the lme4 package (Bates et al., 2015) in R (R Core Team, 2018) with type III sums of squares (set using the afex package; Singmann, Bolker, Westfall, & Aust, 2017) to estimate linear mixed-effects models. Significance for these models was assessed using the lmerTest package (Kuznetsova et al., 2015; Luke, 2017) with Satterthwaite’s approximation for degrees of freedom. We have specified the exact fixed and random effects structure used for each model below. Baseline pupil size was controlled for in all analyses involving infants’ pupil diameter. The pupillometry study and analysis plan were preregistered (see OSF repository at LINK), however a few minor deviations from the preregistration occurred. None of these deviations influenced our general pattern of results, but nonetheless details of these deviations are described in further detail in the supplementary material.

## Did motionese enhance infants’ overall attention to action?

We first explored the extent to which motionese, relative to adult-directed action, influenced infants’ overall attention to unfolding activity by examining both (1) infants’ looking duration to motionese versus adult-directed action, and (2) their overall average pupil diameter (i.e., tonic pupil size) in response to motionese versus adult-directed action. For these analyses we focused only on frames corresponding to the video portion of each trial (i.e., ignoring the grey screen and still frames) as this was where differences between motionese versus adult-directed action should emerge most clearly given the stimuli employed in the present study.

For analysis of infants’ looking to motionese versus adult-directed action, any frame for which the Pi/Matlab program detected a pupil was classified as “looking” and any frame for which Pi/Matlab did not detect a pupil was classified as “not looking.” We then created a proportion of time spent looking to each trial by summing the number of frames participants spent looking during each trial and dividing that value by the total number of frames in the trial. We conducted a linear mixed-effects model predicting the proportion of time spent looking to a given trial from a fixed effect of demonstration type (i.e., motionese versus adult-directed action) and random intercepts for subjects and videos. Because looking times were strongly positively skewed (i.e., most infants looked for almost the full trial duration), looking times were z-scored prior to analysis. In line with previous research (i.e., (**???**)), we found that infants looked significantly longer to motionese (*M* = 93.02%, *SD* = 11.45%) over adult-directed action (*M* = 91.31%, *SD* = 13.14%), (), .

To examine the influence of motionese versus adult-directed action on infants’ tonic pupil size, we ran a linear mixed effects model predicting infants’ z-scored, filtered pupil diameter from a fixed effect of demonstration type (motionese versus adult-directed action) with random intercepts for subjects and videos, controlling for infants’ baseline pupil diameter. This time, contrary to our predictions, we did not find a significant effect of demonstration type, (), , though infants’ average pupil diameter tended to be larger in response to motionese ( *M* = 0.02, *SD* = 0.78) over adult-directed activity sequences (*M* = -0.11, *SD* = 0.81). Thus, while our analysis of infants’ *looking* to motionese versus adult-directed action replicated prior results, our analysis of infants’ *pupil diameter* (an attempt to replicate previous findings with a novel methdology) did not reveal a systetmatic motionese preference.

## Did infants selectively attend to action boundaries in continuous activity sequences?

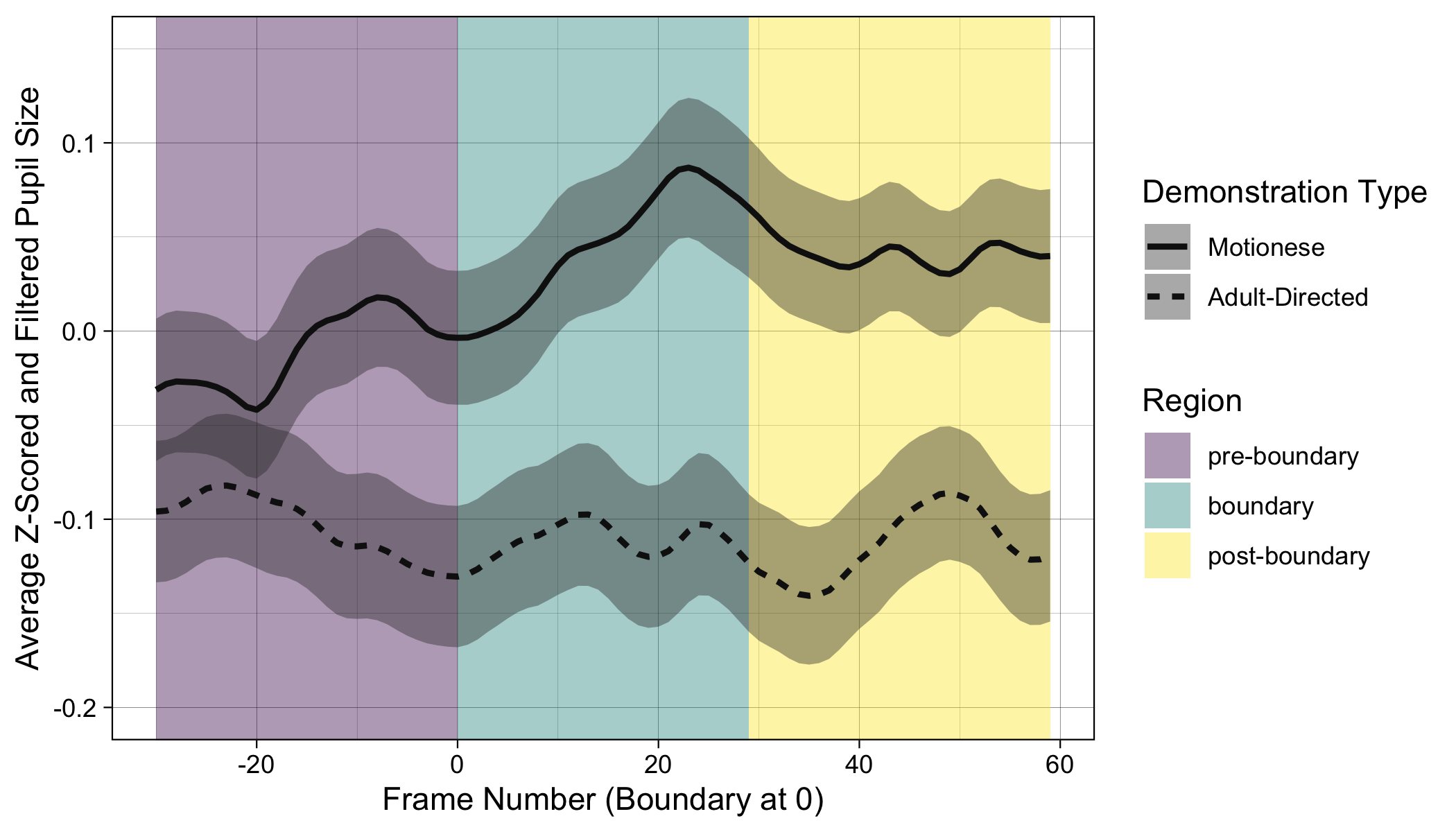
In our next set of analyses, we explored the extent to which infants preferentially attended to boundaries in unfolding activity sequences, as indexed by changes in pupil diameter. For these analyses, we focused in particular on activity surrounding the one major action boundary depicted within each video. We first defined pre-boundary, boundary, and post-boundary regions in each video. The pre-boundary region covered the one second of activity (or 30 frames) occurring prior to the action boundary. The boundary region began at the action boundary and extended for the next one second (30 frames), and the post-boundary region began at the end of the boundary region and continued 1 additional second, or 30 more frames. In previous research exploring adults’ PDR to action boundaries (e.g., Tanaka and colleagues, in preparation), researchers used half-second pre-boundary, boundary, and post-boundary regions. However, this time window might miss infants’ pupillary response to the action boundary, because there is evidence that infants’ pupils respond to cognitive stimuli more slowly than adults’ (e.g., Verschoor, Spapé, Biro, & Hommel, 2013; Verschoor, Paulus, Spapé, & Hommel, 2015; Zhang, Jaffe-Dax, Wilson, & Emberson, 2018). Thus, we opted to extend the windows to one-second regions. This timing is also consistent with prior work in which researchers incidentally provided information about the timing of infants’ response to action boundaries. Jackson and Sirois (2009) measured infants’ PDR to a train entering and emerging from a tunnel. Visual examination of infants’ PDR to the boundary at which the train emerged from the tunnel suggested that the response peaked and began returning to baseline within one second after the action boundary.

To test for a possible boundary-related PDR across all videos (regardless of whether activity depicted was motionese versus adult-directed action), we ran a linear mixed effects model predicting z-scored, filtered pupil diameter from a fixed effect of region (pre-boundary, boundary, post-boundary) and random intercepts for subjects and videos, controlling for baseline pupil size. Because we were specifically interested in boundary effects, the video frames included in these analyses were limited to those occurring in pre-boundary, boundary, and post-boundary regions. Video frames outside of these regions were eliminated from the current analyses. We found a significant main effect of region, , . This effect remained when luminance was added to the model as a covariate in a follow-up analysis.

To explore the locus of this effect of region, we ran a set of Bonferroni-corrected pairwise comparisons. Infants’ pupil diameter was significantly larger to boundary ( *M* = -0.03, *SD* = 0.81) and post-boundary regions ( *M* = -0.04, *SD* = 0.77) than to pre-boundary regions ( *M* = -0.06, *SD* = 0.81) ( and , respectively). However, pupil diameter to post-boundary regions did not significantly differ from pupil diameter to boundary regions (). To summarize, infants’ pupil diameter did increase to action boundaries as expected. However, infants’ pupils did not constrict in the one second region after the boundary.

## Did motionese enhance infants’ response to boundaries within continuous activity?

To explore the extent to which motionese influenced infants’ response to boundaries, we ran the same mixed-effects model described above, but now included fixed effects of demonstration type and an interaction between region and demonstration type, while still controlling for baseline pupil size. As in previous analyses, we found no significant effect of demonstration type, , , but a significant effect of region , . We also found a significant interaction between demonstration type and region, , .

To explore this interaction, depicted in Figure 4, we ran two separate mixed-effects models for motionese and adult-directed demonstrations. In adult-directed demonstrations, there was no systematic effect of region, , . Pupil diameter did not differ significantly across pre-boundary ( *M* = -0.01, *SD* = 0.80), boundary ( *M* = 0.04, *SD* = 0.79), and post-boundary regions ( *M* = 0.04, *SD* = 0.75). In contrast, for motionese demonstrations we observed a significant effect of region, , , that we followed up with a set of Bonferroni-corrected pairwise comparisons. PDR to pre-boundary slides ( *M* = -0.01, *SD* = 0.80) was lower than PDR to boundary slides ( *M* = 0.04, *SD* = 0.79) and post-boundary slides ( *M* = 0.04, *SD* = 0.75), . However, PDR did not differ significantly between boundary and post-boundary slides, . To summarize, in response to motionese demonstrations, infants’ pupil size increased within boundary regions (relative to pre-boundary regions) and remained high post-boundary. These effects were not observed in infants’ PDR to adult-directed demonstrations. Again, a follow-up analysis controlling for luminance indicated it did not in any way systematically influence these patterns. 

# Discussion

To briefly review, we first explored the extent to which infants preferred to view motionese over adult-directed demonstrations, using both looking time and pupil size measures. Previous studies (e.g., Brand & Shallcross, 2008) found that infants prefer to view motionese over adult-directed demonstrations. In the current study, our analysis of looking times replicated this effect while mean pupil size only trended in this direction. A subsequent set of analyses examined whether infants displayed a PDR to major action boundaries across the videos, as previously documented for adults by Tanaka and colleagues (in preparation). We measured infants’ pupil size during pre-boundary, boundary, and post-boundary regions of unfolding activity sequences. Overall, infants’ pupil diameter displayed systematic increase in boundary relative to pre-boundary regions, and then remained high afterwards. Interstingly, this effect only held for motionese action demonstrations and was absent when infants viewed adult-directed adult-directed action. Thus, infants indeed displayed a boundary-related PDR as was previously observed in adults, but for this was only the case for motionese demonstrations, supporting our prediction that motionese facilitates infants’ detection of segmental structure within unfolding activity sequences.

A collection of analyses examined our first major research question: whether infants would prefer motionese over adult-directed action. While we did replicate this previously observed preference (e.g., Brand & Shallcross, 2008) using looking time data, we found no significant difference in infants’ average pupil size in response to infant- versus adult-directed demonstrations. \* I FEEL LIKE I NEED TO GIVE SOME INTERPRETATION OF THIS HERE, BUT I’M A LITTLE STUCK \*

Another collection of analyses addressed our second and third questions: whether infants would display a PDR in response to major action boundaries, and whether motionese would facilitate such boundary-related responding. As it turned out, a significant boundary effect – increased pupil diameter during boundary regions – was observed, but only for motionese demonstrations. These findings suggest that motionese indeed enhance infants’ detection of segmental structure in unfolding activity. It is worth noting, however, that infants’ boundary-related PDR differed in other ways from the comparable pattern observed in one previous study with adults. For one thing, infants’ boundary-related PDR during motionese demonstrations was slower (occurring on average within a one-second region after the boundary) than adults’ (occurring on average within a half-second region after the boundary). This was consistent with other evidence that infants’ pupil response is generally slower than adults’ (e.g., Verschoor et al., 2013, 2015; Zhang et al., 2019). Another difference was that a linear trend provided the best characterization of infants’ boundary-related PDR; in contrast, while Tanaka and colleagues (in preparation) did find both significant linear and quadratic trends, the quadratic trend was stronger and suggested that adults’ pupil diameter began to return to baseline shortly after their boundary-related PDR. Why might infants’ pupil diameter remain high after the boundary? One likely explanation is again, that infants’ pupil response – including the return to baseline – may simply be slower than adults’. Also, as we observed from visual examination of the videos, there was often considerable post-boundary movement in infant-directed demonstrations. Upon examination of the videos, we observed that caregivers depicted in the videos frequently did things like spreading their arms to exaggerate the fact that a boundary had occurred. Often caregivers would also make excited and exaggerated facial expressions after finishing a unit of action. These features of the stimuli are of course characteristic of motionese, and could serve to sustain infants’ arousal, thereby reducing a tendency for pupil diameter to return to baseline levels after a boundary. Overall, these findings provide the first evidence to date that motionese action modifications alter infants’ online action processing. In particular, motionese scaffolds infants’ detection of segmental structure within dynamically unfolding action. – MAYBE WANT TO CHANGE RESULTS TO TALK ABOUT LINEAR AND QUADRATIC TRENDS RATHER THAN PAIRED COMPARISONS (REALLY JUST A DIFFERENT WAY OF DOING THE SAME THING) –

These findings raise an obvious next question: Precisely what is it about motionese that facilitates infants’ detection of action boundaries? One hypothesis is that it’s something about the demonstration itself – perhaps caregivers move and manipulate objects in ways that highlight structure within dynamic activity. Another possibility is that motionese simply heightens infants’ attention, which increases the chances that they will detect structure within unfolding activity. Perhaps instead, or in addition, motionese indicates to the infant that this demonstration is “for me.” When infants can infer that an action demonstration is directed to them, this might further enhance their attention and thus facilitate their detection of structure as activity unfolds. Evidence from the current study speaks to all of these interpretations.

First, could something about the motionese demonstrations have enhanced infants’ detection of action boundaries? In the related domain of motherese, or infant-directed speech, it has been suggested that specific characteristics of motherese input promote infants’ ability to find structure in speech (e.g., Kemler-Nelson et al., 1989; Gleitman, Newport, & Gleitman, 1984). Perhaps characteristics of motionese similarly facilitate infants’ detection of structure within dynamic activity. As described previously, however, the steps we took to match infant- and adult-directed demonstrations reduced some characteristics of motionese that might otherwise serve to highlight action boundaries. For example, shorter action sequences – often characteristic of motionese – might highlight boundaries with pauses or repetition of shorter units of action. However, these dimensions of motionese were reduced in our stimuli. We did find, though, that pixel values – sometimes used as an index of motion change (e.g., Hard et al., 2011; Loucks & Baldwin, 2009) – were greater both before and, after, action boundaries in motionese demonstrations. Visual examination of our videos confirmed that this large degree of pixel change often corresponded to body movements that might highlight the fact that a boundary had just occurred (such as large, emphatic arm movements). Additionally, enthusiasm and interactiveness were high in our infant-directed demonstrations. There is some evidence (e.g., Brand et al., 2013) that caregivers’ gaze toward infants, reflected in the “interactiveness” coding, coincides systematically with action boundaries. It is additionally possible that exaggerated facial expressions, which likely contributed to higher observed enthusiasm ratings in our findings, frequently coincided with action boundaries. These features of our motionese demonstrations could have facilitated infants’ detection of structure. Further coding of the video stimuli will be necessary to fully explore these possibilities; this represents an interesting future direction.

A second hypothesis is that motionese increases infants’ attention overall and, if infants’ attention is increased, they might be better able to attend to action and thus to detect segmental structure. Prior research supports this hypothesis: when infants are in an attentive state (as indexed by heart rate) during stimulus presentation, they are more readily able to recognize that stimulus at later test (Richards, 1997; Frick & Richards, 2001). While that research focused on infants’ recognition memory, and not their sensitivity to structure as in the present reseach, there is reason to believe that the two might be related. For example, a substantial body of evidence suggests that infants’ (and adults’) attention to structure within action is linked to later memory (e.g., Sonne et al., 2016, 2017; Hard et al., 2011; Zacks et al., 2006). While we did not find a significant difference in infants’ overall attention to infant- over adult-directed action, there were a number of hints that a motionese preference was at least weakly present. Thus, despite the fact that these comparisons did not reach statistical significance, infants may have been in a more attentive state in response to motionese demonstrations, enhancing their processing of the unfolding activity.

The final alternative we’ve suggested above is that motionese indicates to infants that this demonstration is “for me.” Information presented to infants in a social context appears to facilitate learning (e.g., Baldwin, Markman, Bill, Desjardins, Irwin, & Tidball, 1996; Baldwin, 2000; Akhtar & Tomasello, 2000; Sage & Baldwin, 2011; Csibra & Gergely, 2009), which seems to be either illustrative of, or closely related to, a phenomenon that Kuhl and colleagues (e.g., Kuhl, Tsao, & Liu, 2003; Kuhl, 2007) call “social gating” following a similar phenomenon in bird-song learning (e.g., Doupe & Kuhl, 1999; Kuhl, 2003). One interpretation of social gating is that a social context simply elicits an increase in infants’ overall attention, analogous to our second alternative account outlined above. However, it has been demonstrated that infants presented with stimuli in both social and non-social contexts learn better from the social context, despite equivalent attention to stimuli across contexts (e.g., Baldwin et al., 1996; Sage & Baldwin, 2011). Thus, there is likely to be something more driving infants’ learning from social stimuli like the motionese demonstrations in the current research. Perhaps contributing to this effect, Gergely, Csibra, and colleagues (Csibra & Gergely, 2006, 2009, 2011; Gergeley, Egyed, & Kiraly, 2007) suggest that pedagogical cues, which abound in motionese, signal to infants that they are being taught and, consequently, infants adopt a “pedagogical stance” that primes them to learn. Perhaps motionese promoted infants’ adoption of a pedagogical stance, and thereby enhanced their detection of segmental structure in unfolding activity sequences.

The current study provided evidence consistent with all of these alternative accounts, without singling out any particular account as the most plausible mechanism by which motionese could enhance infants’ attention to structure. At this juncture, it seems unlikely that any one of the mechanisms proposed above can fully explain why infants displayed a pupillary response to action boundaries within infant-directed demonstrations but not to comparable boundaries within adult-directed action. In contrast, it seems plausible, and perhaps even likely, that all these mechanisms operated in concert to enhance infants’ processing of dynamic action.

In conclusion, this research extends current understanding of the ways in which motionese benefits infants’ development. Previous research has documented the motionese phenomenon, that human caregivers spontaneously modify motion when demonstrating action to infants. As well, prior work demonstrated both that (1) infants prefer motionese over adult-directed action, and (2) motionese promotes infants’ imitation of novel activity sequences. What had remained mysterious, however, was the precise ways in which motionese might alter infants’ processing of dynamically unfolding activity. This question had been difficult to address, in part, because existing methodologies were not well-suited to probing infants’ moment-to-moment action processing. The research reported here offers a signal advance on this methodological front, and at the same time provides the first evidence to date that motionese promotes infants’ detection of segmental structure within dynamically unfolding activity. Put another way, the current findings indicate that, by providing motionese demonstrations, caregivers spontaneously enhance infants’ detection of boundaries within continuous activity. This facilitates infants’ discovery of action units that are amenable to encoding in memory for later recall, and likely promotes their efficient processing of similar activity sequences when subsequently encountered.

# References