Egocentricity in infants' play with familiar objects in caregiver-child interactions

Rajalakshmi Madhavan[1,2*†], Ming Yean Sia[2,4*], Xiaoyun Chen[5], Roger Mundry[3,6,7] and Nivedita Mani[2,3]

[1] Department of Language and Linguistic Science, University of York, Heslington, York YO10 5DD, United Kingdom
[2] Psychology of Language Department, University of Göttingen, Goßlerstraße 14, 37073, Göttingen, Germany
[3] Leibniz ScienceCampus Primate Cognition, Kellnerweg 4, 37077, Göttingen, Germany
[4] Department of Child and Family Science, National Taiwan Normal University, Taipei, Taiwan
[5] Department of Psychology, Lancaster University, LA1 4YF, Lancaster, United Kingdom
[6] Cognitive Ethology Laboratory, German Primate Center, Leibniz Institute for Primate Research, Göttingen 37077, Germany
[7] Department for Primate Cognition, Johann-Friedrich-Blumenbach Institute, University of Göttingen, 37077, Germany

Keywords: parent-child interactions, joint attention, learning, sustained attention, head-mounted eye-tracking

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Grant/Award Number: 254142454/GRK 2070, and DFG Project-ID 454648639 - SFB "1528". This work was also supported by the Leibniz Association through funding for the Leibniz ScienceCampus Primate Cognition (W45/2019 – Strategische Vernetzung). We thank all research assistants involved in collecting, curating, and coding our video data.

Data availability statement

All data (and corresponding analyses scripts) can be found of the OSF project page (https://osf.io/2ruac/). We cannot share the video data to avoid breach of anonymity and data protection.

Ethics statement

Ethics approval was granted by the ethics committee for Psychology at the University of Göttingen.

Conflict of interest

The authors state that there is no conflict of interest.

* Shared first authors
† Corresponding author at: Department of Language and Linguistic Science, University of York, Heslington, York YO10 5DD, United Kingdom. E-Mail: rajalakshmi.madhavan@york.ac.uk

Egocentricity in infants' play with familiar objects in caregiver-child interactions

Research highlights

- Who leads joint-attention (JA) episodes when playing with both novel and familiar objects, and how do parents label these objects? Does it influence learning?
- Children lead most JA instances, and mostly to familiar objects; parents labelled familiar objects more frequently.
- Children did not show successful learning of novel word-object associations from the interactive play sessions.
- The findings underscore the contingent nature of parent-child social interactions, where children lead and parents follow, especially in the case of familiar objects.

Abstract

Parents and children appear to coordinate their attention to objects in their environment, often via mutual gaze. Children, however, display a novelty bias in interacting with objects, looking longer at novel objects relative to familiar objects. This may allow parents to follow in on their child's focus of attention and label the novel objects, creating moments of optimal learning. The current study examined this with regard to whether parents are more likely to lead instances of joint attention to novel relative to familiar objects and how children learn from periods of child-led or parent-led joint attention. In particular, we investigated whether (i) parents lead more instances of joint attention when playing with novel relative to familiar objects, (ii) parents preferentially label novel relative to familiar objects, and (iii) children's learning of novel word-object associations is affected by the frequency of labelling and children's sustained attention towards the objects. We found that not only do children lead more instances of joint attention, but, relative to their caregivers, children lead more instances of joint attention to familiar objects relative to novel objects. Parents also appeared to follow in on their child's attention and labelled familiar objects more often than novel objects. Furthermore, we found no evidence for children's recognition of the novel word-object associations. Our findings highlight the contingent nature of social interactions between caregivers and infants, with children leading and parents following their child's lead, especially with regard to more familiar objects in the child's environment.

Keywords: parent-child interactions, joint attention, learning, sustained attention, head-mounted eye-tracking

Introduction

Even young infants actively seek information about objects they are interested in (Begus & Southgate, 2012) and caregivers respond to such biddings in a contingent manner (Tamis-LeMonda et al., 2013, 2018; Z. Wu & Gros-Louis, 2015). Socially contingent interactions yield instances in which caregiver and child gaze at the same object at the same time, creating a period of joint attention (JA). Children fixate objects longer during episodes of JA (McQuillan et al., 2020; Yu & Smith, 2016), allowing them to encode such objects better and caregivers the opportunity to label the object when it dominates the child's visual field (Suanda et al., 2019). JA also provides children with unambiguous cues as to the referent of a word, especially in cases where the child may not be certain of the meaning of a word. While previous research demonstrated differences in children's seeking and parents' provision of information pertaining to novel and familiar objects (Chen et al., 2021), here, we zoom in on such interactions to examine whether there are differences in caregiver-child JA to novel and familiar objects as well as children's learning from such periods of JA. In particular, we examine whether (a) parents are more likely to lead episodes of JA when playing with novel compared to familiar objects, (b) we replicate previous differences in caregivers' provision of information regarding novel and familiar objects in more varied settings, and (c) children learn novel word-object associations during child-led or caregiver-led JA.

Child-led or parent-led joint attention

As explained, JA refers to interactions where the child and their social partner mutually engage with an object, either by gazing at it, pointing towards it, or gazing at the object that the partner is handling (Akhtar & Gernsbacher, 2007; Baldwin, 1995; Tomasello, 1995). JA may be initiated by the child (child-led JA) such that the adult follows the gaze of the child, or initiated by the parent (parent-led JA) such that the child responds to the adults'

gaze. Although caregivers and children seem equally likely to lead or follow each other's visual attention in natural interactions (Yu & Smith, 2013), distinguishing between these two types of JA is crucial to understanding the relationship between JA and word learning. Indeed, while children follow adults' gaze from early on (Reid & Striano, 2005), children's ability to follow adult gaze improves with age. In contrast, children's initiation of JA remains relatively constant in the first and second year of life, likely due to adults following their child's gaze (Mundy, 2006 in Akhtar & Gernsbacher, 2007). Gaze following in children is also positively associated with later language skills (Brooks & Meltzoff, 2005; Tenenbaum et al., 2015). Such findings speak about the importance of parent-led JA in early development, especially given that parent-led JA – but not child-led JA – correlates with later vocabulary development (Mundy, 2006 in Akhtar & Gernsbacher, 2007). However, if caregivers' attempts to redirect their child's attention early in development do not lead to JA, there may to be reduced effects of such redirected attention on learning (Mason et al., 2019). Against this background, the current study explores the dynamics of caregiver-child interactions in terms of who is leading episodes of JA in play and the influence of such parent- and child-led JA on learning and parent behaviour.

Joint attention and parental labelling

Recent advances in head-mounted eye-tracking and motion tracking technology for young infants and children provide finer detail about how children lead and respond to bids for JA in natural interactions with their caregivers. This work suggests that parents label visually dominant objects in the child's visual field more often (Yu & Smith, 2012), track infants' attention (Abney et al., 2020) to synchronise their speech with the object of their child's interest (Custode & Tamis-LeMonda, 2020), and respond contingently to infants during JA, thereby enhancing children's learning (Cleveland et al., 2007; Tamis-LeMonda et al., 2014). Indeed, parents' utterances during JA coincide with longer visual attention of

infants (Lawson et al., 1992; Suanda et al., 2016), potentially explaining the reported positive correlation between JA and later vocabulary development (Carpenter et al., 1998; Morales et al., 2000; Tomasello & Farrar, 1986).In other words, the contingency between parent labelling behaviour and JA may underlie the relationship between JA and language learning. This contingency may, however, be influenced by parents and children's familiarity with the objects.

The role of object familiarity in caregiver-child interactions

Children's attention and engagement with objects is influenced by their familiarity with these objects. While some studies report that children display a novelty preference, e.g., preferentially fixate a novel stimulus over a familiar stimulus (Rose & Feldman, 1987; Kirkham et al., 2007), other studies report finding a familiarity preference in children's visual engagement with objects (Johnson et al., 2009). A novelty preference has also been reported in naturalistic interactions, where children gaze longer at novel objects relative to familiar objects in interactions with their caregivers (Chen et al., 2021).

Parent behaviour is also modulated by their familiarity with the objects that are the focus of the interaction. Thus, one study examining interactions with novel and familiar objects finds that parents label novel objects more than familiar objects when the novel objects are real and identifiable toys (Danis, 1997). However, in a study where the novel objects were novel to both parents and children, parents labelled the familiar objects more frequently than novel objects (Chen et al., 2021). In contrast, novel labels were provided in a more strategic manner, i.e., during optimal moments when children were already gazing at the novel objects (Chen et al., 2021).

Taken together, the findings reviewed above suggest that parents are more likely to follow their child's lead in directing their visual attention to novel objects, in a manner akin to their strategic labelling of novel objects when their child is gazing at them. On the other hand,

given that children and adults are equally likely to initiate episodes of JA to familiar objects (Yu & Smith, 2013), they may be equally likely to initiate episodes of JA to *novel* objects. However, on the premise that caregivers may be inclined to provide children with information about objects they know their child to be unfamiliar with, caregivers may intentionally direct their child's attention to the novel objects to create optimal moments of learning for their child.

Current study

Against this background, the current study examines the extent to which children and parents lead episodes of JA to novel and familiar objects in naturalistic interactions, as well as children's learning of novel word-object associations from child- and parent-led JA. The study was conducted in a naturalistic environment in our lab mimicking a regular play room, allowing parent-child dyads to play with four toys as they would at home. Caregivers reported two of the objects to be familiar to the child and two objects to be novel to the child. Critically, and in contrast to previous work, we ensured that both familiar and novel objects were presented to children during the same session. Given that children's attention to objects is likely to be influenced by their familiarity with other objects present, it is vital to examine children's role in leading instances of JA in settings in which children are presented with both novel and familiar objects at the same time. Indeed, natural play settings are unlikely to be segregated in terms of the novelty of objects, and ensuring that the interaction presents both familiar and novel objects avoids drawbacks of inter-individual differences in children's attention and caregiver labelling behaviour in between-subject manipulations. We also ensured that caregivers were familiar with the objects considered to be novel to the child, given that adults' labelling behaviour may differ based on whether they are confronted with objects novel to both child and parent or with objects novel to the child but not to the parent. Caregivers and infants wore head-mounted eye-trackers throughout the study allowing us to

capture their eye-movements throughout the play situation. In addition, cameras mounted around the play table allowed us to collect behavioural data during the interaction.

We outline three main research questions for this study:

(1) Are parents more likely to lead instances of JA when the object of attention is an object novel to the child? On the one hand, similar to findings of parents' labelling behaviour being led by children's attention to novel objects, children may be more likely to lead JA instances to novel objects relative to familiar objects. However, given that caregivers may be inclined to provide children with more information about objects they know their child to be unfamiliar with, we hypothesised that caregivers will lead more episodes of JA when playing with novel as opposed to familiar objects.

(2) Do caregivers label familiar and novel objects equally frequently? While previous studies showed that caregivers label familiar objects more frequently than novel objects (Chen et al., 2021), these differences may be due to the settings containing either only novel or familiar objects, with the novel labels also being less familiar to the caregivers. We, therefore, predicted that our setting may prompt caregivers to label novel objects more frequently.

(3) Is the robustness of children's recognition of the novel objects associated with how frequently their parents labelled the object or how much sustained attention children paid to the object during the play interaction? We predicted that increased sustained attention and more frequent object labelling of the novel objects will boost children's learning of the novel word-object associations.

Methods

Ethics

Ethics approval was granted by the Psychology Institute's Ethics committee of the authors' university and all caregivers provided informed written consent prior to participation in the study. Children were offered a book as a token of appreciation.

Preregistration

We preregistered our minimum sample size ($N = 20$), included predictor and response variables, hypotheses, and planned analyses on the Open Science Framework (https://osf.io/h8bmt) prior to data analysis. While we initially preregistered analysis of children's manipulation of the objects, we do not report these analyses here as they would be beyond the scope of this study. The datasets and analysis scripts can be found on the OSF page of the project (https://osf.io/2ruac).

Participants

We recruited 51 parent-child dyads, 10 of which were excluded due to fussiness. Another 10 dyads were excluded because the child was bilingual ($n = 1$), or knew all selected toys (hence did not fulfil our study criteria of playing with novel and familiar objects, $n = 3$), or ended the play session early ($n = 6$). These exclusion criteria were preregistered. The final sample size was 31 parent-child dyads (16 boys, 15 girls, age range 14-23 months; $M_{age} =$ 17.74, $SD_{age} =$ 2.92). Children were German-speaking monolinguals, carried to full-term and with no diagnosed developmental disorders. All dyads were recruited from a database of children managed by the laboratory, and came from families living near and around the city. Please note that race/ethnicity information is not commonly collected in Germany and cannot be reported.

Stimuli

We chose objects that are typically familiar or unfamiliar to young children from two object categories, *animals* and *vehicles* (based on Workbank and CHILDES corpora). The categories were chosen from previous studies examining children's learning of objects from

different object categories (Borovsky et al., 2016; Madhavan & Mani, 2024). The familiar objects were cat, bear, car, and bus; and the novel objects were iguana, seal, submarine, and wagon. All objects were 3-D printed in white and painted a different colour to facilitate coding (see Figure 1). The toys were on average 167.7cm³ in size (Suarez-Rivera et al., 2019; Yu et al., 2019) so that they could be held by 18-month-olds but remained visible when held by adults (Supplementary materials S1 for the exact measurement).

Figure 1

Pictures of the (a) familiar and (b) novel toys

Procedure

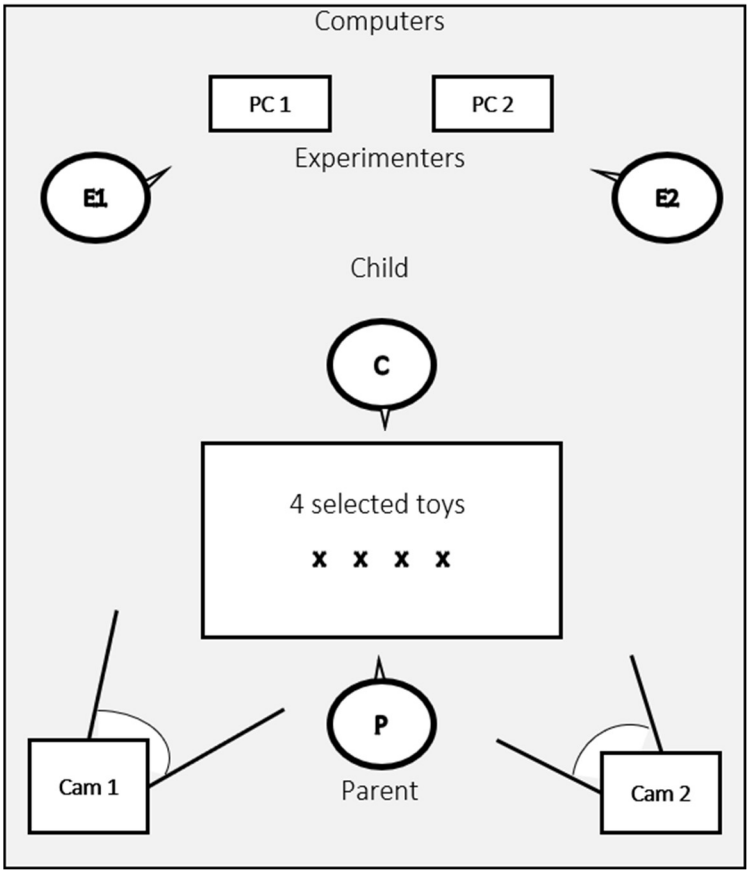
The study was split into two tasks; a play task and an object recognition task. Parents and children wore a head mounted eye-tracker (Positive Science LLC; Model DB9) which captured their eye movements at a sampling rate of 30Hz. In addition, two additional cameras recorded participants' gazes and interactions in case the head-mounted eye tracking failed to record (see Figure 2).

Two experimenters (hereafter E1 and E2) were present at each session, and prior to the sessions, one kept the child occupied while the other explained the study to the parent. When the parent-child dyad arrived at the lab, they were invited to sit across each other at a child-sized table.

Parents were told that the aim of the study was to examine how parents interact with their children and that they should play with their child and the toys as they would at home for six minutes. After parents provided informed consent, they were asked to choose four toys from two different boxes (two toys from each box), where one contained only familiar objects and the other only the novel objects. Parents were not told the difference between these two boxes and the order of presentation was counter-balanced across participants. Parents were asked at the end of the experiment whether the toys were novel to their child. Care was taken to ensure that the child did not see the toys at this stage so that the child was not familiarised with the stimuli before the experiment started. E1 then put the eye-tracker on the parent, after which the parent helped E1 put the eye-tracker on the child. Testing would continue without the child's eye-tracker if the child refused to wear the cap or eye-tracker.

Figure 2

Schematic diagram of the setup.



Note: The play table was approximately 80x50 cm in size.

Next, E1 started recording on two video cameras and the eye-trackers while E2 calibrated the eye-trackers using a red rattle, first a 9-point calibration horizontally on the table and then a 7-point calibration vertically. Then, E2 presented the four toys in a row on the table. The toys were initially hidden inside a box which was removed once E2 had synced cameras and eye-trackers by hitting a rattle on the box. After six minutes had passed or when the child lost interest in the toys, E2 synced the cameras again by hitting a rattle on the table while informing the parent that the play task had ended. E2 now removed the toys from the table. E1 stopped recording and requested the parent to sit quietly beside the child for the object recognition task, while E2 took the parent's place opposite the child.

E2 began the object recognition task once recording started again. This task tested children's retention of novel word-object associations used in the play task across six trials. Specifically, every trial began with E2 placing two objects on a table, to the left and the right of the child. The experimenter then asked the child for one of the objects in a child-friendly manner using the label used in the play task, e.g., "where is the cat?". Children were neither praised nor corrected for their responses. The experimenter always thanked the child and then removed both objects from the table to end the trial. If the child did not provide any explicit response such as pointing or reaching for one of the objects, the experimenter would repeat the question up to three times, after which she would thank the child and remove the objects. Of these six trials, the first two were warm-up trials which presented children with the two familiar objects, such that each of the familiar objects was tested once. The other four trials were test trials which presented children with the two presumably novel objects[1]. Thus, each of the novel objects was tested twice. The target location, i.e., left or right, was

[1] We used "presumably" because the object recognition task was conducted before verifying children's knowledge of the four toys. In case children already knew one of the novel objects, their data in the object recognition task would be excluded from the analysis.

counterbalanced across children. At the end of the experiment, the child was given a certificate and a book as a token of appreciation.

Coding

We coded parent and child gaze toward objects during the play session, from which we derived the duration and number of joint attention episodes. Parents' labelling of familiar and novel objects throughout the play session were also coded, as well as children's looking and pointing behaviour during the test session. Of the 31 dyads, only seven children were willing to wear the eye-tracker. Therefore, children's gaze during the play task was coded using ELAN (v6.3) regardless of whether the eye-tracking data were available. In addition, parents' and children's object manipulation were coded using ELAN. Parents' gaze was coded using GazeTag (build 0.94+), which is a software that codes gaze data obtained from the Positive Science eye-trackers. Table 1 summarises the coding scheme and criteria used to define the variables of interest during the play task, which were pre-registered.

Table 1

Coding scheme for the variables of interest in the play task.

Variables	Criteria
Gaze	Eye-gaze was coded separately for parents and children. It is the duration of fixations towards any of the five areas of interest (AOI), namely, the four chosen toys and the face of the social partner. Each fixation lasts at least 100ms (or three frames).
Joint Attention	Joint Attention was measured using gaze. Joint attention is defined as an episode in which both parent and child jointly fixate their gaze on the same object at the same time. In accordance with previous studies (Yu & Smith, 2016, 2017), fixations need to be longer than

	500ms to be counted as joint attention, with looks away of less than 300ms permitted.
Leader	For each joint attention episode, we coded the individual who initiated the joint attention, i.e., the parent or the child.
Sustained	Sustained attention was measured only in terms of children's eye-gaze, i.e., children's visual attention towards an object which lasted at least three seconds. In accordance with Yu and Smith (2016), no leeway of looks away was allowed
Object labelling	The number of times parents labelled each of the four toys throughout the interaction.
Proportion of object labelling (for novel objects)	The proportion of novel object labelling was calculated by dividing the number of labelling instances for a particular novel object by the total number of labelling instances for all novel objects.
Proportion of Target Looking (PTL)	To determine children's recognition of the novel objects, we measured the proportion of target looking (PTL) (which was coded according to the Gaze coding scheme – see top of the table). We first identified a pre-naming and a post-naming window within each recognition task. Pre-naming was defined as the 2000ms before the experimenter uttered the label of the target novel object, and post-naming was defined as the 4000ms after the experimenter finished labelling the target novel object (see Methods for an expanded view of the task). We calculated the PTL for each phase by dividing the total gaze duration towards the target object, i.e., the labelled object,

by the total gaze duration towards both target and distractor in each trial.

Analysis

We fitted a series of General or Generalised Linear Mixed Models (LMM or GLMM; Baayen, 2008) to analyse the data. The details of how we built the models can be found in the supplementary materials (S2).

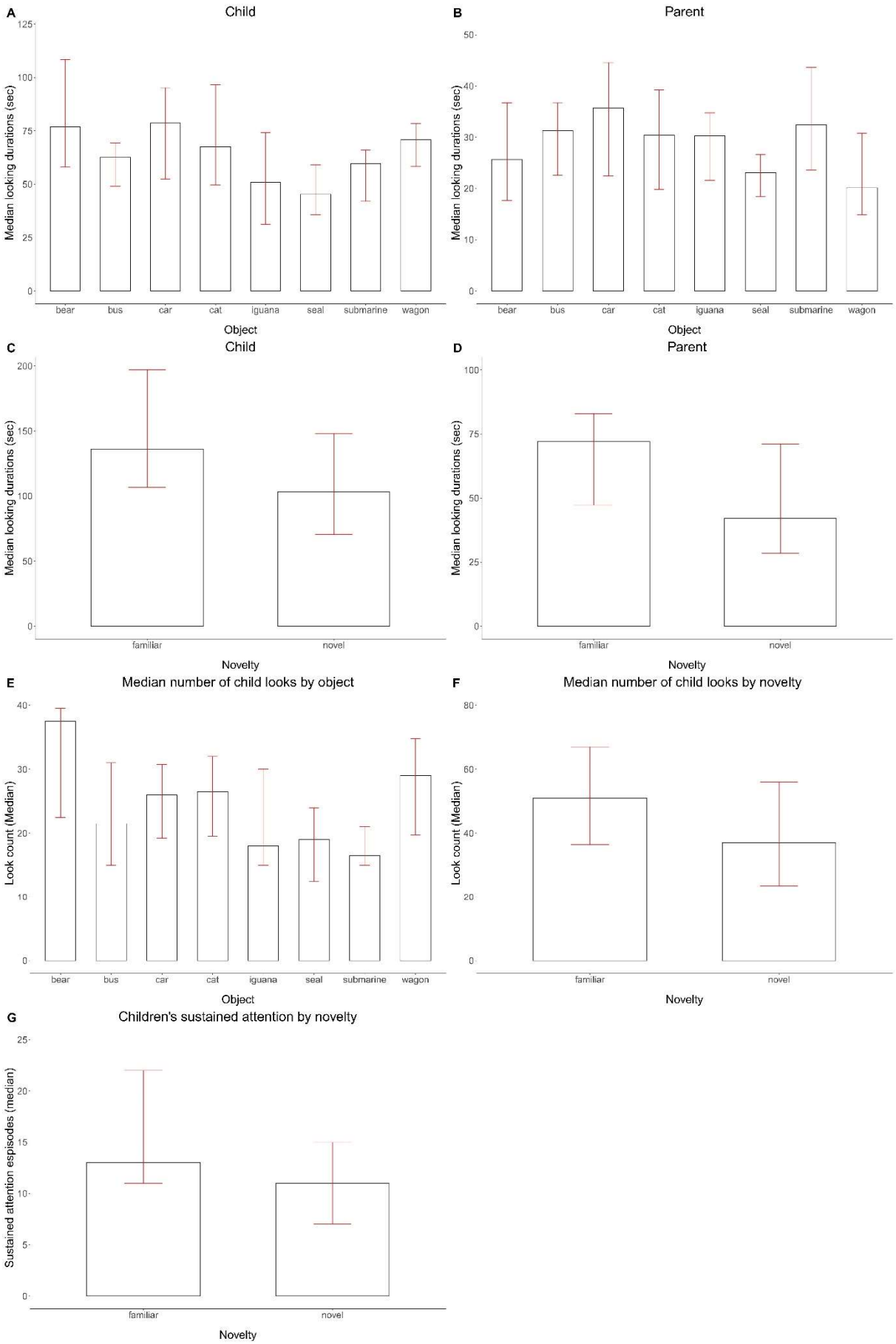
Results

Descriptive statistics

First, we report descriptive statistics for our data in the form of seven plots. We provide here descriptives for parent and child looking data, pooled both on the item-level and novelty (familiar versus novel).

Figure 3

Item-level descriptives (medians)


A
Child
Median looking durations (sec)
125
100
75
50
25
0
bear
bus
car
cat
iguana
seal
submarine
wagon
Object
B
Parent
Median looking durations (sec)
50
40
30
20
10
0
bear
bus
car
cat
iguana
seal
submarine
wagon
Object
C
Child
Median looking durations (sec)
200
150
100
50
0
familiar
novel
Novelty
D
Parent
Median looking durations (sec)
100
75
50
25
0
familiar
novel
Novelty
E
Median number of child looks by object
Look count (Median)
40
30
20
10
0
bear
bus
car
cat
iguana
seal
submarine
wagon
Object
F
Median number of child looks by novelty
Look count (Median)
80
60
40
20
0
familiar
novel
Novelty
G
Children's sustained attention by novelty
Sustained attention episodes (median)
25
20
15
10
5
0
familiar
novel
Novelty

Note: the red error bars represent 25% and 75% quantiles. We depict medians and quartiles instead of means and SDs since these cope much better with skewed distributions.

Hypothesis 1: Caregivers lead more episodes of JA when the dyads are playing with novel objects as opposed to familiar objects

We evaluated whether children or their caregivers are more likely to lead episodes of JA to novel and familiar objects. We found that parents led 104 JA episodes for familiar objects compared to 110 for novel objects; while children led 812 JA episodes for familiar objects compared to only 604 for novel objects. To further examine this pattern, we fitted a Generalised Linear Mixed Model (GLMM) with binomial error distribution and logit link function. Our response variable was a binary variable coding each instance of JA for whether it was parent-led (0) or child-led, (1; for further information, see Supplementary Materials S2) and our predictor variable *obj.novelty* coded for the familiarity of the object (novel or familiar). We also controlled for the order of JA episodes for each object (hereafter *JA order*, i.e., n^{th} JA episode between caregiver and child for a given object), as we wanted to account for the potential decrease in overall attention and engagement of the dyads with the toy play as time progressed. In keeping with the pattern described above, we found a significant positive intercept and a significant effect of *obj.novelty* on the leader of episodes of JA (hereafter **Model 1a**). These results suggest that children typically lead more episodes of JA than adults overall and and that children were more likely to lead episodes of JA involving familiar objects relative to novel objects (Table 2, Figure 4).

Table 2

Results of the GLMM examining JA leader (by mutual gaze).

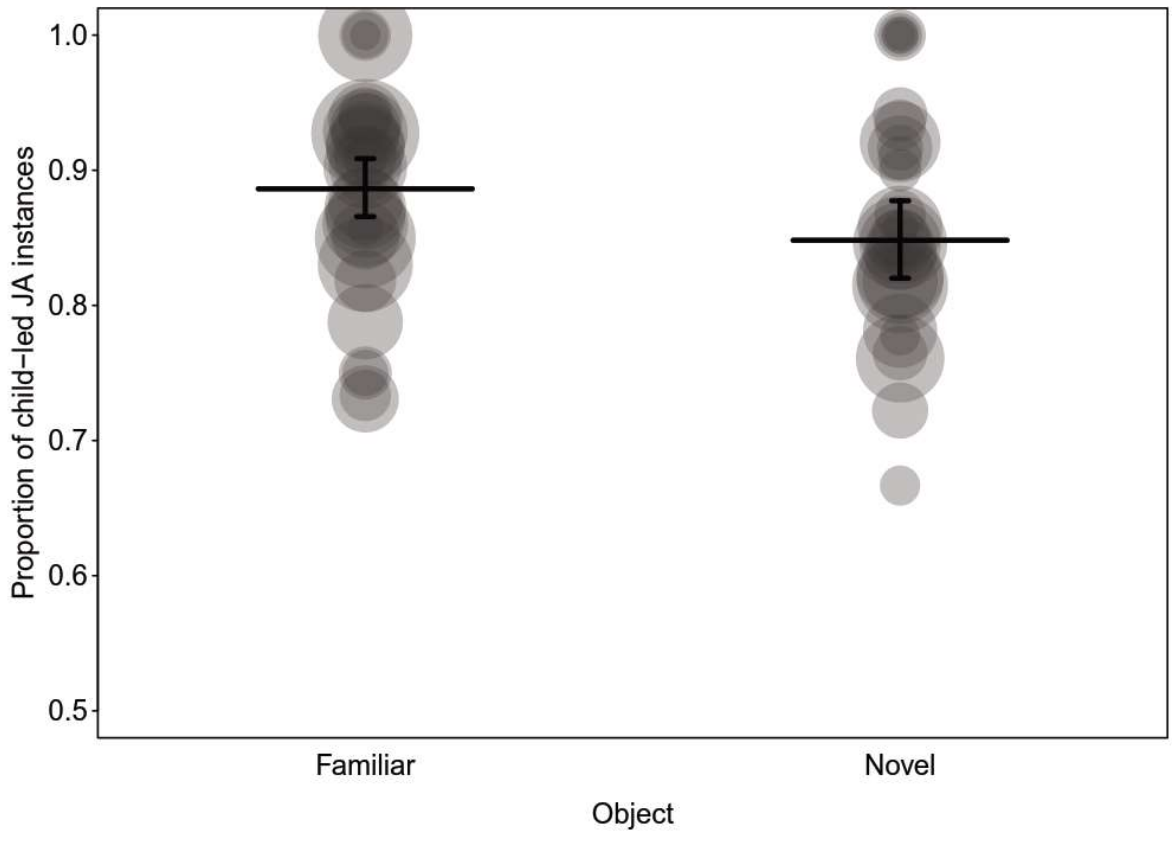
				Leader			
Predictors	*Estimates*	*std.Error*	χ^2	*p*	*CI*	*min*	*max*

(Intercept)	2.045	0.107		<.001*	1.864 – 2.241	1.983	2.101
obj.novelty	-0.330	0.150	3.822	.051*	-0.597 – -0.075	-0.409	-0.182
JA order	0.029	0.086	0.131	.717	-0.119 – 0.208	-0.038	0.071

The predictor JA order was z-transformed to a mean of zero and a standard deviation (SD) of one; original mean and sd were 9.75 and 7.71. In this and subsequent tables, indicated are estimates, together with standard errors, 95% confidence intervals, significance tests, as well as minimum and maximum of model estimates after excluding children and objects one at a time. P-values for the intercept were derived from the model summary. Obj.novelty was dummy coded with familiar being the reference level.

Figure 4

Proportion of child-led JA episodes (as defined by mutual gaze) during the play session separately for novel and familiar objects.



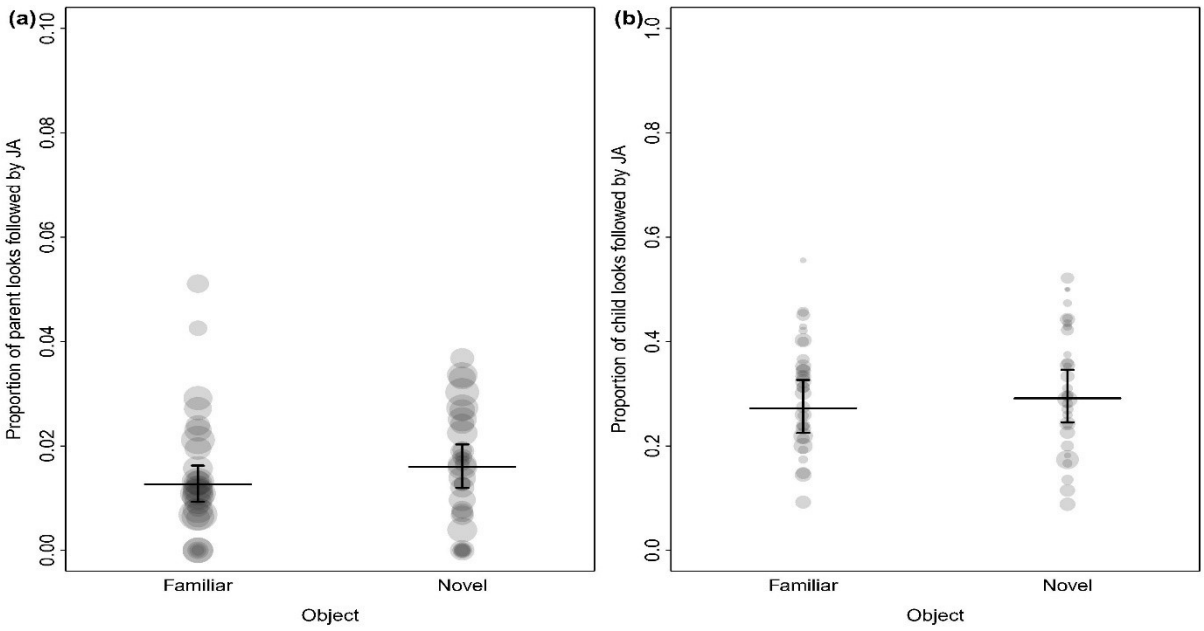
Note. The proportion of child-led instances of JA relative to the number of instances of JA overall, i.e., including child-led and parent-led JA episodes. In this and subsequent figures, for data visualisation purposes, each dot represents one participant or one dyad. The area of the dots depicts the number of observations per participant or dyad(range: 5 to 69) since different dyads had different number of instances of JA. The horizontal line depicts the fitted value, while error bars depict its 95% confidence limits.

Exploratory analysis examined the dynamics of the interaction further. First, we examined whether parent looks to novel or familiar objects were more likely to be followed by periods of joint attention (**Model 1b**). We fitted a GLMM with binomial error distribution and logit link function. Our response variable was a binary variable coding each parent fixation for whether this look was followed by JA (1) or not (0; for further information, see Supplementary Materials S2) while the predictor variable *obj.novelty* coded for the familiarity of the object. We found that parents' looks toward familiar objects were about equally likely to be followed by JA episodes compared to parents' looks toward novel objects (p = .136). (Figure 5a; Table in Supplementary S3).

Secondly, we examined whether child looks to novel or familiar objects were more likely to be followed by periods of joint attention (**Model 1c**). We fitted GLMM with binomial error distribution and logit link function, with response and predictor variables as above using the data for child fixations (for further information, see Supplementary Materials S2). We found that child looks toward familiar objects were about equally likely to be followed by JA episodes compared to child looks toward novel objects (p = .537) (see Figure 5b; Table in Supplementary S3).

Figure 5

Proportion of parent (a) and child (b) looks toward objects followed by JA instances as a function of familiarity of objects to the child.



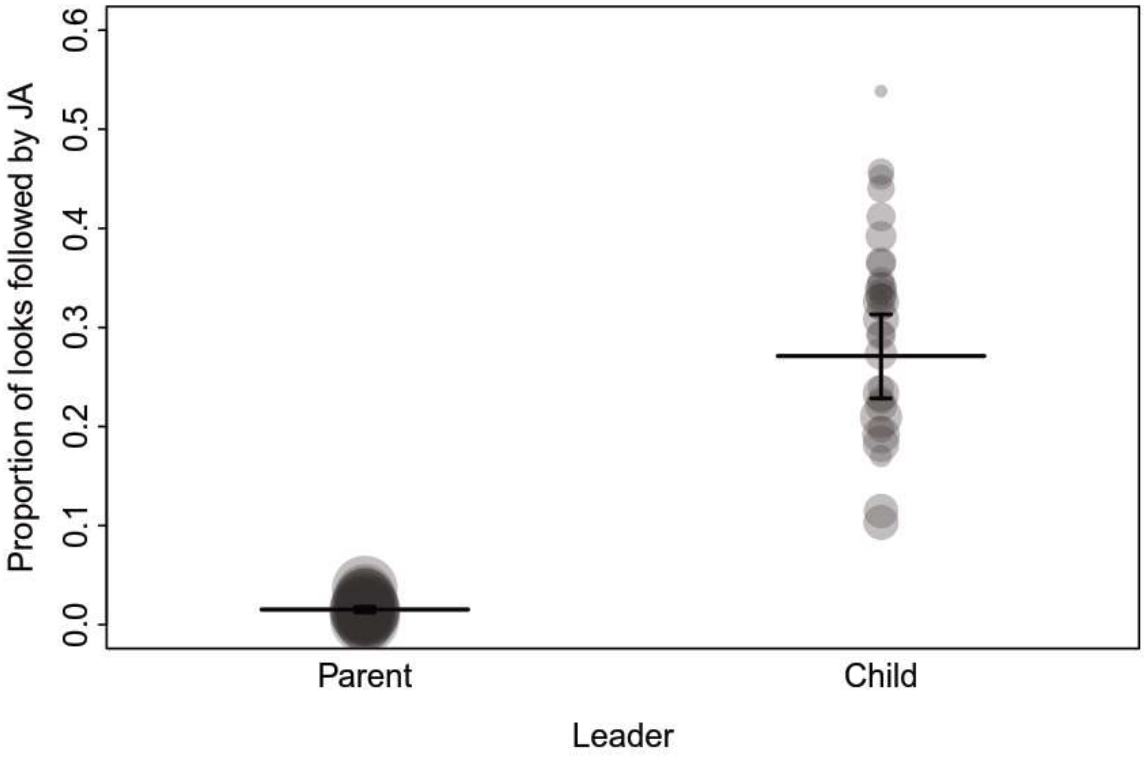
Note: Y-axes limits are different on (a) and (b). The proportion of looks is the number of looks followed by an instance of JA divided by the number of fixations overall. The number of observations per participant and object ranged from 41 to 438 in (a) and from 4 to 144 in (b).

Next, we examined whether children lead more episodes of JA because parents were more likely to follow their child's attention than children were likely to follow their parents' attention. That is, we examined all parent and children's looks towards an object and analysed whether parent or children's fixations were more likely to be followed by a JA episode. (**Model 1d**). We found that 823 children's looks were followed by a JA episode relative to 181 parent looks that were followed by a JA episode. We examined this pattern using a GLMM with binomial error distribution and logit link function. Our response variable was a binary variable coding each independent look for whether it was followed by JA (1) or not (0; for further information, see Supplementary Materials S2), while the predictor variable coded who looked (parent or child). We found that children's looks toward objects were more likely to be followed by episodes of JA compared to parent looks towards objects ($p < .001$). Thus, while children actively fixated objects in their environment, parents actively followed their

child's attention and created moments of JA. Children, on the other hand, very rarely followed their parent's visual attention to establish joint attention to objects in their environment (Figure 6; Table in Supplementary S3).

Figure 6

Proportion of looks that were followed by a period of joint attention, as a function of the leader of joint attention (i.e., parent or child).



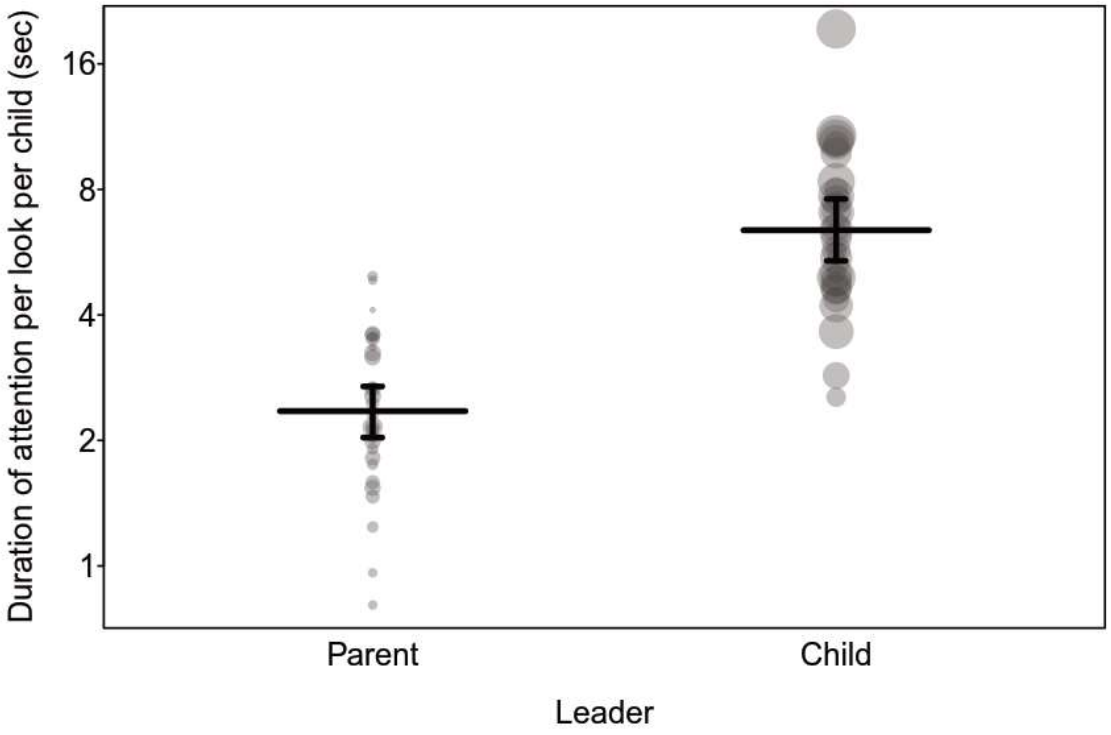
Note. The number of observations looks per parent or child participant ranged from 13 to 555. The elevation of the dots depicts the proportion of looks of each participant was were followed by an instance of JA relative to that participant's total number of looks overall.

Finally, we examined whether child-led or parent-led JA episodes were more likely to be followed by the child paying increased attention to the object (**Model 1e**). To do so, we examined the duration of the child's visual attention to the object that was the focus of the Joint Attention episode, for every JA instance. We fitted a Linear Mixed Model (LMM) with Gaussian error distribution and identity link function. Our response variable was the duration of time the child spent looking at an object following a period of JA to that object for each

episode of JA (for further information, see Supplementary Materials S2), while the leader (parent of child) of the episode of JA towards that object was the predictor variable. We found that child-led JA instances toward objects were more likely to lead to the child paying increased attention to the object (Figure 7; Table in Supplementary S3), relative to parent-led JA instances ($p < .001$).

Figure 7

Children's duration of attention to objects that were preceded by a JA episode as a function of leader of JA episode (i.e., parent or child).



Note. Depicted are the average latencies per participant (range of number latencies per participant: 1 to 81).

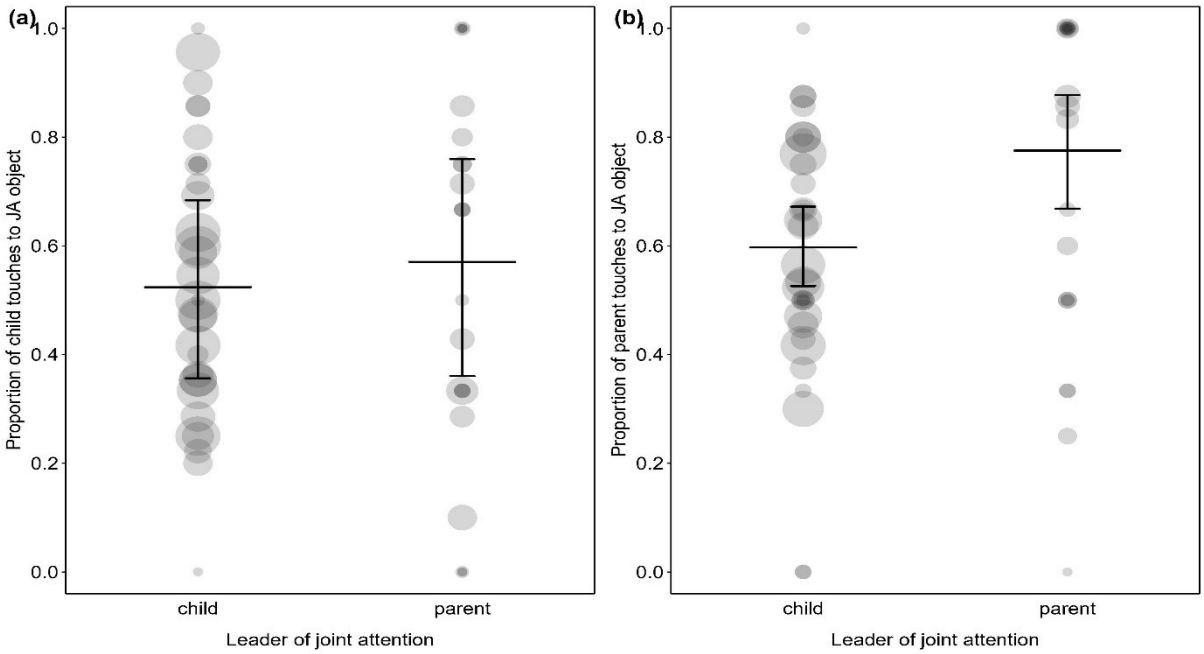
We also explored parent and child manipulations of objects, to understand the multimodality of parent-child interactions during the JA episodes. First, we looked at whether children were more likely to dextrally manipulate objects during child-led JA compared to parent-led instances of JA (**Model 1f**). We fitted a GLMM with binomial error distribution and logit link function. Our response variable was a binary variable coding for each Joint

attention episode for whether the object of the JA was touched by the child (1) or not (0; for further information see Supplementary Materials S2). Our predictor variable coded for who led the JA episode (*leader* of JA, parent or child). We found that children were about equally likely to manipulate objects that were the focus of child-led JA and parent-led JA ($p = .584$) (Figure 8a; Table in Supplementary S3).

Next, we examined whether parents were more likely to manipulate target objects (*parent touches*, response variable) during parent-led instances of JA compared to child-led instances of JA (**Model 1g**). We fitted a GLMM which was identical to model 1g with the exception that the response variable indicated whether the parent manipulated the object of the JA episode (for further information, see Supplementary Materials S2). We found that parents were more likely to manipulate the object that was the focus of parent-led JA relative to manipulating objects that were the focus of child-led JA ($p < .01$) (Figure 8b; Table in Supplementary S3).

Figure 8

Proportion of JA episodes with touches to objects in focus by the child (a) or parent (b) as a function of leader of JA episode (i.e., parent or child).



Note: The proportion of touches is the number of touches to the object of the focus of JA relative to of all JA episodes. The number of JA episodes per participant ranged from 1 to 25 for (a) and from 1 to 26 for (b).

Hypothesis 2: Caregivers label novel objects more frequently when the setting includes novel and familiar objects and caregivers are familiar with the novel objects.

Here we examined whether parents were more likely to label novel as compared to familiar objects (**Model 2**). On average, parents named familiar objects 16.06 times (*SD* = 9.74), compared to naming novel objects an average of 8.68 times (*SD* = 5.41). We fitted a GLMM with binomial error distribution and logit link function (for further information, see Supplementary Materials S2). Our response variable indicated the number of labelling instances of novel and familiar objects by the parents. We also added the predictor variable of sustained attention to control for its effects. We found a significant negative intercept, i.e., that caregivers labelled novel objects less likely than familiar objects during play (Table 3, Figure 9).

Table 3

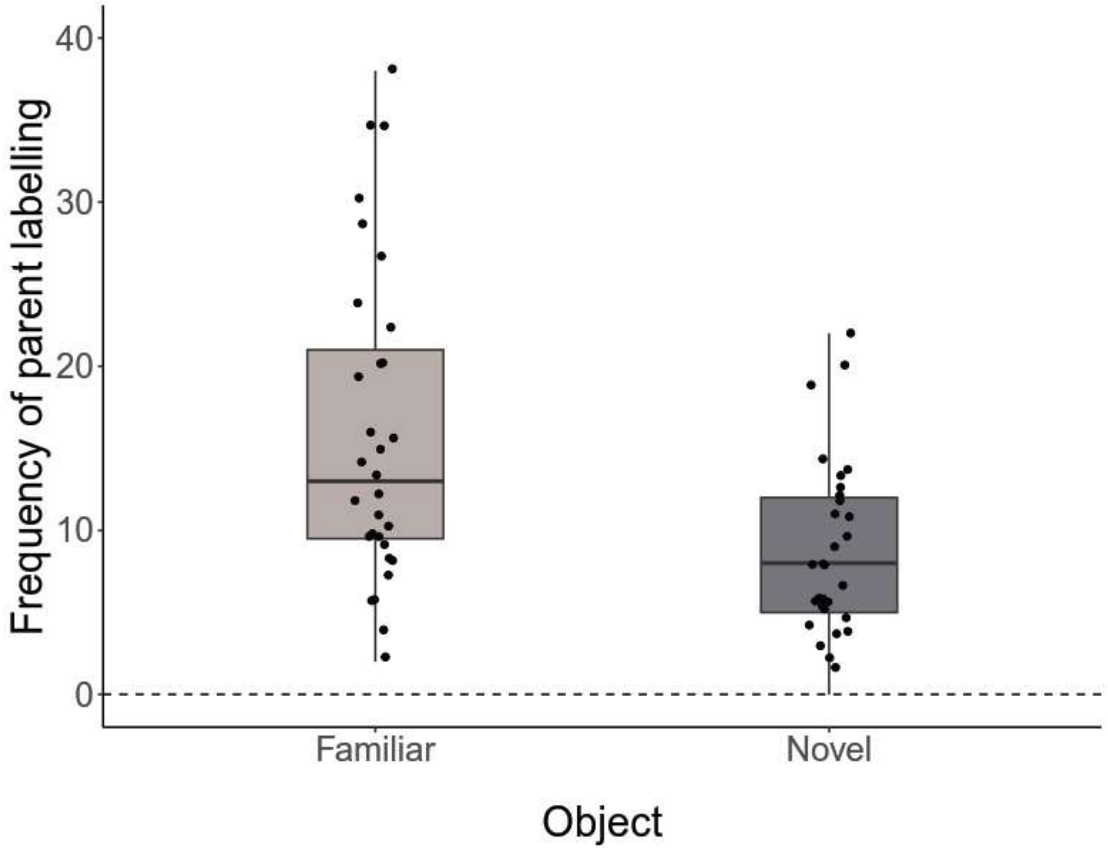
Results of the GLMM examining labelling

				Labelling		
Predictors	*Estimate*	*std.Error*	*p*	*CI*	*min*	*max*
(Intercept)	-0.485	0.096	<.001*	-0.681 – -0.309	-0.527	-0.446
Sustained.att	0.151	0.103	0.145	-0.045 – 0.345	0.189	0.283

The predictor Sustained.att was z-transformed to a mean of zero and a standard deviation (sd) of one; original mean and sd were 74.71 and 39.2.

Figure 9

Number of instances of parent object labelling by familiarity of the object to the children.



Note. The dots show individual observations while horizontal black lines and boxes depict the median and quartiles of the response variable.

Hypothesis 3: Children show improved recognition of novel word-object associations for objects they showed greater sustained attention towards and whose labels they heard more frequently

Here, we only examined a subset of our original data since some children already knew one or more of the novel objects requiring us to exclude data from these dads from the analysis. We examined whether children's recognition of novel objects was modulated by the frequency with which they heard the labels for these objects as well as children's sustained attention toward the objects. We examined two indices of novel word-object association, namely children's proportion of looking towards the labelled object as well as their pointing or handling the target object. Thus, we examined whether children's proportion of looking toward the named novel object (PTL, see Coding section; also response variable here) or their

pointing towards the named novel object varied as a function of parent's frequency of labelling that object and children's sustained attention toward the objects during play (predictor variables). We first report the results of children's recognition as indexed by gaze behaviour (**Model 3a**), followed by recognition measured by pointing behaviour (**Model 3b**).

For Model 3a, we fitted a GLMM with beta error distribution and logit link function, with the response variable being *Proportion of target looking* ranging from 0 to 1 (for further information, see Supplementary Materials S2). Apart from the two predictors mentioned above, we also included a main effect of time window (pre-label versus post-label; see Coding for more details), and interactions between window and the two predictor variables. There was no significant difference between the full and null models excluding the key predictor variables ($\chi^2(5) = 4.73$, $p = .45$). We also found no significant interactions between window and the other fixed effects predictors. A reduced model excluding the interaction with window also did not find a significant effect of any of the three predictors (see Supplementary materials S3 for the reduced model). Thus, we found no evidence that children recognized the novel word-object associations they were introduced to in the play phase or that their recognition was impacted by caregivers' labelling behavior or children's sustained attention towards the objects (Table 4).

Table 4

Results of GLMM examining recognition (as gaze behaviour)

						PTL		
Predictors	*Estimates*	*std.Error*	*CI*	χ^2	*df*	*p*	*min*	*max*
(Intercept)	0.068	0.337	-0.535 – 0.759				-0.241	0.477

Sustained.att	0.209	0.164	-0.106 – 0.543			.202	0.104	0.292
Prop.label	-0.515	0.596	-1.757 – 0.553			.388	-1.553	0.168
Window	0.145	0.481	-0.807 – 1.084			.763	-0.350	0.486
*Sustained.att*window*	-0.216	0.232	-0.689 – 0.252	0.867	1	.352	-0.331	-0.042
*Prop.label *window*	0.461	0.839	-1.081 – 2.080	0.302	1	.583	-0.343	1.707

The predictor sustained.att was z-transformed to a mean of zero and a standard deviation (SD) of one; original mean and sd were 41.03 and 24.94.

A second model examined children's recognition of the novel objects as indexed by their pointing behaviour (**Model 3b**; Supplementary materials S3). Children pointed to the labelled object 24 times and did not point or pointed to an incorrect object 17 times. We fitted a GLMM with binomial error distribution and logit link function, with the response variable of *pointing* coded as 0 for incorrect or no pointing and 1 for correct pointing toward labelled object (for further information, see Supplementary Materials S2) and the same predictor variables as in Model 3a with the exception of *window* and all the interactions with it. Here too, we found no significant difference between the full and null model ($\chi^2(2) = 2.47$, $p = .29$) and no effect of the individual predictors in the model.

Discussion

The current study examined whether children's familiarity with objects in their environment influenced the structure of JA episodes, caregivers' labelling of novel and familiar objects, and children's subsequent recognition of the newly-introduced novel word-object associations. We found that episodes of JA were more likely to be led by children than their caregivers. This was due to parents robustly following their children's attention and attending to the object their children were actively fixating, while children were less likely to follow their parent's attention. Children appeared to be quite egocentric in their attention, displaying longer sustained attention to objects to which they initiated JA to relative to objects to which their parents initiated JA. Furthermore, the likelihood of children leading JA episodes was higher when the dyad attended to a familiar object relative to a novel object. Parents also labelled familiar objects more frequently than novel objects throughout the play session and this effect was not influenced by children's sustained attention to the different kinds of objects. Finally, we found no evidence that children recognised the newly-introduced novel objects or that their recognition was influenced by parent's labelling of novel objects or children's sustained attention toward the novel objects. In what follows next, we discuss the implications of our findings for the literature and future studies.

Children more likely to lead episodes of JA to familiar than novel objects

While studies with older children (24 to 37 months) suggested that children and adults both coordinate attention to objects (Chen et al., 2020), with younger infants, we found that children were more likely to lead episodes of JA relative to their caregivers. Separate exploratory analyses suggested that the predominance of child-led instances of JA was due to parents actively following their child's attention, while children appeared quite oblivious to their caregivers' attention. Furthermore, children also displayed greater sustained attention on an object once they initiated an instance of JA towards this object, i.e., children appeared to be in a state of inertia in terms of attention. Taken together, these findings suggest that

children *lead* parent-child interactions, but also that parents *followed* their children in these interactions. This result is consistent with research highlighting the active role that children play in choosing what, when, and from whom to elicit information from (see Mani & Ackermann, 2018 for a review). Indeed, recent work examining both caregiver and child behaviour in caregiver-child interactions suggest that children's interests shape the quality of caregiver-child interactions, with higher quality, more engaging, dynamic interactions concerning objects of greater interest to the child (Madhavan & Mani, 2024). Parents also cater to children's preferences, by engaging in specific activities (e.g., Baroody & Dobbs-Oates, 2011; Lukie et al., 2014) or providing them with objects they prefer to engage with (DeLoache et al., 2007; Leibham et al., 2005). In keeping with such findings, here, we showed that children steered their parents' attention toward objects they want to focus upon, and subsequently maintain sustained attention to these objects. Thus, children may have a preference for engaging in interactions that align with their attentional preferences. At the same time, parents may indulge such preferences, for both the parents and children to come together to create optimal interactions.

We also explored parents and children's object manipulations during episodes of JA. While children's object manipulations were not predicted by who was leading the JA episode, parents' object manipulations were – parents were more likely to touch and manipulate an object when they led the episode of JA to this object relative to when their child led the episode of JA to the object. This result is in line with previous research (Goupil et al., 2024), showing that parents are more likely to touch objects when they lead their child's attention to this object. Indeed, previous research suggested that there are multiple sensori-motor pathways to joint attention (Chen et al., 2020; Yu & Smith, 2017), which together with our findings suggest that parents handling objects may be a particularly effective cue for children

to pay visual attention to what their parent is handling, thereby establishing joint visual attention to the object.

In contrast to previous research, we found that children were more likely to lead instances of JA to a familiar relative to a novel object, in comparison to their caregivers. Note that this does not mean that parents lead more instances of JA to novel objects. Exploratory analyses suggested that parents' looks to novel and familiar objects were equally likely to be followed by an instance of JA. Similarly, children's looks to novel and familiar objects were equally likely to be followed by an instance of JA. Thus, children and their caregivers did not *independently* prioritise attention to familiar and novel objects. However, coordinated attention between children and their caregivers converged on child-led interactions with familiar objects relative to novel objects. Children lead more JA instances overall, and JA instances that were directed towards familiar objects – and led by children – were more frequent than JA instances directed towards novel objects.

To a certain extent, this finding stands in contrast to previous studies suggesting that children display a novelty bias in caregiver-child interactions. We note, however, that the focus of the current study, in terms of examining who leads instances of JA to familiar and novel objects, differs from previous studies. To this extent, our findings may reflect differences in the novelty bias as indexed by JA instances, relative to simply increased looking towards novel objects in play sessions (c.f., Chen et al., 2021). We suggest that differences in the design in terms of the more naturalistic mix of objects presented to children during the play session may explain differences between our findings and previous studies. In particular, we presented children with both novel and familiar objects in the same play session. The presence of familiar and novel objects in the same interaction may reduce differences in looking behaviour to novel objects, and incur a familiarity bias in terms of children initiating and parents' following their child's attention. Furthermore, the novel

objects were familiar to the parents, while unfamiliar to the children in the current study. When presented with entirely novel objects, parents may trigger different kinds of interactions with objects novel even to themselves. In contrast, parents in the current study may not have been aware that these objects were novel to their child. This may have led to a more naturalistic interaction between children and caregivers, where the novel objects were not highlighted in any way, potentially allowing the familiarity of the objects to the child to drive the interaction. The increased naturalness of the setting used here, we suggest, may allow us to uncover additional dynamics in caregiver-child play and learning.

Parent labelling of objects as a function of object novelty

Parents also labelled familiar objects more frequently than novel objects during play. This is in line with previous work examining parent labelling of objects during play; Chen and colleagues (2021) found that parents labelled familiar objects twice as often as novel objects. At first glance, this is a surprising finding. Indeed, we hypothesised that parents, with the aim of teaching children the novel word-object associations, would name novel objects more frequently, drawing children's attention toward the object. However, our results suggest that children showed an overwhelming focus toward familiar objects, and that parents followed children's lead to establish joint focus. Therefore, parents, acting on their children's attentional focus, may simply name and engage with the familiar objects that their children are looking at.

The results paint a very specific picture of parent-child interactions with regard to free-play: parents do not appear to direct or guide their child's attention to objects with the goal of imparting knowledge to them. Were this to be true, then parents ought to have led more JA instances when the object was novel to the child, and named these objects often to help children associate the novel words with the objects in the environment. In contrast, the child appears to hold the reins to the interaction in many ways: most joint attention instances

were led by the child, about thirty percent of all gazes initiated by the child lead to a joint attention instance, and the child decided for familiarity in such interactions.

Children's (lack of) recognition of newly introduced objects

We did not find any evidence that children learned the novel word-object associations introduced to them in the study. Therefore, caution is needed in examining how recognition varies as a function of parent labelling or sustained attention. Our first analysis utilised infants' looks towards the target object, upon hearing its name, as a measure of children's recognition of the novel word-object associations. One possible explanation for the null results obtained here is the limited sample that provided us with data for the novel word recognition task (n = 18). We also examined children's pointing behaviour as an index of their recognition of the novel objects. Here as well, we found no evidence that children recognised the novel word associations for the novel toys they were playing with only a few minutes before. While this analysis was similarly plagued by the sample size issues noted above, the result is in keeping with our results using looking behaviour as an index of word recognition and suggests that children likely did not learn the novel word-object associations.

This is surprising: indeed, in previous research with slightly older children, evidence suggested that parent-child interactions boost the learning of novel word-object associations (Madhavan & Mani, 2024). We explain this with regard to our previous speculation about the nature of the interaction. In the current study, parents and children engaged in a play task, where the instruction was to, "play naturally as at home". While the above-mentioned study (Madhavan & Mani, 2024), utilised shared book reading, and asked parents to "read as naturally as they would at home", it may be that the very nature of the task can steer the goals of the interaction. Indeed, parents may have felt a need to impart knowledge during a book reading session; i.e., the end goal of this task for them may centre around their child learning. In contrast, during the play sessions, the end goal of the task may centre around entertaining

and having fun with their child. Recent studies comparing toy play sessions and book reading indicate that shared book reading promotes more interactive talk; additionally, parents also tend to use a broader vocabulary during the interaction (Clemens & Kegel, 2021; Wu et al., 2024). Additionally, while there were no instructions on either task for the parents to name the novel objects a certain number of times, the books themselves mentioned the label of the novel objects four times, to ensure the mapping of the novel label to the object. In the current study, there was no such guarantee that the parent would even label the novel object, let alone a certain number of times. Indeed, there were a few cases in which the parent did not label one or more of the novel objects at all throughout the play session. Finally, the novel object recognition trials in the book reading study presented children with two objects on screen. In contrast, in the current study, children were required to answer the experimenter sitting in front of them, which may have made them more hesitant and shy, inhibiting their response to the experimenter's question. Taken together, we see multiple explanations for why we did not find evidence for successful novel word-object mapping in children in the current study.

Our study is not without its drawbacks. Our analyses with regard to parent labelling only concerned the frequency (or proportion) of labelling instances; indeed, we did not look at temporal dynamics of labelling alongside gaze or touch behaviour of the children. While this will no doubt shed more light on the micro-features of parent-child interactions, it would have been beyond the aims of this paper, and, therefore, was not examined. Our final sample sizes with regard to our analyses of novel object recognition were too small to reliably accept the lack of recognition (and effects of the predictors) shown. With the rich dataset this study has generated, our next steps would be to look into the temporal co- dynamics of parents' labelling, and children's gaze or touch on the corresponding target objects, and perhaps also look into how children's long-standing interests of certain objects steer these interactions.

To briefly summarize, our results suggest that children actively lead day-to-day interactions with their parents, and initiate most instances of joint attention between themselves and their parents during a play session. Children also appear to be consistently egocentric, with child-led instances of JA triggering sustained attention from the child more often than parent-led instances of JA. Children not only choose what to look at, and trigger their caregiver's attention to this object, but also appear to sustain their attention to their chosen object of interest. Surprisingly, most of these joint attention instances center around objects already familiar to children. This, in turn, may prompt parents to label familiar objects more frequently than novel objects also present in the environment. Thus, children appear to lead social interactions with their caregivers and also appear to ensure that such interactions centre on familiar objects. Indeed, we found no evidence that children recognised the newly-introduced word-object mappings. We attribute this to the smaller sample size for this analysis or parent's limited labelling of the novel objects, which, in turn, may be due to the nature of the play interaction. Thus, in certain kinds of interactions with their children, parents appear to be less pedagogically focussed, in terms of imparting knowledge to their child, and more focussed on the journey that the interaction takes them on.

List of Figure legends

Figure 1. *Pictures of the (a) familiar and (b) novel toys*

Figure 2. *Schematic diagram of the setup.*

Figure 3. *Proportion of child-led JA episodes (as defined by mutual gaze) during the play session separately for novel and familiar objects*.

Figure 4. *Proportion of parent looks toward objects followed by joint attention instances as a function of familiarity of objects to the child*.

Figure 5. *Proportion of child looks toward objects followed by joint attention instances as a function of familiarity of objects to the child.*

Figure 6. *Proportion of looks that were followed by a period of joint attention, as a function of the leader of joint attention (i.e., parent or child).*

Figure 7. *Children's duration of attention on objects that were preceded by a Joint Attention episode as a function of leader of JA episode (i.e., parent or child).*

Figure 8. *Proportion of touches to objects in focus of a JA episode as a function of leader of JA episode (i.e., parent or child).*

Figure 9. *Number of instances of parent object labelling by familiarity of the object to the children.*

References

Abney, D. H., Suanda, S. H., Smith, L. B., & Yu, C. (2020). What are the building blocks of parent–infant coordinated attention in free-flowing interaction? *Infancy*, *25*(6), 871–887. https://doi.org/10.1111/infa.12365

Akhtar, N., & Gernsbacher, M. A. (2007). Joint Attention and Vocabulary Development: A Critical Look. *Language and Linguistics Compass*, *1*(3), 195–207. https://doi.org/10.1111/j.1749-818X.2007.00014.x

Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge University Press.

Baldwin, D. A. (1995). Understanding the Link Between Joint Attention and Language. In *Joint Attention*. Psychology Press.

Baroody, A. E., & Dobbs-Oates, J. (2011). Child and parent characteristics, parental expectations, and child behaviours related to preschool children's interest in literacy. *Early Child Development and Care*, *181*(3), 345–359. https://doi.org/10.1080/03004430903387693

Begus, K., & Southgate, V. (2012). Infant pointing serves an interrogative function: Infant pointing serves an interrogative function. *Developmental Science*, *15*(5), 611–617. https://doi.org/10.1111/j.1467-7687.2012.01160.x

Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016). Lexical leverage: Category knowledge boosts real-time novel word recognition in 2-year-olds. *Developmental Science*, *19*(6), 918–932. https://doi.org/10.1111/desc.12343

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, *8*(6), 535–543. https://doi.org/10.1111/j.1467-7687.2005.00445.x

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social Cognition, Joint Attention, and Communicative Competence from 9 to 15 Months of Age. *Monographs*

of the Society for Research in Child Development, *63*(4), i–174. https://doi.org/10.2307/1166214

Chen, C., Castellanos, I., Yu, C., & Houston, D. M. (2020). What leads to coordinated attention in parent–toddler interactions? Children's hearing status matters. *Developmental Science*, *23*(3), e12919. https://doi.org/10.1111/desc.12919

Chen, C., Houston, D. M., & Yu, C. (2021). Parent–Child Joint Behaviors in Novel Object Play Create High-Quality Data for Word Learning. *Child Development*, *92*(5), 1889–1905. https://doi.org/10.1111/cdev.13620

Clemens, L. F., & Kegel, C. A. T. (2021). Unique contribution of shared book reading on adult-child language interaction. *Journal of Child Language*, *48*(2), 373–386. https://doi.org/10.1017/S0305000920000331

Cleveland, A., Schug, M., & Striano, T. (2007). Joint attention and object learning in 5- and 7-month-old infants. *Infant and Child Development*, *16*(3), 295–306. https://doi.org/10.1002/icd.508

Custode, S. A., & Tamis-LeMonda, C. (2020). Cracking the code: Social and contextual cues to language input in the home environment. *Infancy*, *25*(6), 809–826. https://doi.org/10.1111/infa.12361

DeLoache, J. S., Simcock, G., & Macari, S. (2007). Planes, trains, automobiles--and tea sets: Extremely intense interests in very young children. *Developmental Psychology*, *43*(6), 1579–1586. https://doi.org/10.1037/0012-1649.43.6.1579

Goupil, L., Dautriche, I., Denman, K., Henry, Z., Marriott-Haresign, I., & Wass, S. (2024). Leader–follower dynamics during early social interactions matter for infant word learning. *Proceedings of the National Academy of Sciences*, *121*(38), e2321008121. https://doi.org/10.1073/pnas.2321008121

Horst, J. S., & Samuelson, L. K. (2008). Fast Mapping but Poor Retention by 24-Month-Old Infants. *Infancy*, *13*(2), 128–157. https://doi.org/10.1080/15250000701795598

Lawson, K. R., Parrinello, R., & Ruff, H. A. (1992). Maternal behavior and infant attention. *Infant Behavior and Development*, *15*(2), 209–229. https://doi.org/10.1016/0163-6383(92)80024-O

Leibham, M. E., Alexander, J. M., Johnson, K. E., Neitzel, C. L., & Reis-Henrie, F. P. (2005). Parenting behaviors associated with the maintenance of preschoolers' interests: A prospective longitudinal study. *Journal of Applied Developmental Psychology*, *26*(4), 397–414. https://doi.org/10.1016/j.appdev.2005.05.001

Lukie, I. K., Skwarchuk, S.-L., LeFevre, J.-A., & Sowinski, C. (2014). The Role of Child Interests and Collaborative Parent–Child Interactions in Fostering Numeracy and Literacy Development in Canadian Homes. *Early Childhood Education Journal*, *42*(4), 251–259. https://doi.org/10.1007/s10643-013-0604-7

Madhavan, R., & Mani, N. (2024). The quality of caregiver–child interaction is predicted by (caregivers' perception of) their child's interests. *Royal Society Open Science*, *11*(4), 231677. https://doi.org/10.1098/rsos.231677

Mani, N., & Ackermann, L. (2018). Why Do Children Learn the Words They Do? *Child Development Perspectives*, *12*(4), 253–257. https://doi.org/10.1111/cdep.12295

Mason, G. M., Kirkpatrick, F., Schwade, J. A., & Goldstein, M. H. (2019). The Role of Dyadic Coordination in Organizing Visual Attention in 5-Month-Old Infants. *Infancy*, *24*(2), 162–186. https://doi.org/10.1111/infa.12255

McQuillan, M. E., Smith, L. B., Yu, C., & Bates, J. E. (2020). Parents Influence the Visual Learning Environment Through Children's Manual Actions. *Child Development*, *91*(3). https://doi.org/10.1111/cdev.13274

Morales, M., Mundy, P., Delgado, C. E. F., Yale, M., Messinger, D., Neal, R., & Schwartz, H. K. (2000). Responding to Joint Attention Across the 6- Through 24-Month Age Period and Early Language Acquisition. *Journal of Applied Developmental Psychology*, *21*(3), 283–298. https://doi.org/10.1016/S0193-3973(99)00040-4

Reid, V. M., & Striano, T. (2005). Adult gaze influences infant attention and object processing: Implications for cognitive neuroscience. *European Journal of Neuroscience*, *21*(6), 1763–1766. https://doi.org/10.1111/j.1460-9568.2005.03986.x

Suanda, S. H., Barnhart, M., Smith, L. B., & Yu, C. (2019). The Signal in the Noise: The Visual Ecology of Parents' Object Naming. *Infancy*, *24*(3), 455–476. https://doi.org/10.1111/infa.12278

Suanda, S. H., Smith, L. B., & Yu, C. (2016). The Multisensory Nature of Verbal Discourse in Parent–Toddler Interactions. *Developmental Neuropsychology*, *41*(5–8), 324–341. https://doi.org/10.1080/87565641.2016.1256403

Suarez-Rivera, C., Smith, L. B., & Yu, C. (2019). Multimodal parent behaviors within joint attention support sustained attention in infants. *Developmental Psychology*, *55*(1), 96–109. https://doi.org/10.1037/dev0000628

Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why Is Infant Language Learning Facilitated by Parental Responsiveness? *Current Directions in Psychological Science*, *23*(2), 121–126. https://doi.org/10.1177/0963721414522813

Tamis-LeMonda, C. S., Kuchirko, Y., & Suh, D. D. (2018). Taking Center Stage: Infants' Active Role In Language Learning. In M. M. Saylor & P. A. Ganea (Eds.), *Active Learning from Infancy to Childhood* (pp. 39–53). Springer International Publishing. https://doi.org/10.1007/978-3-319-77182-3_3

Tamis-LeMonda, C. S., Kuchirko, Y., & Tafuro, L. (2013). From Action to Interaction: Infant Object Exploration and Mothers' Contingent Responsiveness. *IEEE Transactions on Autonomous Mental Development*, *5*(3), 202–209. https://doi.org/10.1109/TAMD.2013.2269905

Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Malle, B. F., & Morgan, J. L. (2015). Attention to the mouth and gaze following in infancy predict language development. *Journal of Child Language*, *42*(6), 1173–1190. https://doi.org/10.1017/S0305000914000725

Tomasello, M. (1995). Joint Attention as Social Cognition. In *Joint Attention*. Psychology Press.

Tomasello, M., & Farrar, M. J. (1986). Joint Attention and Early Language. *Child Development*, *57*(6), 1454–1463. https://doi.org/10.2307/1130423

Wu, S.-C., Tzeng, O. J. L., & Wang, S. (2024). Mothers used wider vocabulary and talked to their six-month-old infants more during shared book reading than when they played with toys. *Acta Paediatrica*, *113*(1), 84–90. https://doi.org/10.1111/apa.17004

Wu, Z., & Gros-Louis, J. (2015). Caregivers provide more labeling responses to infants' pointing than to infants' object-directed vocalizations. *Journal of Child Language*, *42*(3), 538–561. https://doi.org/10.1017/S0305000914000221

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*(2), 244–262. https://doi.org/10.1016/j.cognition.2012.06.016

Yu, C., & Smith, L. B. (2013). Joint Attention without Gaze Following: Human Infants and Their Parents Coordinate Visual Attention to Objects through Eye-Hand Coordination. *PLoS ONE*, *8*(11), e79659. https://doi.org/10.1371/journal.pone.0079659

Yu, C., & Smith, L. B. (2016). The Social Origins of Sustained Attention in One-Year-Old Human Infants. *Current Biology*, *26*(9), 1235–1240. https://doi.org/10.1016/j.cub.2016.03.026

Yu, C., & Smith, L. B. (2017). Multiple Sensory-Motor Pathways Lead to Coordinated Visual Attention. *Cognitive Science*, *41*(S1), 5–31. https://doi.org/10.1111/cogs.12366

Yu, C., Suanda, S. H., & Smith, L. B. (2019). Infant sustained attention but not joint attention to objects at 9 months predicts vocabulary at 12 and 15 months. *Developmental Science*, *22*(1), e12735. https://doi.org/10.1111/desc.12735