Behavioral/Cognitive

Selective Action Prediction in Infancy Depending on Linguistic Cues: An EEG and Eyetracker Study

M. Colomer,[1,2] K. Zacharaki,[1] and N. Sebastian-Galles[1]
[1]Center for Brain and Cognition (CBC), Universitat Pompeu Fabra, Barcelona 08018, Spain and [2]Department of Psychology, University of Chicago, Chicago, Illinois 60637

Humans' capacity to predict actions and to socially categorize individuals is at the basis of social cognition. Such capacities emerge in early infancy. By 6 months of age, infants predict others' reaching actions considering others' epistemic state. At a similar age, infants are biased to attend to and interact with more familiar individuals, considering adult-like social categories such as the language people speak. We report that these two core processes are interrelated early on in infancy. In a belief-based action prediction task, 6-month-old infants (males and females) presented with a native speaker generated online predictions about the agent's actions, as revealed by the activation of participants' sensorimotor areas before the agent's movement. However, infants who were presented with a foreign speaker did not recruit their motor system before the agent's action. The eyetracker analysis provided further evidence that linguistic group familiarity influences how infants predict others' actions, as only infants presented with a native speaker modified their attention to the stimuli as a function of the agent's forthcoming behavior. The current findings suggest that infants' emerging capacity to predict others' actions is modulated by social cues, such as others' linguistic group. A facilitation to predict and encode the actions of native speakers relative to foreign speakers may explain, in part, why infants preferentially attend to, imitate, and learn from the actions of native speakers.

Key words: **action prediction; EEG; eyetracker; infancy; μ rhythm; social categorization**

Significance Statement

This article investigates how information about social categories, such as the language people speak, influences action prediction in infancy. Six-month-old infants predicted the actions of an agent based on her knowledge about the location of an object, but only if she previously spoke the infants' native language rather than a foreign language. Previous studies found that infants are less likely to imitate and engage with the actions of foreign speakers. Here, we extend these findings by showing that the tendency to predict others' actions may be weakened when interacting with speakers of a foreign language, potentially due to an association of foreign languages with unexpected behaviors. This selectivity could drive infants' biases to attend to and learn from native speakers.

Introduction

A sensitivity to adult-like social categories such as language, race, and gender emerges early in infancy (Rhodes and Baron, 2019). One of the most salient social categories influencing infants' responses to others is the language people speak (Liberman et al., 2017). Infants prefer to imitate speakers of their native language relative to foreign speakers (Howard et al., 2015; Altınok et al., 2021), and they prefer to select or attend to objects, music, and food introduced by native relative to foreign speakers (Kinzler et al., 2007; Shutts et al., 2009; Soley and Sebastián-Gallés, 2015). However, it remains unclear why infants respond to others selectively depending on the language they speak.

To date, two main theories have been proposed to explain infants' selective responses to speakers of different languages. In one view, infants' preference for native speakers may be related to a preference for in-group members, guided by the principle of

Received July 12, 2023; revised Feb. 10, 2024; accepted Feb. 14, 2024.

Author contributions: M.C. and N.S-G. designed research; M.C. and K.Z. performed research; M.C. analyzed data; M.C. wrote the paper.

This research was supported by grants from the European Community's Seventh Framework Programme (FP7/2007–2013): European Research Council Grant agreement number 323961 (Under Control); the Spanish Ministerio de Ciencia, Innovación y Universidades (PID2021-123416NB-I00); and the Catalan Government (SGR 2021-00911; FI-9015-456763; FI_B2 00093; ICREA ACADEMIA 2019 Award). We thank Ildiko Kiraly and Katarina Begus for their discussions and suggestions; Andrea Riquelme who helped create the stimuli; Alice Drew and Chiara Santolin who proofread the manuscript; Valentina Ferrario who helped code the videos of participants; Xavier Mayoral and Silvia Blanch who provided technical support and helped improve the code to present stimuli; and all families and infants who participated in the experiments.

The authors declare no competing financial interests.

Correspondence should be addressed to M. Colomer at marc.colomer.ca@gmail.com.

https://doi.org/10.1523/JNEUROSCI.1301-23.2024

in-group loyalty (Kinzler et al., 2007; Buttelmann et al., 2013). In another view, the bias toward native speakers may be explained by a motivation to seek relevant information, which infants may expect to receive to a greater extent from the native than foreign speakers (Begus et al., 2016; Marno et al., 2016). These theories shed light on the motivations underlying infants' social preferences, but they do not specify the processes that may be modulated by linguistic group membership. In the current study, we investigated whether a distinct tendency to predict the actions of familiar relative to unfamiliar speakers could underlie infants' selective responses toward others. We hypothesized that infants' emerging ability to predict others' actions may be weakened when interacting with unfamiliar language speakers, as compared with speakers of their native language.

Language can indicate a common culture and mark the limits of shared cultural knowledge (Soley and Spelke, 2016). In fact, already at 2 years of age, toddlers relate a foreign language with nonconventional actions (Oláh et al., 2014). Even earlier in development, 6-month-old infants match other-race faces to unfamiliar languages (Uttley et al., 2013), and 12-month-old infants expect native language speakers to act prosocially, but they do not hold such expectations when presented with foreign language speakers (Pun et al., 2018). These studies do not directly prove that linguistic group influences infants' ability to reason about others' actions, but they suggest that infants associate foreign speakers with behaviors or physical attributes that differ from their existing knowledge of the world. Infants may expect native speakers to perform actions that are consistent with their knowledge about other agents, but they may be unsure of what foreign speakers will do next. Based on this idea, we hypothesized that infants would show a facilitation of predicting the actions of relatively more familiar over unfamiliar individuals. This facilitation may be present already at 6 months when infants begin to show a preference to attend to native speakers (Kinzler et al., 2007).

To address this hypothesis, we focused on a neural correlate of the so-called mirror system (Rizzolatti and Craighero, 2004). Starting from early infancy, motor areas are recruited in a similar way when humans observe and predict the actions of others, as well as when they perform the actions themselves (Southgate and Begus, 2013). Sensorimotor activation is visible in the electroencephalogram (EEG) as a desynchronization of the resting-state μ rhythm measured over sensorimotor scalp areas (Debnath et al., 2019). Based on our hypothesis, we would expect event-related μ desynchronization (μ-ERD) to a greater extent in anticipation of the actions of native speakers relative to the actions of foreign speakers. To induce action anticipation, we adapted a belief-based action prediction task by Southgate and Vernetti (2014) in which 6-month-olds were presented with an agent who performed an action or remained still depending on her (false) epistemic state. We closely followed the design of the original study, except that the agent presented herself before the false belief task. Critically, some infants heard her speaking their native language (Catalan or Spanish), and other infants heard her speaking a foreign language (German: German is prosodically dissimilar from the native language of the participants. Previous research has shown that even neonates can discriminate between rhythmically distinct languages (Nazzi et al., 1998)).

Materials and Methods

Open data

The stimuli, raw eyetracker data, final EEG data, and scripts for data cleaning and analysis can be found here: https://osf.io/rfuve/. The raw EEG data as well as the preprocessed time–frequency EEG data can be found here in the format of EEGLAB datasets (.set) and Matlab files (.mat), respectively: https://doi.org/10.34810/data708.

Experiments and hypothesis

Experiment 1

We analyzed μ-ERD as a neural marker of action prediction in infancy. We hypothesized that infants would show greater μ-ERD when anticipating the actions of native language speakers as compared with the actions of foreign language speakers (action trials). In addition, we expected no desynchronization of μ waves if the agent was not about to act (no-action trials).

Experiment 2

We analyzed infants' overt attention to the stimuli with an eyetracker. First, we tested if infants attended more to the videos of native versus foreign language speakers during the entire trial, which could influence how much information they had available to make online action predictions. We hypothesized that, in general, infants would attend similarly to the videos independently of the language that the person spoke. Second, we tested the hypothesis that, by the end of each trial, infants may lose interest faster if they expected the agent to remain still, rather than to perform an action. We expected such a difference particularly if infants were presented with a native language speaker, since this is the case in which we expected infants to generate online action predictions (for the action trials).

Participants

Participants were recruited by visiting maternity rooms at two private hospitals in Barcelona, Spain (Hospital Quirón and Clínica Sagrada Família), and through social media ads. All participants were healthy, full-term infants (>37 GW). The research reported in this manuscript was conducted in accordance with the principles expressed in the Declaration of Helsinki and approved by the local ethical committee (Parc de la Salut Mar, Barcelona). Written informed consent was obtained from the participants' caregivers before the experiment was conducted. The sample was determined based on a large effect size, based on Begus et al. (2016). A priori sample size calculation with G*Power given a Cohen's d of 0.8, $\alpha = 0.05$, and $\beta = 0.2$ found that we should run 42 participants per experiment (21 per cell) to see significant differences between infants in the native group and infants in the foreign group. This sample is similar or higher to other studies manipulating the influence of language familiarity on attention (Kinzler et al., 2007), selection of food (Shutts et al., 2009), and the readiness to learn from others (Begus et al., 2016).

Experiment 1 (EEG)

The analysis included 42 6- to 7-month-old infants, all of them monolingual (at least 90% exposure to their native language) except for 5 bilinguals (>30% exposure to another language). All the participants had either Catalan (28) or Spanish (14) as their primary native language, and none of them had been exposed to German before. Of these, 21 infants were assigned to the native group (M_{age} = 6 months 15 d; range, 5 m.24 d to 7 m.2 d; 11 females, 1 bilingual), and 21 infants were assigned to the foreign group (M_{age} = 6 months 14 d; range, 6 m.0 d to 7 m.4 d; 13 females, 4 bilinguals). Sixty-eight additional participants were tested but excluded from the analysis due to crying early in the study (15) and experimental error [18: no sound (5), software not working (8), bad placement of the net (5)] or because they did not contribute the minimum number of three trials per condition required (35) due to movement artifacts or fussiness.

Experiment 2 (eyetracker)

The analysis included 42 6- to 8-month-old infants, all monolinguals except for three bilinguals. All the participants had either Catalan (32) or Spanish (10) as their primary native language, and none of them had been exposed to German before. Twenty infants were assigned to the native group (M_{age} = 6 months 30 d; range, 6 m.1 d to 7 m.30 d; 10 females), and 22 infants were assigned to the foreign group (M_{age} = 6 months 29 d; range, 6 m.1 d to 7 m.30 d; 12 females, three bilinguals). Of the 42 participants, 21 infants represented a subgroup of the participants in Experiment 1 who were tested both with an EEG and an eyetracker (native group, 9; foreign group, 12; participants with valid EEG data, 9). Additionally, 21 infants (native group, 11; foreign group, 10) were tested using the same task as in Experiment 1, but without wearing

an EEG cap. We only recorded infants' looking behavior using an eyetracker. Data of 82 additional participants were excluded from the analysis due to bad calibration (25) or not contributing the minimum number of three trials per condition required (due to data loss or fussiness; 57). Note that this high rejection rate is due, in part, to the experimenter's criterion of prioritizing the quality of the EEG signal over the quality of the eyetracker data for participants tested with both methods simultaneously. Specifically, capping the participant, improving the channel impedance, and maintaining the infant calm before the study started were prioritized over calibrating the eyetracker and making sure that the infant was in an optimal position relative to the eyetracker. As a trade-off, a high number of participants tested with both methods were excluded from the eyetracker analysis (58).

Procedure and stimuli

Participants were tested in a sound-attenuated and darkened room at the Universitat Pompeu Fabra. Participants' behavior during the session was recorded using a Sony HDR-HC9E camera (temporal resolution, 25 frames/s). Infants sat on a caregiver's lap at ∼65 cm from a 23″ screen (1,920 × 1,080 pixels). Caregivers were asked to close their eyes when the videos started. Videos were projected onto the screen using Psychtoolbox-3 in Matlab (MathWorks). The experiment was structured in three phases presented in the following order (Fig. 1): familiarization, presentation, and test. The design and stimuli of each phase, except for the presentation, closely followed the design and stimuli of a previous study (Southgate and Vernetti, 2014).

Familiarization

Infants viewed four trials of 7 s in which an actress always held a true belief about the location of a ball. First, infants saw the actress reaching for and grasping a ball (A+O+). The actress appeared in the center of the screen in front of a box. When the lid of the box opened, a ball appeared from the right side of the screen and entered the box. Then the lid of the box closed, and the actress reached toward the box. After that, infants saw a trial in which the actress did not reach for the (absent) ball (A−O−). The trial started with the actress in the center of the screen in front of the box. When the lid of the box opened, the ball appeared inside the box and jumped out until disappearing at the right side of the screen. The lid of the box then closed, and the actress remained still until the end of the trial. The two familiarization trials were presented again in the same order. Each of the four trials was preceded by an attention getter including a circular flash looming at the center of the screen (3 s). Infants who attended <50% of time to the familiarization trials were excluded from the analysis (M_{native} = 83%; $M_{foreign}$ = 87%; see Manual coding section; infants included in the analysis attended at least once to each trial type).

Presentation

An actress appeared at the center of the screen telling a story while making eye contact with participants. For the infants in the native group, the agent spoke either Spanish (∼18 s) or Catalan (∼20 s), depending on the participants' dominant language. For the participants in the foreign group, the agent spoke German, an unfamiliar language for all infants (∼23 s). The actress was highly proficient in the three languages. The presentation trial was preceded by an attention getter scene (3 s). Infants who attended <50% of time to the presentation trial were excluded from the analysis (M_{native} = 88%; $M_{foreign}$ = 79%; see Manual coding section)(An ANOVA with the group as the between-subjects factor and percentage of attention as the dependent variable found a significant effect of the group ($F_{(1,40)} = 4.19$; $p = 0.047$). Infants attended longer at the presentation of the native language speaker).

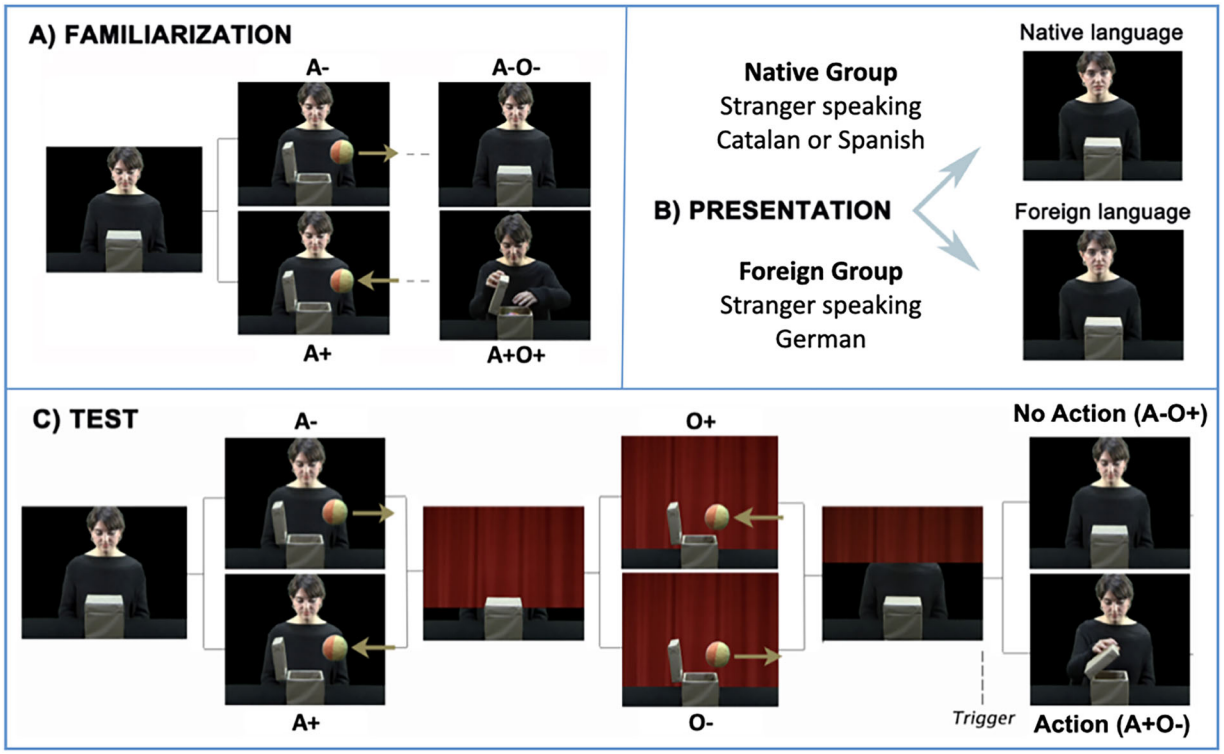

Figure 1. Structure of the videos. ***A***, First, infants saw the agent (A) reaching for a ball (O) when the ball jumped into a box (A+O+) and not moving when the ball jumped out of the box (A−O−). ***B***, Second, the agent introduced herself in the participant's native language or in a foreign language. ***C***, Third, the agent saw the ball jumping into (A+) or out of (A−) the box at the start of the video. Then a curtain came down hiding the agent, and the ball jumped either into (O+) or out of (O−) the box. When the curtain opened again, the agent remained still during 1,500 ms (anticipation period), and then she attempted to reach for the ball (A+O−; action trial) or remained still (A−O+; no-action trial) depending on her (false) belief about the ball's location. A successful prediction was associated with μ desynchronization prior to the agent's behavior (anticipation period) in the action trial (Experiment 1) and a different pattern of attention in the action versus no-action trials during the anticipation period, as measured by an eyetracker (Experiment 2). The WoA was defined as a window with onset at the "trigger" (when the curtain totally opened) and offset 500 ms later. Both trial types were randomly presented until infants became inattentive.

Test

Infants saw multiple times two trial types of 14 s that appeared in random order until they lost attention. The trials consisted of the same scene as in familiarization, except that we manipulated the actress' belief about the location of the ball. In the action trial, the actress first saw the ball jumping into the box (A+). Then, a curtain came down hiding the actress, and the ball jumped out of the scene (O−) unbeknown to the agent. The curtain opened, and the actress reappeared at the back of the display. She stayed 1,500 ms still (anticipation period), and then she reached toward the empty box (1,500 ms). In the no-action trial, the actress first saw the ball jumping outside of the box (A−) and then the ball jumping into the box (O+) when the curtain was closed, and she could not witness the change of the ball's location. When the curtain opened and the actress reappeared, she stayed still for 1,500 ms (anticipation period), and she remained still until the end of the trial (1,500 ms).

Data processing and analysis

Experiment 1

EEG acquisition. Participants' neural activity was recorded using continuous recordings of EEG with a Geodesic Sensor Net composed of 128 electrodes (Electrical Geodesics). Data were recorded with respect to the vertex electrode, sampled at 500 Hz, and stored on a computer disc. Triggers were sent through a parallel port from Matlab to the EEG amplifier and synchronized with the acquisition of EEG data. Triggers were sent at the start and end of each trial and each attention getter. A trigger was also sent in test trials at the point where the actress reappeared at the scene after the curtain had opened (test trigger). The test trigger marked the onset of the anticipation period (Time 0).

Manual coding. Participants' videos were manually coded using Datavyu (Datavyu Team, 2014) by a coder blind to the experimental manipulation. The coding was used to ensure that participants spent enough time looking at the stimuli during familiarization and presentation trials (instances in which babies did not pay attention to the videos during the familiarization and presentation were identified, and then a percentage of attention was calculated based on the length of trials and the total length of the not-looking instances) and to remove test trials in which participants did not pay attention or produced goal-directed actions (e.g., grasping) that could generate artifacts. Moreover, 30% of the videos were double coded and received ratings of good reliability for the not-looking instances [intraclass correlation (ICC) = 0.87] and moderate reliability for the goal-directed actions (ICC = 0.69).

EEG processing. Processing steps to remove artifacts and segment the data followed the Maryland Analysis of Developmental EEG (Debnath et al., 2020). Electrodes in the outer ring were removed due to potential poor connections in pediatric data, with 104 channels remaining. The continuous data were filtered (0.5–25 Hz) using a finite impulse response (FIR) filter provided within the FIRfilt plug-in of EEGLAB (Widmann et al., 2015). Next, we identified and removed bad EEG channels using the fully automated statistical thresholding for EEG artifacts (an EEG plug-in; Nolan et al., 2010). As suggested by Debnath et al. (2020), we applied a hybrid independent component analysis (ICA) approach to filter out non-neural artifacts including blinks, saccades, or EMG. In a copied version of the EEG dataset, a 1 Hz high-pass filter was applied, data were segmented into 1 s epochs, excessive EMG segments were removed, and then ICA was performed on the copied dataset. After ICA decomposition, artifactual independent components (ICs) were automatically identified using the ADJUST EEGLAB plug-in (Mognon et al., 2011). Then, ICs were transferred from the copied dataset to the original dataset, and artifactual ICs were removed from the original dataset. Within our window of interest (see below), if a participant produced a behavior that could influence the signal, as noted by manual video coding, the label of the event was marked to be excluded. These behaviors were coded using Datavyu (Datavyu Team, 2014) and included instances when the infant was performing a goal-directed action (e.g., grasping or reaching), crying, or not looking directly forward or the parents were interfering. The remaining data during the action observation block was segmented into epochs from 1,500 ms before to 1,500 ms after the 0 ms time-locked mark, corresponding to the end of the curtain's opening. A voltage threshold rejection (±220 μV) was applied on six frontal channels (E1, E8, E14, E21, E25, E34) to identify and remove additional eye artifacts. For all other channels, we interpolated data for noisy channels in each epoch. Epochs with >20% of interpolated channels were rejected. Before time–frequency analysis, the data were converted into a current source density (CSD) waveform using the CSD toolbox (Kayser and Tenke, 2006) to minimize the spread of brain activity due to volume conduction and improve the spatial resolution of the EEG signal.

EEG time–frequency analyses. Time–frequency analysis closely followed the analysis of Southgate and Vernetti (2014), including the selection of the channels of interest, baseline, test window of interest, baseline, and frequency range of interest. First, we applied a continuous wavelet transformation from 5 to 20 Hz to free-artifact epochs in each channel, using the function newtimef from EEGLAB toolbox [cycles = (3 0.5), nfreqs = 150, padratio = 8]. To eliminate distortion created by the wavelet transform, the first and last 500 ms of each segment were removed. A 400 ms period beginning 1,000 ms before the test trigger was selected as a baseline [i.e., −1,000 to −600 ms before the curtain opened completely (time 0)]. In both trial types, the event occurring within the baseline was the curtain opening before revealing the actress. Averaged activity within the baseline period was used to convert the time-varying signal to a decibel (dB) scale, using the formula $10*\log_{10}(\text{power}(t)/\text{power}(\text{averaged baseline}))$, where t refers to the signal at each time point. Time windows considered for the analyses included the baseline and a 500 ms window starting at the onset of the anticipation period (window of analysis, WoA). During the WoA, the actress remained stationary in both trial types, and infants could potentially predict her subsequent behavior. We analyzed activity across the frequency range from 5 to 7 Hz (μ) in left sensorimotor areas (electrodes 30, 36, 37, and 42). We focused on left central electrodes (electrodes around C3 based on the 10–20 system) given that, in infants, motor activation during action prediction has been typically localized over left rather than right sensorimotor areas (Southgate et al., 2010; Southgate and Begus, 2013; Southgate and Vernetti, 2014). Average event-related spectral power (ERSP) data for each participant in the baseline and the WoA were calculated by taking the mean across trials, electrodes, and frequency range of interest. To plot the results, we grand-averaged the data across participants. Southgate and Vernetti (2014) found sensorimotor activation reflecting infants' prediction of the actions of the agent when analyzing participants with a minimum of three free-artifact trials per condition, a common criterion in studies of action mirroring (Debnath et al., 2019; Chung et al., 2022). Thus, infants with less than three valid artifact-free trials in either condition were excluded. The native group contributed with an average of 6 trials in the action trials (range, 3–13) and 6.2 trials in the no-action trials (range, 3–14). The foreign group contributed with an average of 5.6 trials in the action trials (range, 3–11) and 5.5 trials in the no-action trials (range, 3–11).

Experiment 2

Eyetracker processing. A Tobii X120 Eye Tracker was used in stand-alone mode and with a sampling rate of 120 Hz. Eyetracking data were saved with a set of triggers that defined which part of the experiment was being presented to participants, allowing us to know when each trial started and ended (same triggers as used in Study 1). Prior to the beginning of the experiment, a five-point calibration was carried out. Participants who did not provide accurate data for at least four points were rejected from analysis.

Eyetracker analysis. Data from the right and left eyes were averaged when both eyes had valid data. In cases in which only one eye provided valid data, we used the signal from the corresponding valid eye. Eyetracking data was segmented across trials. Trials with <50% of valid data were excluded from the analysis. Following our previous criteria with EEG data, participants who contributed to less than three trials per trial type were excluded from the analysis. Using the temporal dynamics of gaze x and gaze y coordinates, we then calculated the amount of time in which infants looked at four areas of interest (AOIs;

see Extended Data Fig. 4-1): an area including the entire video, an area including the agent's face, an area including the box, and an area including the ball trajectory. Before statistical tests, data of each participant for each trial type were averaged across valid trials. The native group contributed with an average of 6 trials in the action trials (range, 3–12) and 6.4 trials in the no-action trials (range, 3–11). The foreign group contributed with an average of 6.45 trials in the action trials (range, 3–11) and 6.41 trials in the no-action trials (range, 3–12).

Statistical analysis plan

Experiment 1

To test if infants showed greater μ-ERD when predicting the actions of native over foreign language speakers (but not when the agents remained still), we ran a two-factor mixed ANOVA with trial type (action vs no action) as the within-subjects factor, linguistic group (native vs foreign) as the between-subjects factor, and the average μ-ERD of each participant as the dependent variable. If the interaction between the two factors was significant ($p < 0.05$) or marginally significant ($p < 0.1$), follow-up t tests were performed to compare groups in each condition separately. Additionally, multiple one-sample t tests were used to investigate if the power of μ waves changed significantly (relative to baseline) during the WoA. All tests were two sided.

Experiment 2

Two types of statistical analysis were performed with eyetracker data. A first analysis tested whether infants, in general, attended differently at the videos of native versus foreign language speakers. Eyetracker data were averaged across time to generate a single score per participant, trial type (action vs no action), and AOI (all video, face, box, and ball). The scores of each AOI were included as dependent variables in a two-factor mixed ANOVA with trial type as the within-subjects factor and language group as the between-subjects factor. A second analysis investigated changes in attention over time. A cluster mass test was performed for each linguistic group to explore periods in which the percentage of looking time (LT) between the action and no-action trials would differ significantly. First, for each group and AOI, we computed a difference vector across time and participants defined as the percentage of LT at the AOI in action trials minus the percentage of LT in no-action trials. We then computed the t values of the difference vector across time bins considering only the participants that provided data in each time bin. Due to the abrupt changes in the t value vector originating from the missing data of some participants in some time bins, the vector was smoothed using a moving average filter with the function smooth from Matlab. The moving average filter was computed every 100 ms. Next, we identified clusters as the sum of consecutive t values above a threshold corresponding to $p = 0.05$ (two tailed). We then tested the significance of the larger cluster by computing the same analysis on 2,000 permutations. The permutations were computed by randomly assigning the participants to one of the two trial types. In each permutation, the t value vector was smoothed with the same filter as the one used in the original test. Finally, we checked if the main cluster of each group of participants was >95% of what we would find by chance. If the main cluster was >95% of the permutation clusters, we reported the cluster as being significant. Thus, a significant cluster refers to a cluster that is unlikely to be found by chance.

Results

Experiment 1: μ event-related desynchronization (μ-ERD)

A marginally significant interaction between trial type (action vs no action) and linguistic group (native vs foreign) was obtained ($F_{(1,40)} = 3.54$, $p = 0.067$, $\eta_p^2 = 0.081$, ANOVA). Post hoc t tests compared μ-ERD as a function of the participants' group in each trial type. As hypothesized, in the action trials, infants in the native group ($M = -1.15$, SD = 0.48) showed greater μ-ERD than infants in the foreign group ($M = 0.41$, SD = 0.5; $t_{(1,40)} = -2.24$, $p = 0.031$; unpaired t test; Fig. 2). Significant μ-ERD was found for infants in the native group in the action trials ($t_{(1,20)} = -2.38$, $p = 0.027$; one-sample t test). However, for infants in the foreign group, μ

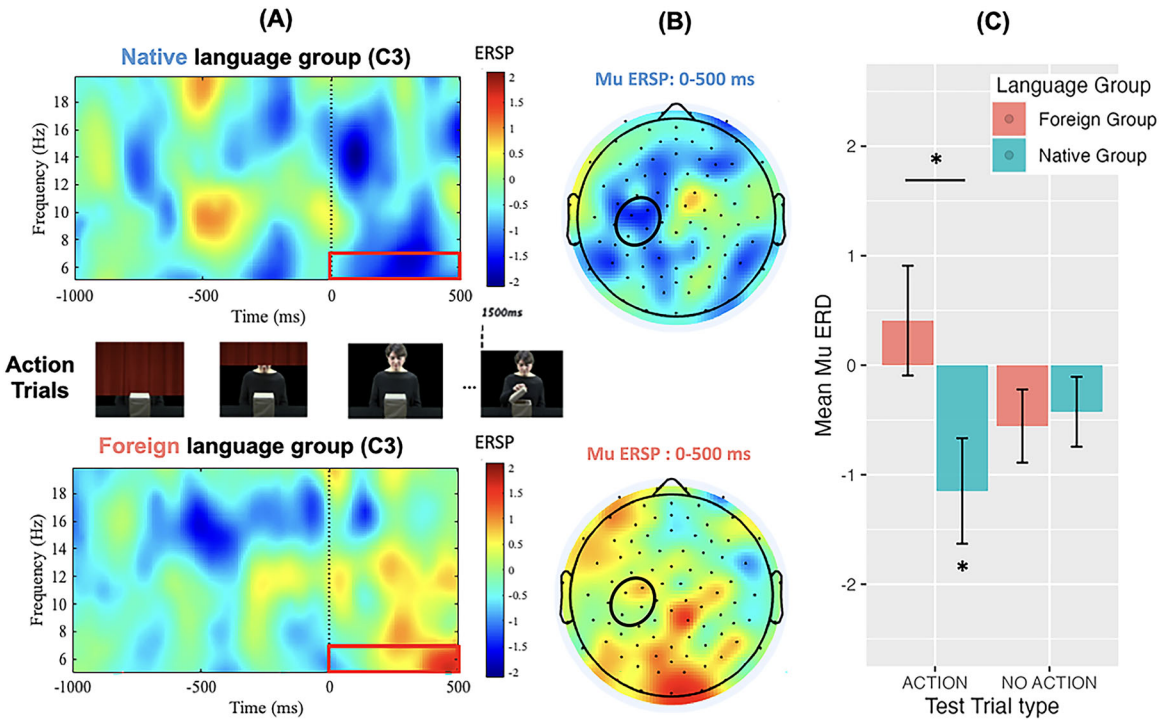

Figure 2. ERSP in decibels relative to baseline. ***A***, Spectrogram of the action trial separated by group (top, native group; bottom, foreign group), averaged across trials and channels (E30, E36, E37, and E42), and grand-averaged across participants in each group. The vertical line at Time 0 indicates the start of the WoA. The red rectangle defines the frequency (5–7 Hz) and time (0–500 ms) limits over which the analyses were computed (WoA). ***B***, Topographic maps showing the average amplitude for all channels within the μ frequency range and 0–500 ms time window (WoA). The black circle delimits the four sensorimotor channels selected for the statistical analysis. ***C***, Bars represent the averaged ERSP over the WoA with ±1 standard error, separated by trial type (*x*-axis) and group (color). A paired-sample t test found a significant difference between groups in the action trials; a one-sample t test found a significant decrease of μ only in the native group, but not in the other conditions (the symbol "*" indicates a $p < 0.05$). See Extended Data Figure 2-1 for details about the no-action trial.

activity did not significantly change relative to baseline ($t_{(1,20)} = 0.81$, $p = 0.43$; one-sample t test). In the no-action trials, there were no differences between groups of participants ($t_{(1,40)} = 0.28$, $p = 0.78$; unpaired t test; $M_{native} = -0.43$, $SD_{native} = 0.32$; $M_{foreign} = -0.56$, $SD_{foreign} = 0.33$). In addition, as expected, a one-sample t test found no significant μ-ERD for any of the two groups in the no-action trials (all $p > 0.1$; see Extended Data Fig. 2-1 for spectrograms and topoplots of the no-action trials).

In conclusion, infants in the native group showed a neural correlate of action anticipation in the action trials that suggests they predicted the agent's forthcoming behavior. However, infants in the foreign group showed no evidence that they generated action predictions. Critically, differences between groups were not found if the agent was not about to perform an action, suggesting that the effects are specific to the action anticipation. However, the interaction between group and condition was only marginally significant. Exploratory analysis further investigated the differences in infants' neural activation depending on the presented condition and their language group.

Experiment 1: exploratory analysis using permutation tests

To further address differences in neural activity depending on group and condition, we applied a nonparametric test with 2,000 permutations with false discovery rate correction for multiple comparisons using the function std_stat from the EEGLAB package (Widmann et al., 2015) in Matlab. The model included the main effects of condition and language groups, as well as their interaction.

The results showed significant interactions between condition and language within the window of interest (0–500 ms; Fig. 3*A*). Specifically, two main clusters were identified. One cluster appeared approximately from 0 to 200 ms and 9 to 14 Hz, which in infants has been defined as the β band (Xie et al., 2018). The other cluster appeared approximately from 250 to 500 ms and 5 to 6 Hz, which is part of the μ band (5–7 Hz) we previously defined. The follow-up analysis focused on the differences between groups in each condition based on the results of the nonparametric permutation test. Consistent with the main analysis, infants showed a greater event-related μ desynchronization in the native language group than that in infants in the foreign language group (200–500 ms), and this significant effect was only found in the action condition (Fig. 3*B*). The results also revealed that the difference between groups was not limited to the 5–7 Hz frequency band but extended up to 9–10 Hz. Several studies in infancy have defined the μ band up to 9 Hz (Debnath et al., 2019; Colomer et al., 2023a), and thus it is possible that the μ band recruited in the current study was wider than the 5–7 Hz range defined a priori. In addition, as suggested by the interaction between condition and group, the analysis found a cluster that appeared mainly within the β band (9–14 Hz). Infants in the native language group showed a greater β desynchronization than that in infants in the foreign language group in the action condition, but the opposite pattern was found in the no-action condition.

The results of the exploratory analysis supported the findings of the main analysis: infants showed greater μ desynchronization—and β desynchronization—in anticipation of the action of the

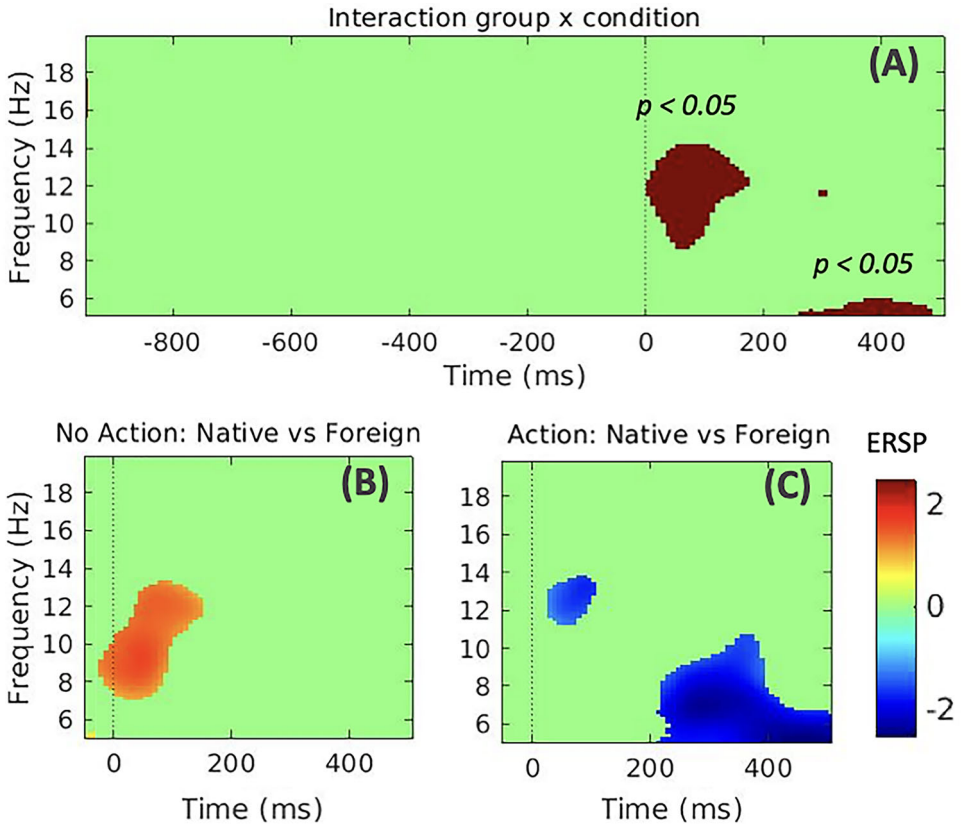

Figure 3. Results from a nonparametric test with 2,000 permutations (std_stat from EEGLAB) comparing ERSP depending on group and condition. ***A***, The colored clusters indicate a significant ($p < 0.05$) interaction between the two factors. ***B***, ***C***, ERSP difference (in decibels) in the window of interest (0–500 ms) between the native group and the foreign group (native–foreign) in the no-action condition (***B***) and in the action condition (***C***). The green images depict the nonsignificance ($p > 0.05$) between the two groups ($p < 0.05$), and the colored clusters indicate a significant difference ($p < 0.05$). The color indicates the level of ERSP difference in decibels. As indicated in the color bar (bottom-right), the colder colors indicate a greater desynchronization in the native than the foreign group, and the warmer colors indicate the opposite.

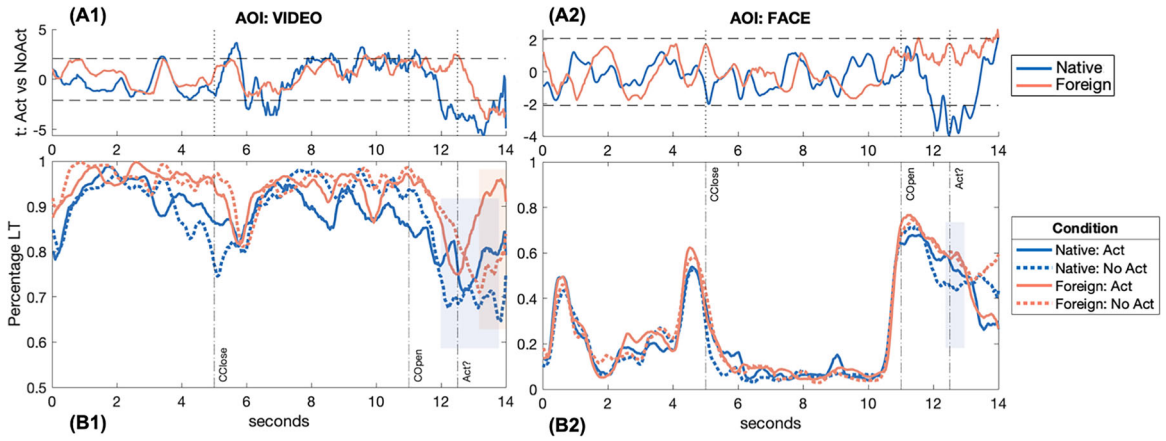

Figure 4. ***A***, Time course (in seconds) of the *t* value describing the statistical difference in percentage of LT at the video (A1) and face (A2) between action relative to no-action trials for both language groups (blue, native group; salmon, foreign group). The horizontal dashed lines indicate the threshold from which $p < 0.05$ for each group. ***B***, Time course of the mean percentage of LT at the video (B1) and face (B2) separated by group (color) and trial type (line type). The transparent rectangles indicate the clusters that were significant in a cluster mass test analysis with 2,000 permutations (blue, native group; orange, foreign group). The vertical dashed lines indicate the periods during the video when the curtain closed completely (Line 1, CClose), the curtain opened completely (beginning of the anticipation period; Line 2, Copen), and the agent acted in the action trials—but remained still in the no-action trials (end of the anticipation period; Line 3, Act?). See Extended Data Figures 4-1, 4-2, and 4-3 for more details about the AOIs, the average LT at each area, and the time course of infants' LT at the box and the ball.

native language speaker relative to the action of the foreign language speaker, a pattern that was specific to the action condition. These findings suggest that infants are less likely to anticipate the actions of people who previously spoke a foreign language. Experiment 2 tested if participants in the foreign group attended less at the stimuli as compared with participants in the native group. This would increase the likelihood of missing out on information about the location of the ball, thus making action predictions more challenging. In addition, Experiment 2 tested if infants' patterns of attention during the anticipation period differed as a function of the speaker's linguistic group.

Experiment 2: averaged attention at each AOI across all the trial

Mixed ANOVAs for each AOI, with group and trial type as factors and percentage of LT at the AOI as dependent variable, found no significant main effects or interactions of any of the factors (all $p > 0.1$; see Extended Data Fig. 4-2). In general, infants in both groups looked similarly at the videos of each trial type, suggesting that the differences between the two groups in Experiment 1 were not due to differences in overt attention to the stimuli.

Experiment 2: attention at each AOI over time

Infants in both groups showed very similar patterns of differences between the action and no-action trials across time in all the AOIs (see Extended Data Fig. 4-3 for detailed plots about the box and ball AOIs). However, important differences between the two groups were found in the period when infants could anticipate the agent's behavior (See Fig. 4). In the native group, infants looked longer at the action than no-action trials ($p < 0.001$) starting from 500 ms before the agent's behavior was revealed (anticipation period). In contrast, infants in the foreign group started attending differently to the two trials ($p < 0.001$) only 700 ms after the agent's behavior had been revealed, suggesting no prediction of her actions. Similarly, infants in the native group, but not in the foreign group, showed a significant cluster ($p = 0.015$) starting from 150 ms before the agent's behavior that indicated greater attention to her face in the action than the no-action condition. These findings suggest that infants in both groups attended similarly at the screen in general, except for the critical period of the video (anticipation of the agent's behavior) in which only infants in the native group distinguished between the action and no-action trials. The results provide further support for the possibility that infants predicted the agent's action only if she previously spoke their native language, but not if she spoke a foreign language.

Discussion

The current study investigated infants' ability to generate belief-based action predictions depending on others' linguistic group. The results on μ-ERD (Experiment 1) and eyetracker (Experiment 2) provided converging evidence that infants in the native language group, but not infants in the foreign language group, predicted the action of the agent as a function of her epistemic state. In the EEG analyses, infants showed significant μ-ERD in anticipation of the agent's action if the agent previously spoke their native language, but not if the agent spoke an unfamiliar language. In the eyetracker analyses, infants attended differently to the videos in the action relative to no-action trials already in the anticipatory period only when presented with a native speaker. When presented with a foreign speaker, infants differentiated between the two trial types after the agent's action had been revealed, but not earlier, suggesting that they did not generate predictions during the anticipation window. Together, these results suggest that the mechanisms supporting action understanding in infancy can be modulated by top–down processes, such as the recognition of relatively more familiar versus unfamiliar agents as a function of their linguistic group.

There is substantial evidence that infants are more likely to imitate and engage with the information provided by native rather than foreign speakers (Kinzler et al., 2007; Buttelmann et al., 2013). Recent evidence also indicates that infants are more likely to visually track the goal-directed actions of the native than those of the foreign language speakers (Colomer et al., 2023b). However, from these studies, it is unclear what neural mechanisms respond selectively to others' actions depending on the person's linguistic group. The current results suggest that this selectivity is supported by a modulation of the mirror system, as indicated by μ desynchronization over sensorimotor areas. Additionally, the exploratory analysis also suggested a potential role of β oscillations in selective action anticipation. Previous studies in adults have also linked β

oscillations to action processing and particularly to the mirror system (Milston et al., 2013). However, the functional significance of the β band in infancy is not clear (Cuevas et al., 2014), and thus future work is needed to elucidate how β oscillations relate to infants' action predictions.

The results indicate that even before others' actions are presented, infants show a different mental representation of these actions depending on the person's linguistic group. This selective response during action anticipation may influence infants' ability to interpret, evaluate, and learn from others' actions and may, consequently, bias infants toward the actions of native speakers. In fact, previous research has found a link between infants' ability to generate action predictions and their capacity to imitate others' actions (Gampe et al., 2016) and identify when actions are unexpected (Gredebäck et al., 2018).

A further question is why infants did not predict the action of the foreign speaker. One possibility is that infants showed less interest in the agent and her actions if she previously spoke a foreign language. In fact, infants in the native group attended longer at the presentation than infants in the foreign group. If infants also showed a different level of interest in the agent's actions, this could result in a reduced predisposition to anticipate and mirror what she would do next. In line with this, previous research has typically interpreted infants' biases toward native speakers in terms of a preference for native speakers, either because they are identified as in-group members (Kinzler et al., 2007; Buttelmann et al., 2013) or because they are evaluated as better sources of information (Begus et al., 2016). Research in adults shows similar patterns as people are less likely to represent the actions of out-group relative to in-group members during a joint action (McClung et al., 2013; Sacheli et al., 2015), and they are less likely to mirror the actions of other-race individuals (Gutsell and Inzlicht, 2010), a bias that correlates with adults' racial prejudice. However, the findings from the eyetracker in our study challenge the less interesting interpretation. First, infants attended similarly to the test videos of the native and foreign speakers. Second, infants in the foreign language group showed increased attention to the video in the action as compared with the no-action condition after the agent began to produce her action. Particularly, the descriptive data shows a peak in attention after infants saw the action of the foreign speaker at the end of the video, suggesting an interest in seeing and encoding her action.

Alternatively, infants' social selectivity may not be based on a preference toward native speakers but on a facilitation to predict and interpret the actions of native speakers as compared with foreign speakers. At 6 months of age, infants begin to show an ability to generate basic expectations about the agents and their actions, and they are surprised if the agents violate their expectations (Woodward, 1998; Csibra, 2008; Vouloumanos et al., 2014). We propose that infants may be less likely to generate basic expectations about the agents if presented with a stranger who demonstrated to behave in ways that are inconsistent with infants' knowledge about the social world (for a similar argument, see Colomer and Woodward, 2023). Infants may associate foreign speakers with behaviors and properties that are unfamiliar (Uttley et al., 2013; Oláh et al., 2014), and thus they may be less likely to use their basic knowledge about the agents to predict the actions of foreign speakers. For example, it has been proposed that infants solve false belief tasks by expecting agents to act toward the last location in which they saw an object (Tomasello, 2018). The current results, together with the results of Southgate and Vernetti (2014), suggest that 6-month-old infants expect agents to act consistently with what they see or have seen in the past, but such an assumption may not apply to agents who previously spoke in a foreign language.

Our findings have implications for theories about the neural underpinnings of action understanding and encourage further investigation into the role of the agent's identity in action prediction. In one account, infants' action understanding is grounded in their ability to produce actions and represent the observed actions in their mirror system (Sommerville et al., 2005; Woodward and Gerson, 2014; Gredebäck and Falck-Ytter, 2015). For example, infants predict at an earlier age the goals of actions of human rather than nonhuman agents (Cannon and Woodward, 2012; Adam et al., 2016) and show a greater ability to anticipate others' actions the more competent they are producing these actions (Kanakogi and Itakura, 2011; Colomer et al., 2023a). However, bottom–up sensorimotor information cannot explain our findings, since infants in both groups saw the same set of actions. One possibility is that a larger hierarchical system is involved in action prediction, which engages low-level (e.g., sensory information) and high-level (e.g., context information) computations that, together, allow infants to infer others' intentions (Kilner et al., 2007; Southgate et al., 2009; Urgen and Miller, 2015). Linguistic group cues may change how infants interpret the agent's behavior at the context level, depending on whether the agent's past behavior was consistent or inconsistent with the infants' prior model of the social world.

In conclusion, the present study provides important insights into our understanding of two fundamental processes of social cognition that had been considered separately: the precursors of social categorization and the mechanisms supporting action prediction. First, the results show that the mechanisms supporting action prediction, in this case the neural correlates of the mirror system, are flexible from early infancy and may be modulated by social cues such as the language people speak. Second, the results suggest that infants' biases to attend to, imitate, and learn from the actions of native speakers may be due, in part, to a facilitation to predict their actions relative to the actions of foreign speakers.

References

Adam M, Reitenbach I, Papenmeier F, Gredebäck G, Elsner C, Elsner B (2016) Goal saliency boosts infants' action prediction for human manual actions, but not for mechanical claws. Infant Behav Dev 44:29–37.

Altınok N, Király I, Gergely G (2021) The propensity to learn shared cultural knowledge from social group members: selective imitation in 18-month-olds. J Cogn Dev 23:273–288.

Begus K, Gliga T, Southgate V (2016) Infants' preferences for native speakers are associated with an expectation of information. Proc Natl Acad Sci U S A 113:12397–12402.

Buttelmann D, Zmyj N, Daum M, Carpenter M (2013) Selective imitation of in-group over out-group members in 14-month-old infants. Child Dev 84:422–428.

Cannon EN, Woodward AL (2012) Infants generate goal-based action predictions. Dev Sci 15:292–298.

Chung H, Meyer M, Debnath R, Fox NA, Woodward A (2022) Neural correlates of familiar and unfamiliar action in infancy. J Exp Child Psychol 220: 105415.

Colomer M, Woodward A (2023) Should I learn from you? Seeing expectancy violations about action efficiency hinders social learning in infancy. Cognition 230:105293.

Colomer M, Chung H, Meyer M, Debnath R, Morales S, Fox NA, Woodward A (2023a) Action experience in infancy predicts visual-motor functional connectivity during action anticipation. Dev Sci 26:e13339.

Colomer M, Hwang HG, Burke N, Woodward A (2023b) Development of infants' preferential looking toward native language speakers across distinct social contexts. Dev Psychol 60:124–134.

Csibra G (2008) Goal attribution to inanimate agents by 6.5-month-old infants. Cognition 107:705–717.

Cuevas K, Cannon EN, Yoo K, Fox NA (2014) The infant EEG mu rhythm: methodological considerations and best practices. Dev Rev 34:26–43.
Datavyu Team (2014). *Datavyu: video coding and data visualization tool*. Databrary Project, New York University. Available at:
Debnath R, Salo VC, Buzzell GA, Yoo KH, Fox NA (2019) Mu rhythm desynchronization is specific to action execution and observation: evidence from time-frequency and connectivity analysis. NeuroImage 184: 496–507.
Debnath R, Buzzell GA, Morales S, Bowers ME, Leach SC, Fox NA (2020) The Maryland analysis of developmental EEG (MADE) pipeline. Psychophysiology 57:e13580.
Gampe A, Prinz W, Daum MM (2016) Measuring action understanding: relations between goal prediction and imitation. Br J Dev Psychol 34: 53–65.
Gredebäck G, Falck-Ytter T (2015) Eye movements during action observation. Perspect Psychol Sci 10:591–598.
Gredebäck G, Lindskog M, Juvrud JC, Green D, Marciszko C (2018) Action prediction allows hypothesis testing via internal forward models at 6 months of age. Front Psychol 9:290.
Gutsell JN, Inzlicht M (2010) Empathy constrained: prejudice predicts reduced mental simulation of actions during observation of outgroups. J Exp Soc Psychol 46:841–845.
Howard LH, Henderson AME, Carrazza C, Woodward AL (2015) Infants' and young children's imitation of linguistic in-group and out-group informants. Child Dev 86:259–275.
Kanakogi Y, Itakura S (2011) Developmental correspondence between action prediction and motor ability in early infancy. Nat Commun 2:341.
Kayser J, Tenke CE (2006) Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: II. Adequacy of low-density estimates. Clin Neurophysiol 117:369–380.
Kilner JM, Friston KJ, Frith CD (2007) Predictive coding: an account of the mirror neuron system. Cogn Process 8:159–166.
Kinzler KD, Dupoux E, Spelke ES (2007) The native language of social cognition. Proc Natl Acad Sci U S A 104:12577–12580.
Liberman Z, Woodward AL, Kinzler KD (2017) The origins of social categorization. Trends Cogn Sci 21:556–568.
Marno H, Guellai B, Vidal Y, Franzoi J, Nespor M, Mehler J (2016) Infants' selectively pay attention to the information they receive from a native speaker of their language. Front Psychol 7:1150.
McClung JS, et al. (2013) Group membership affects spontaneous mental representation: failure to represent the out-group in a joint action task. PLoS One 8:e79178.
Milston SI, Vanman EJ, Cunnington R (2013) Cognitive empathy and motor activity during observed actions. Neuropsychologia 51:1103–1108.
Mognon A, Jovicich J, Bruzzone L, Buiatti M (2011) ADJUST: an automatic EEG artifact detector based on the joint use of spatial and temporal features. Psychophysiology 48:229–240.
Nazzi T, Bertoncini J, Mehler J (1998) Language discrimination by newborns: toward an understanding of the role of rhythm. J Exp Psychol Hum Percept Perform 24:756–766.
Nolan H, Whelan R, Reilly RB (2010) FASTER: fully automated statistical thresholding for EEG artifact rejection. J Neurosci Methods 192:152–162.
Oláh K, Elekes F, Bródy G, Király I, Quinn P (2014) Social category formation Is induced by cues of sharing knowledge in young children. PLoS One 9: e101680.
Pun A, Ferera M, Diesendruck G, Kiley Hamlin J, Baron AS (2018) Foundations of infants' social group evaluations. Dev Sci 21:e12586.
Rhodes M, Baron A (2019) The development of social categorization. Annu Rev Dev Psychol 1:359–386.
Rizzolatti G, Craighero L (2004) The mirror-neuron system. Annu Rev Neurosci 27:169–192.
Sacheli LM, Christensen A, Giese MA, Taubert N, Pavone EF, Aglioti SM, Candidi M (2015) Prejudiced interactions: implicit racial bias reduces predictive simulation during joint action with an out-group avatar. Sci Rep 5:8507.
Shutts K, Kinzler KD, McKee CB, Spelke ES (2009) Social information guides infants' selection of foods. J Cogn Dev 10:1–17.
Soley G, Sebastián-Gallés N (2015) Infants prefer tunes previously introduced by speakers of their native language. Child Dev 86:1685–1692.
Soley G, Spelke ES (2016) Shared cultural knowledge: effects of music on young children's social preferences. Cognition 148:106–116.
Sommerville JA, Woodward AL, Needham A (2005) Action experience alters 3-month-old infants' perception of others' actions. Cognition 96:B1–B11.
Southgate V, Begus K (2013) Motor activation during the prediction of nonexecutable actions in infants. Psychol Sci 24:828–835.
Southgate V, Vernetti A (2014) Belief-based action prediction in preverbal infants. Cognition 130:1–10.
Southgate V, Johnson MH, Osborne T, Csibra G (2009) Predictive motor activation during action observation in human infants. Biol Lett 5:769–772.
Southgate V, Johnson MH, El Karoui I, Csibra G (2010) Motor system activation reveals infants' on-line prediction of others' goals. Psychol Sci 21:355–359.
Tomasello M (2018) How children come to understand false beliefs: a shared intentionality account. Proc Natl Acad Sci U S A 115:8491–8498.
Urgen BA, Miller LE (2015) Towards an empirically grounded predictive coding account of action understanding. J Neurosci 35:4789–4791.
Uttley L, de Boisferon AH, Dupierrix E, Lee K, Quinn PC, Slater AM, Pascalis O (2013) Six-month-old infants match other-race faces with a non-native language. Int J Behav Dev 37:84–89.
Vouloumanos A, Martin A, Onishi KH (2014) Do 6-month-olds understand that speech can communicate? Dev Sci 17:872–879.
Widmann A, Schröger E, Maess B (2015) Digital filter design for electrophysiological data – a practical approach. J Neurosci Methods 250:34–46.
Woodward A (1998) Infants selectively encode the goal object of an actor's reach. Cognition 69:1–34.
Woodward AL, Gerson SA (2014) Mirroring and the development of action understanding. Philos Trans R Soc Lond B Biol Sci 369:20130181.
Xie W, Mallin BM, Richards JE (2018) Development of infant sustained attention and its relation to EEG oscillations: an EEG and cortical source analysis study. Dev Sci 21:e12562.