Running head: SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

The role of social signals in segmenting observed actions in eighteen-month-old children

Christian Kliesch[1,2,3], Eugenio Parise[3,4], Vincent Reid[3,5], & Stefanie Hoehl[2,6]

[1] Department of Psychology, Potsdam University, Potsdam, Germany

[2] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[3] Department of Psychology, Lancaster University, United Kingdom

[4] CIMeC – Center for Mind/Brain Sciences, University of Trento, Italy

[5] School of Psychology, University of Waikato, New Zealand

[6] Department of Developmental and Educational Psychology, University of Vienna, Austria

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

Correspondence concerning this article should be addressed to Christian Kliesch, Department of Psychology, Potsdam University, Campus Golm, Karl-Liebknecht-Str. 24-25, 14476 Potsdam. E-mail: christian.kliesch@uni-potsdam.de

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

Abstract

Learning about actions requires children to identify the boundaries of an action and its units. Whereas some action units are easily identified, parents can support children's action learning by adjusting the presentation and using social signals. However, currently little is understood regarding *how* children use these signals to learn actions.

In the current study we investigate the possibility that communicative signals are a particularly suitable cue for segmenting events. We investigated this hypothesis by presenting 18-month-old children (N=60) with short action sequences consisting of toy animals either hopping or sliding across a board into a house, but interrupting this two-step sequence either (a) using an ostensive signal as a segmentation cue, (b) using a non-ostensive segmentation cue, and (c) without additional segmentation information between the actions.

Marking the boundary using communicative signals increased children's imitation of the less salient sliding action. Imitation of the hopping action remained unaffected. Crucially, marking the boundary of both actions using a non-communicative control condition did not increase imitation of either action. Communicative signals might be particularly suitable in segmenting non-salient actions that would otherwise be perceived as part of another action or as non-intentional. These results provide evidence of the importance of ostensive signals at event boundaries in scaffolding children's learning.

Keywords: Action Segmentation, Communicative Signals, Ostensive Communication, Action Imitation

Word count: 6948

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

The role of social signals in segmenting observed actions in eighteen-month-old children

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

Research highlights:

- We investigate the role of direct gaze and infant directed speech in segmenting action sequences in 18-month-old children.

- Children are more likely to imitate parts of an action sequence if the boundaries of its units are marked by a short burst of gaze and infant-directed-speech, compared to a non-ostensive signal of the agent pausing and saying 'hmm,' or no boundary marking.

- The increase in imitation is only seen for the non-salient action type; the salient action type shows high levels of imitation irrespective of the boundary marking condition.

- Ostensive signals act as communicative signals of temporal reference, delineating the different units of an action sequence for toddlers.

Introduction

Children are avid social learners. From the age of 9 months they anticipate the goals of other social agents (Biro, 2013; Gredebäck et al., 2009; Reid et al., 2009) and start intentionally imitating other people's actions during their second postnatal year (Jones, 2007, 2009). Caregivers support children's learning of novel information using a wide range of behavioural adaptations and social signals. So far, parents' modifications of child directed actions have been studied without direct reference to how these adaptations may contribute to the segmentation of action sequences. There are a number of recent findings, however, that suggest that communicative signals may also contribute to the segmentation of event sequences, and may be particularly important for children's action learning. In the present study, we tested whether 18-month-old children can use a brief exposure to a communicative signal to segment an action sequence and subsequently adjust which parts of the action sequence they imitate.

Children's action imitation is selective

Actions are intentional, goal-directed movements. These features distinguish them from incidental movements and events. Already from an early age, infants show an understanding of actions as goal-directed (Biro et al., 2014; Csibra, 2003; Verschoor et al., 2013) and distinguish between actions that are carried out in an efficient and non-efficient manner in relation to the action goal (Gergely & Csibra, 2003; Gredebäck & Melinder, 2011). Children do not blindly imitate all actions exactly the way they are shown (cf. Csibra, 2008), but selectively imitate the parts of the action (Carpenter et al., 1998; Király et al., 2013). For example, they will imitate a model turning on a lamp with their head, but only if there is no other rational explanation for this inefficient action manner (e.g. the model's hands were occupied Gergely et al., 2002; Király et al., 2013).

Children's selective imitation is also evident in studies by Carpenter et al. (2005) and Southgate et al. (2009): Here, children were presented with a toy animal hopping or sliding into a toy house. When the hopping and sliding motions were presented on their own, children chose to imitate these actions to a high degree. However, children were less likely to imitate the manner of the action when it was presented together with a clear goal, e.g. putting the animal in the house, even though children were still more likely to imitate the hopping than sliding (Carpenter et al., 2005). Furthermore, children who were told about the goal of the action, were more likely to imitate the action's manner, compared to children who observed manner and outcome of the action, or discovered the outcome on their own (Southgate et al., 2009). Data from younger children indicates that 6-month-old infants are less likely to notice changes in the transition between different outcomes, compared to a change of the target locations. For example, they are

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

less likely to notice when a self-propelled ball bounces down before reaching its goal location after seeing it go on a straight line during habituation, compared to a change from one target goal location to another (Hespos et al., 2009). Therefore, children's focus on the outcome of the action demonstrated in Carpenter et al. (2005) and Southgate et al. (2009) could be explained by perceptual failure of segmenting both actions and/or assigning intentionality to both parts of the action sequence.

The structure of actions determines how actions are understood

Like events, most actions can be described on different levels. They are hierarchically organised, with simpler action units nested in higher-order action plans (Elman, 1990; Zacks & Tversky, 2001). The action of making a cup of coffee, for example, may be described by said overarching goal ('Making coffee') or by the units that comprise the action, e.g. grinding the coffee, boiling the water, filling the coffee press with coffee, and subsequently water and pressing the coffee press down. These actions comprise the *action sequence* of the action *'making coffee'*, which in turn may be part of the overarching action sequence of *'preparing breakfast'*. As we can see, most actions consist of different action units with their own goals and sub goals. At the same time, they do not exist as isolated events, but are part of a wider action sequence (Zacks et al., 2009). In the current paper, we will therefore refer to the highest relevant level of analysis of the description as the *action*, which comprises of *action units*. Therefore, observing actions (and potentially carrying out actions Hommel et al., 2001) shares many features with a general description of events, enriched by goals and intentions (Zacks et al., 2009).

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

According to Event Segmentation Theory (Zacks et al., 2009; Zacks & Swallow, 2007) segmenting a stream of events plays an important role in the comprehension, anticipation and subsequent imitation of event sequences (Baldwin et al., 2001; Zacks et al., 2001, 2007, 2009). The identification and segmentation of event boundaries is just as important as the identification of objects in space. Events can be segmented based on low-level features, such as motion cues, or prior, higher-order knowledge of the event (Zacks & Swallow, 2007). Adults are able to segment actions into finer or coarser units when requested (Zacks, 2004; Zacks et al., 2009). Already 10- to 11-month-old infants demonstrate sensitivity to event structures, looking longer at actions paused mid-stream, compared to those where the pause coincided with an event boundary (Baldwin et al., 2001). It is likely that visual features of the action stream are important contributors in identifying potential goals.

Understanding actions as hierarchical structures has important implications for children's imitation of actions as well: As we have seen in the study by Carpenter et al. (2005), children ignore parts of an action sequence when focusing on another part, even though they are capable of imitating the sequence on its own when no overarching goal is present. From an event segmentation perspective, the children who ignored the action manner potentially did not segment both actions, but saw the manner as only instrumental in getting the animal into the house. Interestingly, although imitation of both manner actions was reduced when the goal of putting the animal into the house was present, overall fewer children imitated the sliding compared to the hopping action. This might be because the hopping action can be more readily identified as an action on its own, as it is more repetitive, has a larger movement range and the additional effort and energy expenditure suggests that it marks it as an intentional action. The

children in the study by Southgate et al. (2009) on the other hand did not ignore the manner of the action if they were made aware of the outcome of the action ostensively prior to the action demonstration. Knowing the outcome of the action before might have helped children to perceive the action as its own event. Using this information, children are then able to recognise the manner as a separate event, in the same way that recognising familiar words within a stream of syllables allows for the recognition of new candidate words.

Caregivers modify and adapt actions in teaching contexts and play

Caregivers adapt and adjust how they present actions to children to make them more accessible for learning. Infant-directed actions are often presented in an exaggerated manner (Brand et al., 2002; Koterba & Iverson, 2009; Rutherford & Przednowek, 2012; Schaik et al., 2020; Williamson & Brand, 2014), are highly repetitive (Brand et al., 2009) and caregivers interrupt and emphasise the boundaries of action units (Williamson & Brand, 2014). This balance between variation and repetition may be particularly suitable for learning and retaining attention (Brand et al., 2009; Goldstein et al., 2010; Twomey et al., 2017; Twomey & Westermann, 2017) and increases the capacity to learn from socially presented actions, compared to actions that children only observe incidentally.

In addition, caregivers also use other social signals to ease and support infant learning. According to Natural Pedagogy Theory (Csibra & Gergely, 2009, 2011) children have an innate sensitivity to some social signals that inform them of the presence of a communicative interaction. Natural Pedagogy builds upon dual intentions of communication — the intention to communicate, and the intention transmitting the content of the message (c.f. Sperber & Wilson,

1995; Sperber & Wilson, 2002). Although young infants are unlikely to already possess the meta-cognitive skills to interpret these informative intentions, Csibra (2010) suggests that infants have an early sensitivity towards direct gaze, infant-directed speech and contingent interactions that signal caregivers' communicative intention. When addressed with these *ostensive* signals, children expect that caregivers will provide them with generalisable information (Csibra & Gergely, 2009; Csibra & Shamsudheen, 2015). This very simple code-based system allows children to rapidly learn and acquire culturally-relevant knowledge (Csibra & Gergely, 2009, 2011). Their use is not just restricted to explicit teaching contexts, but also free play (Sage & Baldwin, 2012).

Caregivers also use a second type of communicative signals, *referential signals*, to link the content of an interaction to the world around them. For example, through the use of pointing and gaze following, caregivers can restrict the number of possible referents (Senju et al., 2008; Woodward & Guajardo, 2002). Children start to follow pointing and gaze during their first year of life and (Morissette et al., 1995; Senju & Csibra, 2008) and identify the targets of pointing between 15-18 months (Morissette et al., 1995). These signals fulfill an important role of providing spatial reference and thereby restrict the number of possible referents in space (cf. Clark, 2003).

Learning about actions also requires temporal information to identify potential candidate units for predicting and imitating an action sequence. Currently, there is no research into whether a similar referential signal exists in the temporal domain. Such a marker may be particularly important in segmenting events and action units that cannot reliably be identified through observation and therefore may be particularly important for event learning in general, and action

learning in particular. A temporal marker to denote the beginnings and ends of actions and their units may be just as important for learning about actions in the same way that protodeclarative pointing enables the disambiguation of different objects.

Especially when the boundaries of events and actions cannot be reliably identified through observation, caregivers' use of signals or cues that help to segment individual action units could potentially support children's learning of these actions. Indeed, previous research has found that parents are more likely to look toward and address children at action boundaries (Brand et al., 2007, 2013; Williamson & Brand, 2014). Children are also more likely to imitate an action if a model looks at them during event boundaries (Williamson & Brand, 2014) and parents reduce the amount of direct gaze at action boundaries with increasing age (and presumably prior knowledge) of their child (Brand et al., 2007).

Many of these signals used by parents at event boundaries also signal the presence of communication, such as direct gaze and infant-directed speech. Direct gaze may be a particularly suitable signal to segment events and actions. Newborn infants already prefer looking at eyes with the contrast polarity of the human pupil/sclera, but not the inverse contrast polarity (Farroni et al., 2005). Faces and gaze also play an important role during early development, as during early infancy, faces are dominant stimuli (Fausey et al., 2016). This sensitivity and high frequency of exposure might help children to pay attention to faces as signifiers of interactions that are directed towards them. In adults, brief periods of direct gaze interrupts working memory and delays response times in a visual search task. Direct gaze also provides the listener with feedback during dialogue (e.g. Bavelas et al., 2002; Hömke et al., 2017). Furthermore, direct gaze is common even in cultures considered to be gaze-avoidant (Haensel et al., 2021). Because of its

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

privileged role and potentially disruptive properties, adults' use of direct gaze and infant-directed speech at action boundaries may help children to break up and identify individual parts of an incoming stream of actions. Therefore, children might not only use ostensive signals to infer the presence of communication, but also segment an incoming event or action sequence. In the following study we investigate whether children use a social signal that is traditionally seen as an ostensive signal to segment an action stream by directly comparing the effect of a social marker, a non-social marker and a baseline condition without a boundary marker in segmenting a short action sequence.

We investigated this question by adapting the paradigm used by Carpenter et al. (2005) and Southgate et al. (2009). Their finding, that children frequently imitate the outcome of the action but are much less likely to imitate its manner, could be interpreted as a failure to segment both action units. Instead, children perceive the action sequence of hopping or sliding into the house as part of a single action unit and consequently focus only on the outcome. If that was indeed the case, children's imitation of the action's manner might increase if the action sequence of hopping/sliding the animal into the house is interrupted between both action units, for example by addressing the children using direct gaze and infant-directed speech. This raises the interesting question of whether certain signals, such as communicative signals, are particularly suitable to segment such an event sequence, or whether any interruption of the action sequence is sufficient to mark both action units as separate units and subsequently increase the imitation of the action's manner. We investigated this question in a between-subjects experiment by presenting three groups of 18-month-old toddlers the same action sequences in three different conditions.

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

The *Baseline condition* was comparable to the mouse-hopping/sliding-into-a-house sequence used by Carpenter et al. (2005) and the no-prior-information condition reported in the study by Southgate et al. (2009). Like Southgate et al. (2009), we used only one house rather than two houses like in Carpenter et al. (2005). Additionally, we added a communicative signal (a short exclamation of "Wow" combined with direct gaze) after the entire hopping/sliding-into-the-house sequence was completed, to ensure that the same amount of ostensive signals are used compared to the communicative condition.

In the *Communicative condition*, exactly the same linguistic and communicative signals were used. However, the short communicative signal was now presented *after* the hopping/sliding action and *before* putting the animal into the house. If social signals aid children's action segmentation, establishing direct gaze and addressing the child with infant-directed speech at the boundary between the hopping/sliding event and putting the animal into the house should help them interpret both action units as separate units worthy of imitation. We predicted that, if segmented in such way, imitation of the manner should increase for both types of action, because children now perceive both actions as distinct action events.

Not all kinds of interruptions may, however, lead to the successful segmentation and subsequent imitation of the manner action unit. In the *Control condition*, we investigated whether a non-communicative signal—the actor paused, looked down and said 'Hmmm'—would also affect children's manner imitation. Therefore, the condition is identical to the Communicative condition, except for the signal used to mark the boundary between both action units. If the results mirror the baseline condition, we can infer that children do not use any interruption to segment the action sequence. If the results mirror the communicative condition, we can infer that

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

children also use non-communicative signals to segment action sequences. This would suggest that children use such interruptions to segment and learn about actions and whether communicative signals are more suitable than other, non-communicative, interruptions.

Finally, because imitation of the putting-the-animal-into-the-house action was already high in previous studies, we do not expect that children's imitation of the action outcome will be affected by additional segmentation information.

Methods

The methodology, hypotheses and analyses were preregistered on `aspredicted.org`, reference number #5771 for the baseline and marked-communicative conditions and #19880 for the marked-control condition. To account for missing data on the subject level, we deviated from the original hypothesis #5771 by using a Generalised Linear Models with item and subject random effects instead of ANOVAs.

Participants

The final sample contained 60 18-month-old toddlers (Mean: 18m, Min: 17.5m, Max: 18.5m, 28 female), with 20 children in each condition. An additional 11 toddlers were tested, but excluded due to being unwilling to engage with the game (8), parental or sibling interference during all trials (2), incorrect age at time of testing (1). Eight of these children were in the communicative condition, 2 in the baseline condition and 1 in the non-communicative control condition. Written informed consent was provided by the caregivers and procedures were in accordance with institutional protocols.

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

Materials

The actions were presented on a green cardboard mat (42 × 60 cm) with a small cardboard house (yellow, red). Four small toy animals (fox, rabbit, hedgehog, squirrel, all approximately 6–8 cm tall) were used to act out the actions. The animals were kept in a small, colourful box prior to the experiment. Additionally, we used a wooden stacking game during the warm-up phase.

Procedure

Toddlers were sitting on their caregiver's lap. After a warm-up session to familiarise the toddlers with the room and the experimenter, the experimenter presented each animal to the toddler with a short statement (e.g. "The squirrel has a bushy tail"). The toddler was allowed to play with all animals for approximately one minute. Afterwards, the animals were returned to the box and the experimenter revealed the board with the house. The modelling phase began during which the experimenter took out one animal, placed it on the board and said: "Look what the [animal] does!" (German original: "Schau mal, was das [Tier] macht!"). He then moved the animal across the table with either the sliding or the hopping action.

In the *marked-communicatively* condition, the experimenter looked up to the toddler and said "Wow" *after* the hopping/sliding movement, but *before* putting the animal into the house. In the *baseline* condition, the adult put the animal into the house before looking towards and addressing the child.

In the *marked-control* condition, the experimenter looked down, put his hand to the chin and said "Hmmmm." *after* the hopping/sliding movement, but *before* putting the animal into the

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

house.

In the *baseline* condition, the experimenter did not pause between the hopping/sliding and putting animal into the house, but said "Wow" *after putting the animal into the house* to ensure the type and amount of verbal information remained the same when contrasted with the other conditions.

After the animal was put into the house, the experimenter said "Great, the animal went into the house. Now it's your turn!" (German original: "Toll, das Tier ist ins Haus gegangen. Jetzt bist du dran!") before pushing the board to the child. Each trial demonstration lasted approximately 10 seconds, and the child had 30 seconds to respond. If the child did not engage with the animal, the experimenter encouraged the child by saying "Now you can play with it!" ("Jetzt kannst du damit spielen"), "Now it's your turn" ("Jetzt bist du dran") or similar. If the child attempted to pull the house off the board, the experimenter said: "That's attached" ("Das ist fest."). A visual illustration of the procedure is shown in Figure 1.

fig1here

Each child was presented with up to four trials of the actions. The actions were shown in a fixed order of sliding–hopping–hopping–sliding, (as recommended by Southgate et al., 2009). Boundary-marking was presented as a between-subjects factor: one third of the children saw the action in the boundary-marked communicatively condition, one third in the boundary-marked communicatively condition, and one third in the boundary unmarked control condition.

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

Coding

Infants were scored on whether they (1) imitated the action manner (2) imitated the goal/outcome of the action. In line with previous research (Carpenter et al., 2005; Southgate et al., 2009), the action manner was coded as sliding when the animal moved continuously without breaking contact with the mat. The child imitated the hopping action, when the animal broke contact and made contact at least once again with the mat. For the analysis, only the previously modelled behaviour was coded as 1, any other non-modelled behaviour (e.g. hopping during a sliding trial) was coded as 0. The goal of putting the animal into the house was achieved if the child put the animal into the house at least once, even if the child removed the animal afterwards.

Children were included in the analysis if they contributed at least any 2 out of the 4 trials. After coding, a total of 220 trials were included in the analysis (111 hopping, 109 sliding). An additional 20 trials (9 hopping, 11 sliding) were excluded from the analysis due to the child refusing to touch the animal or being fussy (11), parental interference (7) and experimenter error (2).

A second coder naïve to the hypothesis coded manner and outcome in the videos of 39 children. The interrater agreement was excellent for both manner $\kappa = 0.90, p < .0001$ and outcome $\kappa = 0.94, p < 0.001$.

Data analysis

Because some trials were missing, we decided to compute a Generalised Linear Mixed Effects model based on the binomial distribution using R [Version 4.0.3; R Core Team (2019)]. Preprocessing was conducted using the *tidyverse* (Wickham et al., 2019) packages and statistical

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

models were built in *lme4* (Bates et al., 2015). Based on our pre-registration, we compared a model with Boundary Marker (factor levels: baseline, marked-communicative, marked-control) to a null model containing the intercept only. In an exploratory analysis, we also tested the interaction between Action Type (hopping/sliding) and Boundary Marking. We attempted to use the maximally converging random effects structure in line with recommendations by Barr et al. (2013). However, all models that were more complex than a single intercept with participant number at random effect level failed to converge due to singularity or the optimizer failing to establish a reliable solution to the model.

Results

fig2here

Manner imitation: For the manner imitation in the Baseline condition, 18/36 children imitated the hopping, and 5/39 the sliding action. In the Marked-Communicative condition, 16/38 imitated the hopping, and 14/35 the sliding action. In the Marked-Control condition, 11/37 imitated the hopping, and 8/35 the sliding action.

We are primarily interested in whether children were more likely to imitate the first action step if the boundary between the two action steps was marked either communicatively or non-communicatively. Overall, the model containing only the boundary marker conditions did not perform better than the null model containing the intercept-only null model ($\chi^2(3.50) = 2.00$, $p = .174$). However, the model containing the interaction between boundary marker and action type performed significantly better than the marker-only model ($\chi^2(13.70) = 3.00$, $p = .003$)

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

and the intercept only model ($\chi^2(17.19) = 5.00$, $p = .004$). A detailed overview of the full

comparison can be seen in Table 1.

Table 1:

Model comparison between baseline, marked-communicative and marked-control condition for

the first part of the action sequence, the manner of putting the animal into the house.

	Par	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
Null Model	2	281.577	288.364	-138.788	277.577			
Marker Model	4	282.082	295.656	-137.041	274.082	3.495	2	0.174
Marker x Action Type Interaction	7	274.385	298.140	-130.192	260.385	13.697	3	0.003

To investigate the effect of the boundary marking on children's imitation of the action

manner, we first compared the manner imitation in the experimental groups to the baseline

condition. To disentangle the interaction effect, we investigated these differences on the subsets

of the hopping/sliding action factor separately. Our results show that the sliding action was

imitated by only few children at baseline (M = 12%, 95% CI[5%, 27%]). Here, the marked-

communicative condition shows a statistically significant increase in imitation (M = 39%, 95%

CI[24%, 58%], $\beta = 1.57$, p = .015), but this effect could not be found in the marked-control

condition (M = 21%, 95% CI[10%, 39%], $\beta = 0.72$, p = .271). The hopping action was imitated

by half the children at baseline level (M = 50%, 95% CI[33%, 67%]). Marking the boundary

communicatively (M = 42%, 95% CI[26%, 59%]) or with the control intervention (M = 29%,

95% CI[16%, 46%]) did not significantly affect their imitation (Baseline – Marked-

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

communicative: β = -0.38, p = .524, Baseline – Marked-control: β = -1.05, p = .103). A visual representation of these results can be found in Figure 2 A.

We investigated the effect of communication against the other two conditions by investigating the subsets of the hopping and sliding data. We found that children imitate the sliding manner significantly more (β = 1.20, p = .022) than in the other two conditions, but marking actions in the communicative condition did not have the same effect for the hopping action (β = 0.13, p = .797).

Finally, we analysed the difference between the hopping/sliding action for each of the condition levels separately. In line with previous studies using this paradigm (Carpenter et al., 2005) we found a difference in the imitation of the hopping and sliding action in the baseline condition, i.e. the condition that replicates the original study most closely (β = -2.01, p = .002). This difference was neither apparent in the marked-communicative (β = -0.10, p = .843), nor the marked-control condition (β = -0.41, p = .473).

Outcome imitation: For the outcome imitation in the Baseline condition, 33/36 children put the animal into the house after the hopping, and 34/39 after the sliding action. In the Marked-Communicative condition, 36/38 imitated the outcome action after hopping, and 32/35 after sliding action. In the Marked-Control condition, 34/37 imitated the outcome after the hopping, and 31/35 the sliding action.

The analysis using a mixed effects model with the same specifications as in the manner analysis suggested that the null model explains the data best. Imitation of the action outcome was very high across all conditions and 90.91% of the children imitated the outcome (see Figure 2 B).

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

Adding additional factors to account for boundary marker ($\chi^2(0.09) = 2.00$, $p = .957$) and the interaction between boundary marker and prior manner demonstration ($\chi^2(1.98) = 5.00$, $p = .852$) did not lead to a significantly better model.

Combined analysis of manner and outcome depending on condition: We investigated the imitation of manner and outcome in the same model to compare the effects of the experimental conditions for both segments of the action sequence. In the first step, we compared a model containing the interaction between Condition (Baseline, Communicative, Control) and the action segment (Manner, Outcome) with a null model. Due to a failure of convergence at the null model, we compared both models with trial number as the only random effect. The model containing the interaction between Condition and Segment was significantly better than the null model ($\chi^2 (5) = 178.97$, p < .001). Within the full model, the only significant effect was that of Segment, with Outcome segments being significantly more likely to be imitated than Manner segments at baseline ($\beta = 3.13$, SE = 0.48, t = 6.49, p < .001). All other p values are p > 0.20. These results are also visualised by Figure 2 A and B.

Discussion

We were interested in the role of communicative and non-communicative signals in segmenting action sequences during child-directed action presentations. Our results showed that separating a short action sequence by briefly interrupting the action steps using child-directed speech and direct gaze increased children's imitation of the less salient sliding action. For the more salient hopping action, marking the boundary between the action manner and the outcome did not increase the level of imitation that was already high in the unmarked condition. Despite

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

this, we do not find such an effect for the non-communicatively marked control condition. These results provide support for the hypothesis that communicative signals help children segment action sequences, but these effects are only observable in less salient actions.

Our results suggest that any effect of communicative signals at event boundary markers only generalises to the sliding action in our study. In contrast, the hopping action remained unaffected, and was copied at a comparatively high level independently of condition. It is possible that the hopping action used in our study is more readily identified as an action unit due to its salience. Previous research has already highlighted the role of salience in children's imitation of actions. For example, toddlers between 12–30 months were more likely to imitate a hammering action compared to a less salient pulling action (Gampe et al., 2016) and 12-month-olds were better at learning to anticipate reaches towards large, compared to small objects (Adam et al., 2016; Henrichs et al., 2012). Therefore, the contribution of communicative signals to goal-directed action segmentation might be to identify action boundaries particularly in low salience actions.

As in previous studies (Carpenter et al., 2005; Southgate et al., 2009), children imitated the outcome to a very high degree and our manipulation did not affect whether children put the animal into the house. Previous studies attributed this difference to the importance of distinct goals versus the manner of the action (Bekkering et al., 2000; Carpenter et al., 2005). As previous studies by Hespos and colleagues (Hespos et al., 2009, 2010) have shown, children find it easier to recognise event with clear distinctive outcomes, rather than transitions between events.

These results are compatible with theories emphasising the segmentation and chunking of incoming information, such as Event Segmentation Theory (Zacks et al., 2007; Zacks &

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

Swallow, 2007) and domain-general theories of chunking and bottlenecks (Christiansen & Chater, 2016; Isbilen et al., 2020). Event Segmentation Theory also suggests that bottom-up and top-down processes influence the perception of an incoming stream of events (Zacks & Swallow, 2007), such as the action sequence observed by the children in our study. According to Event Segmentation Theory, the way that an event is segmented affects how it is interpreted. Consequently, the segmentation of an action should lead to a different interpretation of the action sequence. Whereas the uninterrupted action sequence is interpreted as "putting the animal into the house" (by any means), the interrupted event sequence is interpreted as two separate events of "sliding" and "putting the animal into the house." Crucially, because the hopping action is more readily recognised, more salient, more repetitive and potentially has clearer event boundaries, the hopping action may have already provided sufficient information to be recognised as its own separate action sequence, and additional segmentation cues were not helpful.

To explain why the communicative interruption increased the imitation of the sliding action when the non-communicative interruption did not, it is possible to appeal to low-level perceptual features of direct gaze and infant-directed speech or higher order inferences about the communicative intent of the interruption. Communicative signals have many properties that make them particularly suitable for signalling event boundaries. For example, infants show a stimuli-specific preference towards gaze (Farroni et al., 2000, 2002; Michel, Wronski, et al., 2017; Michel, Pauen, et al., 2017) and infant-directed speech (Dominey & Dodane, 2004), and direct gaze appears to interrupt working memory in adults (Wang & Apperly, 2016). However, so far we do not know why these specific properties of direct gaze and child-directed-speech make them suitable for action and event segmentation. In fact, other, non-social signals may have

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

similar effects on action segmentation and increase the imitation of action manner in this

paradigm as long as the boundary-marking event is clearly identifiable as an intentional action.

For instance, it is possible that a non-social, but clearly intentional 'beep' initiated by a button

press may have had a similar effect on action segmentation, comparable to the way that children

flexibly use an adult's intentional placement of an item as a referential signal to identify the

location of a hidden reward (Moore et al., 2013, 2015). If this is the case, even a simple pause

might be sufficient to induce the segmentation of an action, if marked as intentional. For

example, 6- (Sharon & Wynn, 1998) and 10– to 11-month old (Baldwin et al., 2001) infants that

have been familiarised with video sequences of everyday actions look longer if the video

sequence is paused within an intentional action, compared to a pause between intentional actions.

Therefore, the effect might not be specific to social signals and any pause may be sufficient to

segment non-salient actions for children and subsequently increase their imitation.

On the other hand, children may have interpreted the boundary marker in the control

condition differently to the communicative interruption, due to differences in the valence and

expressed intention. For example, children might have interpreted the non-communicative control

marker to indicate hesitation. In other studies, children of a similar age were sensitive to such

information and imitated the intended goals of failed actions, instead of faithfully copying

accidental actions (Carpenter et al., 1998). Therefore, they may have interpreted the part of the

action preceding the "Hmmm" as accidental and non-intentional, and were therefore less inclined

to imitate it. Similarly, the interruption used in the communicative and control conditions also

differed in terms of valence of the interruption itself, and children may have perceived the

"Hmmm" as showing less positive affect towards the previously executed action and this may

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

have lead to the—albeit not significant—decrease in imitated behaviour even for the hopping action. However, in the baseline condition and the communicative condition, identical linguistic information was used, and the enthusiastic "Wow!" was used in both the baseline and the communicative condition, albeit at different temporal locations. Therefore, any difference observed between these conditions would be due to the position of the cue, rather than the amount and direction of emotional valence expressed by the model in the interaction.

The communicative interruption of the action sequence in our study does not support the interpretation of direct gaze and social signals as markers of ostension in this particular context. According to empirical evidence, in young infants below 6 months of age, communicative signals function like a 'switch' (Csibra, 2010) for learning, that is either on or off, instead of providing a gradual response towards different levels of "communicative-ness" (Parise & Csibra, 2013). This function of communicative signals cannot explain the results, because (1) in our study, the entire action sequence was sandwiched between the experimenter speaking to the child using communicative signals at the beginning and the end of each trial. (2) The amount of communicative signals remained the same, and only the temporal location of one of them was varied. (3) Communicative signals typically *precede* relevant information, but in our study they followed the part of the action sequence that children were more likely to imitate.

Given that the interaction between the child and the experimenter was already marked as a pedagogical interaction by the presence of communicative signals at the beginning of each trial, any additional communicative signals could only be used to further interpret the presented action sequence. Children might have been actively looking for an interpretation of the communicative interruption, and determined that the signal intends to communicate the importance of the

otherwise ignored manner of the action as a relevant sub-action. Therefore, children may have used the already established relevance of the action demonstration to interpret the position of the communicative marker.

Previous studies on the effect of communicative signals as indicators of ostension have investigated the processing of objects (e.g. Michel, Wronski, et al. (2017); Yoon et al. (2008), but see Silverstein et al. (2019)), or relatively short and simple actions (Hernik & Csibra, 2015). But objects, unlike actions, are static. The temporal progression of an unfolding action sequence might benefit from regular bursts of communicative signals to retain focus and attention. If this is the case, the uninterrupted sliding action was potentially too long to sustain attention, and the communicative interruption provided a boost soon enough *after* the action segment had ended.

Whether we appeal to a lower-level or a higher-level interpretation of the results, our results also point towards a novel function of communicative signals within action sequences. Interrupting an action sequence communicatively provides important information on interpreting the meaning of the event by providing *temporal reference*, in addition to *spatial reference* that can be established through referential gaze to object locations (Butler et al., 2000) or pointing (Gliga & Csibra, 2009; Melinder et al., 2014; Morissette et al., 1995). In just the same way that gaze and pointing can be used to indicate the location of an object in space, communicative signals can be used to signal event boundaries. Following lower-level accounts, such as cognitive chunking, the communicative interruption provided a possibility to store the manner of the action separately rather than chunk it together with the more salient goal and subsequently forget it (Christiansen & Chater, 2016). Appealing to higher-level inferential accounts, children actively search for the relevance of the communicative interruption, taking into account its position within

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

the action stream. However, as discussed before, they are not using the communicative interruption as a signal of ostension [which already exists in all conditions], but a marker of an event boundary.

Conclusions

Our study shows that communicative signals can help to increase the imitation of a non-salient action unit in 18-month-old children. Importantly, our results also show that children do not use any interruption to segment these actions. In our study, the communicative marker, but not the non-communicative control marker, increased the imitation of the sliding action. These findings open up a new way of looking at the role of caregiver-child interactions by highlighting the role of structural information that guides pedagogical actions.

The preceding discussion of our results has also shown the importance of further investigating how children use social and non-social signals to segment events. As the results of our experiment only show an effect of marker location on one of the two actions, future investigations need to systematically broaden the range of actions and control their salience to generalise the findings to other actions.

In the current study we investigated the effect of communicative signals on the segmentation of a short action sequence. We have argued that this may be a crucial aspect of teaching novel actions to children. However, it is likely that these signals may also be useful in segmenting events in general, for example when teaching verbs.

Additionally, by looking at the information provided by caregivers in action teaching contexts, we find many similarities with the segmentation and processing of linguistic

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

information (see Christiansen & Chater, 2016; Hilton et al., 2019). Processing of action steps

may be incremental and lower levels of representation may be chunked together to form larger

representations of actions and their meanings. Children as well as adults may use such an

incremental, bottom-up approach to identify and segment basic action units and re-assemble them

into larger chunks to facilitate prediction. By drawing on the structural information that is

inherent in action-, event- and language processing, it may be possible to develop and enhance

domain-general models of learning.

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

Supplementary Information
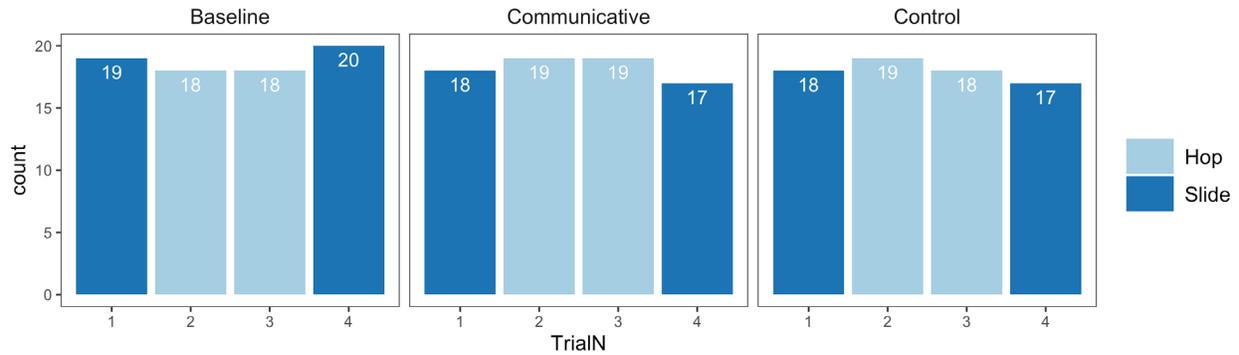
Figure 3.   Total number of trials contributed by participants across trials and conditions. In total, 111 trials were included for the hopping action and 109 trials for the sliding action.
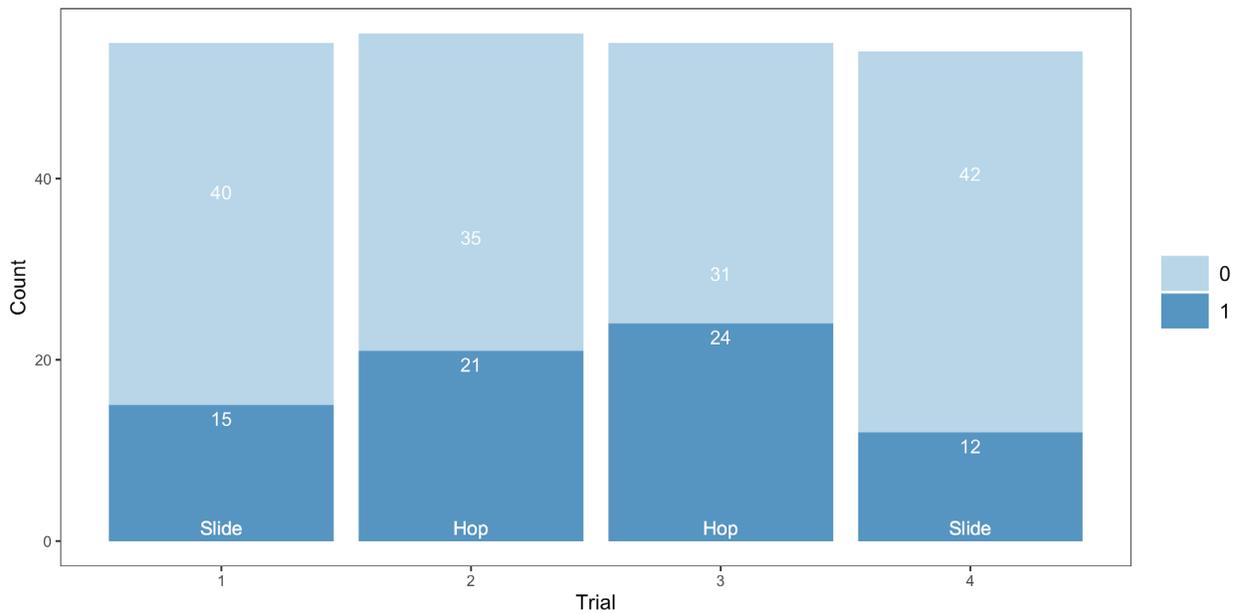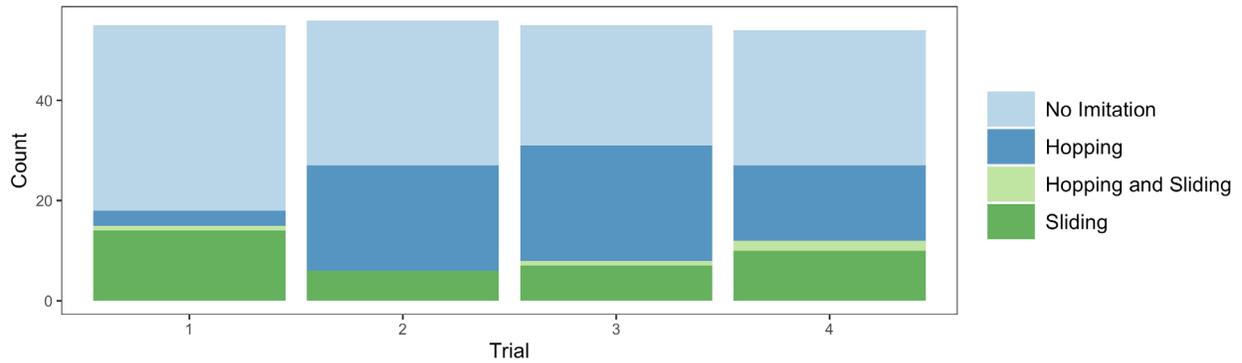

Figure 4.   Counts of imitated trials per trial number. If a child imitated the modelled action during a trial, their behaviour was coded as 1, if they did not imitate the behaviour, it was coded as 0.

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

Figure 5. Overview of all behaviours for each trial number. In the analysis, an action was only coded as 1 if it was modelled by the experimenter during the trial.
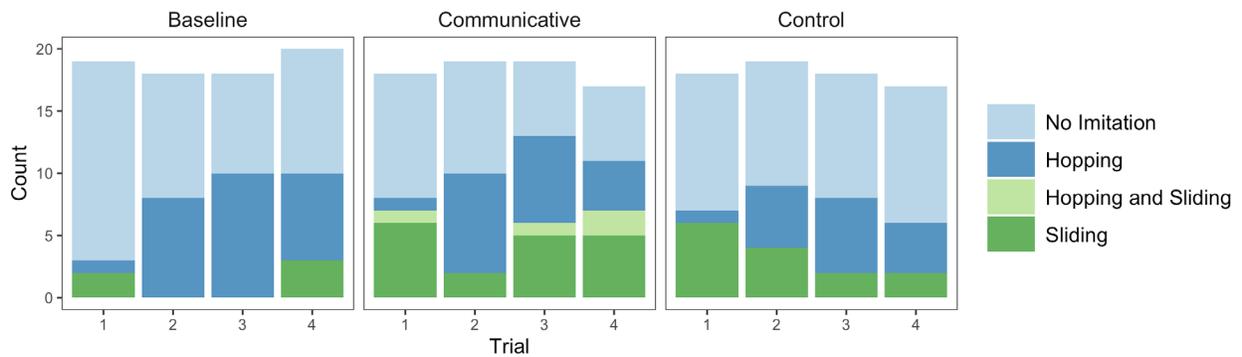
Figure 6. Overview of all behaviours for each trial number across conditions. In the analysis, an action was only coded as 1 if it was modelled by the experimenter during the trial.

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS
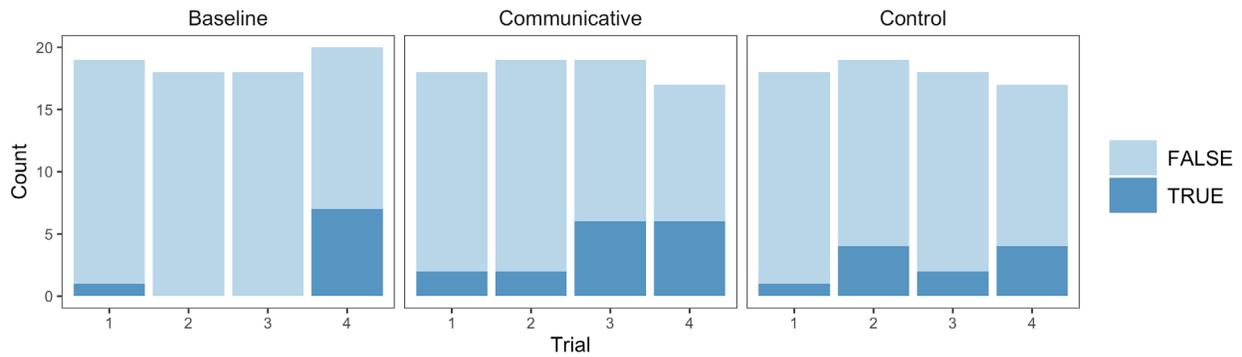
Figure 7.   Plot for children that imitated the inverse action, i.e. hopping during a sliding trial, sliding during a hopping trial (Label: TRUE; Label: FALSE indicates the number of trials where children either chose to imitate the target action or did not imitate at all). Trials 1 and 4 were sliding trials, trials 3 and 4 were hopping trials.

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 Behavior and Development*, *44*, 29–37. https://doi.org/10.1016/j.infbeh.2016.05.001

Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic action. *Child Development*, *72*(3), 708–717. https://doi.org/10.1111/1467-8624.00310

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure in mixed-effects models: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bavelas, J. B., Coates, L., & Johnson, T. (2002). Listener responses as a collaborative process: The role of gaze. *Journal of Communication*, *52*(3), 566–580. https://doi.org/10.1111/j.1460-2466.2002.tb02562.x

Bekkering, H., Wohlschläger, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *The Quarterly Journal of Experimental Psychology Section A*, *53*(1), 153–164. https://doi.org/10.1080/713755872

Biro, S. (2013). The role of the efficiency of novel actions in infants' goal anticipation. *Journal of Experimental Child Psychology*, *116*(2), 415–427. https://doi.org/10.1016/j.jecp.2012.09.011

Biro, S., Verschoor, S., Coalter, E., & Leslie, A. M. (2014). Outcome producing potential influences twelve-month-olds' interpretation of a novel action as goal-directed. *Infant Behavior and Development*, *37*(4), 729–738. https://doi.org/10.1016/j.infbeh.2014.09.004

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for 'motionese': Modifications in mothers' infant-directed action. *Developmental Science*, *5*(1), 72–83. https://doi.org/10.1016/j.infbeh.2003.09.004

Brand, R. J., Hollenbeck, E., & Kominsky, J. F. (2013). Mothers' infant-directed gaze during object demonstration highlights action boundaries and goals. *IEEE Transactions on Autonomous Mental Development*, *5*(3), 192–201. https://doi.org/10.1109/TAMD.2013.2273057

Brand, R. J., McGee, A., Kominsky, J. F., Briggs, K., Gruneisen, A., & Orbach, T. (2009). Repetition in infant-directed action depends on the goal structure of the object: Evidence for statistical regularities. *Gesture*, *9*(3), 337–353. https://doi.org/10.1075/gest.9.3.04bra

Brand, R. J., Shallcross, W. L., Sabatos, M. G., & Massie, K. P. (2007). Fine-grained analysis of motionese: Eye gaze, object exchanges, and action units in infant-versus adult-directed action. *Infancy*, *11*(2), 203–214. https://doi.org/10.1111/j.1532-7078.2007.tb00223.x

Butler, S. C., Caron, A. J., & Brooks, R. (2000). Infant understanding of the referential nature of looking. *Journal of Cognition and Development*, *1*(4), 359–377. https://doi.org/10.1207/S15327647JCD0104_01

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development*, *21*(2), 315–330. https://doi.org/10.1016/s0163-6383(98)90009-1

Carpenter, M., Call, J., & Tomasello, M. (2005). Twelve- and 18-month-olds copy actions in terms of goals. *Developmental Science*, *8*(1), F13–F20. https://doi.org/10.1111/j.1467-7687.2004.00385.x

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*, 1–72. https://doi.org/10.1017/s0140525x1500031x

Clark, H. H. (2003). Pointing and placing. In S. Kita (Ed.), *Pointing. Where language, culture, and cognition meet* (pp. 243–268). Erlbaum.

Csibra, G. (2003). Teleological and referential understanding of action in infancy. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 447–458. https://doi.org/10.1098/rstb.2002.1235

Csibra, G. (2008). Action mirroring and action understanding: An alternative account. In P. Haaggard, Y. Rossetti, & M. Kawato (Eds.), *Attention and performance XXII: Sensorimotor foundations of higher cognition* (pp. 435–459). Oxford University Press.

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Brain*, *25*(2), 141–168. https://doi.org/10.1111/j.1468-0017.2009.01384.x

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*, 148–153. https://doi.org/10.1016/j.tics.2009.01.005

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1149–1157. https://doi.org/10.1098/rstb.2010.0319

Csibra, G., & Shamsudheen, R. (2015). Nonverbal generics: Human infants interpret objects as symbols of object kinds. *Annual Review of Psychology*, *66*(1), 689–710. https://doi.org/10.1146/annurev-psych-010814-015232

Dominey, P. F., & Dodane, C. (2004). Indeterminacy in language acquisition: The role of child directed speech and joint attention. *Journal of Neurolinguistics*, *17*(2–3), 121–145. https://doi.org/10.1016/S0911-6044(03)00056-3

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*(2), 179–211. https://doi.org/10.1207/s15516709cog1402_1

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences*, *99*(14), 9602–9605. https://doi.org/10.1073/pnas.152159999

Farroni, T., Johnson, M. H., Brockbank, M., & Simion, F. (2000). Infants' use of gaze direction
to cue attention: The importance of perceived motion. *Visual Cognition*, *7*(6), 705–718.
https://doi.org/10.1080/13506280050144399

Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D., & Csibra, G. (2005).
Newborns' preference for face-relevant stimuli: Effects of contrast polarity. *Proceedings
of the National Academy of Sciences*, *102*(47), 17245–17250.
https://doi.org/10.1073/pnas.0502205102

Fausey, C. M., Jayaraman, S., & Smith, L. B. (2016). From faces to hands: Changing visual input
in the first two years. *Cognition*, *152*, 101–107.
https://doi.org/10.1016/j.cognition.2016.03.005

Gampe, A., Prinz, W., & Daum, M. M. (2016). Measuring action understanding: Relations
between goal prediction and imitation. *British Journal of Developmental Psychology*,
34(1), 53–65. https://doi.org/10.1111/bjdp.12125

Gergely, G., Bekkering, H., & Kiraly, I. (2002). Rational imitation in preverbal infants. *Nature*,
415(6873), 755–755. https://doi.org/10.1038/415755a

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational
action. *Trends in Cognitive Sciences*, *7*(7), 287–292. https://doi.org/10.1016/S1364-
6613(03)00128-1

Gliga, T., & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science*, *20*(3), 347–353. https://doi.org/10.1111/j.1467-9280.2009.02295.x

Goldstein, M. H., Waterfall, H. R., Lotem, A., Halpern, J. Y., Schwade, J. A., Onnis, L., & Edelman, S. (2010). General cognitive principles for learning structure in time and space. *Trends in Cognitive Sciences*, *14*(6), 249–258. https://doi.org/10.1016/j.tics.2010.02.004

Gredebäck, G., & Melinder, A. (2011). Teleological reasoning in 4-month-old infants: Pupil dilations and contextual constraints. *PLoS ONE*, *6*(10), e26487. https://doi.org/10.1371/journal.pone.0026487

Gredebäck, G., Stasiewicz, D., Falck-Ytter, T., Rosander, K., & Hofsten, C. von. (2009). Action type and goal type modulate goal-directed gaze shifts in 14-month-old infants. *Developmental Psychology*, *45*(4), 1190–1194.

Haensel, J. X., Smith, T. J., & Senju, A. (2021). Cultural differences in mutual gaze during face-to-face interactions: A dual head-mounted eye-tracking study. *Visual Cognition*, 1–16. https://doi.org/10.1080/13506285.2021.1928354

Henrichs, I., Elsner, C., Elsner, B., & Gredebäck, G. (2012). Goal salience affects infants' goal-directed gaze shifts. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00391

Hernik, M., & Csibra, G. (2015). Infants learn enduring functions of novel tools from action demonstrations. *Journal of Experimental Child Psychology*, *130*, 176–192. https://doi.org/10.1016/j.jecp.2014.10.004

Hespos, S. J., Grossman, S. R., & Saylor, M. M. (2010). Infants' ability to parse continuous

actions: Further evidence. *Neural Networks*, *23*(8-9), 1026–1032.

https://doi.org/10.1016/j.neunet.2010.07.010

Hespos, S. J., Saylor, M. M., & Grossman, S. R. (2009). Infants' ability to parse continuous

actions. *Developmental Psychology*, *45*(2), 575–585. https://doi.org/10.1037/a0014145

Hilton, M., Räling, R., Wartenburger, I., & Elsner, B. (2019). Parallels in processing boundary

cues in speech and action. *Frontiers in Psychology*, *10*.

https://doi.org/10.3389/fpsyg.2019.01566

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding

(TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*,

24(5), 849–878. https://doi.org/10.1017/s0140525x01000103

Hömke, P., Holler, J., & Levinson, S. C. (2017). Eye blinking as addressee feedback in face-to-

face conversation. *Research on Language and Social Interaction*, *50*(1), 54–70.

https://doi.org/10.1080/08351813.2017.1262143

Isbilen, E. S., McCauley, S. M., Kidd, E., & Christiansen, M. H. (2020). Statistically induced

chunking recall: A memory-based approach to statistical learning. *Cognitive Science*,

44(7). https://doi.org/10.1111/cogs.12848

Jones, S. S. (2007). Imitation in infancy: The development of mimicry. *Psychological Science*,

18(7), 593–599. https://doi.org/10.1111/j.1467-9280.2007.01945.x

Jones, S. S. (2009). The development of imitation in infancy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2325–2335. https://doi.org/10.1098/rstb.2009.0045

Király, I., Csibra, G., & Gergely, G. (2013). Beyond rational imitation: Learning arbitrary means actions from communicative demonstrations. *Journal of Experimental Child Psychology*, *116*(2), 471–486. https://doi.org/10.1016/j.jecp.2012.12.003

Koterba, E. A., & Iverson, J. M. (2009). Investigating motionese: The effect of infant-directed action on infants' attention and object exploration. *Infant Behavior and Development*, *32*(4), 437–444. https://doi.org/10.1016/j.infbeh.2009.07.003

Melinder, A. M. D., Konijnenberg, C., Hermansen, T., Daum, M. M., & Gredebäck, G. (2014). The developmental trajectory of pointing perception in the first year of life. *Experimental Brain Research*, *233*(2), 641–647. https://doi.org/10.1007/s00221-014-4143-2

Michel, C., Pauen, S., & Hoehl, S. (2017). Schematic eye-gaze cues influence infants' object encoding dependent on their contrast polarity. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-07445-9

Michel, C., Wronski, C., Pauen, S., Daum, M. M., & Hoehl, S. (2017). Infants' object processing is guided specifically by social cues. *Neuropsychologia*, *126*, 54–61. https://doi.org/10.1016/j.neuropsychologia.2017.05.022

Moore, R., Liebal, K., & Tomasello, M. (2013). Three-year-olds understand communicative intentions without language, gestures, or gaze. *Interaction Studies*, *14*(1), 62–80. https://doi.org/10.1075/is.14.1.05moo

Moore, R., Mueller, B., Kaminski, J., & Tomasello, M. (2015). Two-year-old children but not domestic dogs understand communicative intentions without language, gestures, or gaze. *Developmental Science*, *18*(2), 232–242. https://doi.org/10.1111/desc.12206

Morissette, P., Ricard, M., & Décarie, T. G. (1995). Joint visual attention and pointing in infancy: A longitudinal study of comprehension. *British Journal of Developmental Psychology*, *13*(2), 163–175. https://doi.org/10.1111/j.2044-835X.1995.tb00671.x

Parise, E., & Csibra, G. (2013). Neural responses to multimodal ostensive signals in 5-month-old infants. *PLoS ONE*, *8*(8), e72360. https://doi.org/10.1371/journal.pone.0072360

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Reid, V. M., Hoehl, S., Grigutsch, M., Groendahl, A., Parise, E., & Striano, T. (2009). The neural correlates of infant and adult goal prediction: Evidence for semantic processing systems. *Developmental Psychology*, *45*(3), 620–629. https://doi.org/10.1037/a0015209

Rutherford, M. D., & Przednowek, M. (2012). Fathers show modifications of infant-directed action similar to that of mothers. *Journal of Experimental Child Psychology*, *111*(3), 367–378. https://doi.org/10.1016/j.jecp.2011.10.012

Sage, K., & Baldwin, D. (2012). Exploring natural pedagogy in play with preschoolers: Cues parents use and relations among them. *Education Research and Perspectives*, *39*(1), 153–181.

Schaik, J. E. van, Meyer, M., Ham, C. R. van, & Hunnius, S. (2020). Motion tracking of parents' infant- versus adult-directed actions reveals general and action-specific modulations. *Developmental Science*, *23*(1), e12869. https://doi.org/10.1111/desc.12869

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*, 668–671. https://doi.org/10.1016/j.cub.2008.03.059

Senju, A., Csibra, G., & Johnson, M. H. (2008). Understanding the referential nature of looking: Infants' preference for object-directed gaze. *Cognition*, *108*, 303–319. https://doi.org/10.1016/j.cognition.2008.02.009

Sharon, T., & Wynn, K. (1998). Individuation of actions from continuous motion. *Psychological Science*, *9*(5), 357–362. https://doi.org/10.1111/1467-9280.00068

Silverstein, P., Gliga, T., Westermann, G., & Parise, E. (2019). Probing communication-induced memory biases in preverbal infants: Two replication attempts of Yoon, Johnson and Csibra (2008). *Infant Behavior and Development*, *55*, 77–87. https://doi.org/10.1016/j.infbeh.2019.03.005

Southgate, V., Chevallier, C., & Csibra, G. (2009). Sensitivity to communicative relevance tells young children what to imitate. *Developmental Science*, *12*, 1013–1019. https://doi.org/10.1111/j.1467-7687.2009.00861.x

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Wiley-Blackwell.

Sperber, D., & Wilson, D. (2002). Pragmatics, modularity and mind-reading. *Mind& Language*, *17*(1-2), 3–23. https://doi.org/10.1111/1468-0017.00186

Twomey, K. E., Ma, L., & Westermann, G. (2017). All the right noises: Background variability helps early word learning. *Cognitive Science*. https://doi.org/10.1111/cogs.12539

Twomey, K. E., & Westermann, G. (2017). Curiosity-based learning in infants: A neurocomputational approach. *Developmental Science*, *21*(4), e12629. https://doi.org/10.1111/desc.12629

Verschoor, S. A., Spapé, M., Biro, S., & Hommel, B. (2013). From outcome prediction to action selection: Developmental change in the role of action-effect bindings. *Developmental Science*, *16*(6), 801–814. https://doi.org/10.1111/desc.12085

Wang, J. J., & Apperly, I. A. (2016). Just one look: Direct gaze briefly disrupts visual working memory. *Psychonomic Bulletin& Review*, *24*(2), 393–399. https://doi.org/10.3758/s13423-016-1097-3

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Williamson, R. A., & Brand, R. J. (2014). Child-directed action promotes 2-year-olds' imitation. *Journal of Experimental Child Psychology*, *118*, 119–126. https://doi.org/10.1016/j.jecp.2013.08.005

Woodward, A. L., & Guajardo, J. J. (2002). Infants' understanding of the point gesture as an object-directed action. *Cognitive Development*, *17*(1), 1061–1084. https://doi.org/10.1016/s0885-2014(02)00074-6

Yoon, J. M. D., Johnson, M. H., & Csibra, G. (2008). Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences*, *105*(36), 13690–13695. https://doi.org/10.1073/pnas.0804388105

Zacks, J. M. (2004). Using movement and intentions to understand simple events. *Cognitive Science*, *28*(6), 979–1008. https://doi.org/10.1016/j.cogsci.2004.06.003

Zacks, J. M., Kumar, S., Abrams, R. A., & Mehta, R. (2009). Using movement and intentions to understand human activity. *Cognition*, *112*(2), 201–216. https://doi.org/10.1016/j.cognition.2009.03.007

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, *133*(2), 273–293. https://doi.org/10.1037/0033-2909.133.2.273

Zacks, J. M., & Swallow, K. M. (2007). Event segmentation. *Current Directions in Psychological Science*, *16*(2), 80–84. https://doi.org/10.1111/j.1467-8721.2007.00480.x

SOCIAL SIGNALS SEGMENT ACTIONS IN TODDLERS

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3–21. https://doi.org/10.1037/0033-2909.127.1.3

Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General*, *130*(1), 29–58. https://doi.org/10.1037/0096-3445.130.1.29