



## OPEN Six-month-old infants use cross-modal synchrony to identify novel communicative signals

Brock Ferguson<sup>1,3</sup>, Alexander LaTourrette<sup>2,3</sup> & Sandra R. Waxman<sup>1</sup>✉

Humans adopt non-linguistic signals, from smoke signals to Morse code, to communicate. This communicative flexibility emerges early: embedding novel sine-wave tones in a social, communicative exchange permits 6-month-olds to imbue them with communicative status, and to use them in subsequent learning. Here, to specify the mechanism(s) that undergird this capacity, we introduced infants to a novel signal—sine-wave tone sequences—in brief videotaped vignettes with non-human agents, systematically manipulating the socio-communicative cues in each vignette. Next, we asked whether infants interpreted new tone sequences as communicative in a fundamental cognitive task: object categorization. Infants successfully interpreted tones as communicative if they were produced in dialogues with one agent speaking (Study 1) or both agents producing tones (Study 2), or in monologues involving only a single agent (Study 3). What was essential was cross-modal temporal synchrony between an agent's movements and the tones: when this synchrony was disrupted (Study 4), infants failed in the subsequent task. This synchrony, we propose, licensed an inference that the tones, *generated* by an agent, were candidate communicative signals. Infants' early capacity to use synchrony to identify new communicative signals and recruit them in subsequent learning has implications for theory and for interventions to support infants facing communicative challenges.

**Keywords** Communication, Language, Categorization, Infants, Objects, Agency

Centuries of spirited inquiry have centered on the power of human language and its inextricable link to human cognition. Recent evidence has added a new dimension, revealing that very young infants, who cannot yet roll over in their cribs, have already begun to link the language they hear to their mental representations of the objects and events that they observe around them<sup>1</sup>. For infants between 3 and 7 months of age, simply listening to language supports their fundamental cognitive capacities, including object categorization and abstract rule-learning. Moreover, language supports cognition in a way that carefully matched non-linguistic signals (e.g., sine-wave tone sequences) do not<sup>2–7</sup>. This precocious language-cognition link, which offers powerful advantages for infants' subsequent learning, appears to emerge from a broader template: at three and four months, listening to vocalizations of non-human primates confers the same cognitive advantage as listening to language. But by 6 months, infants have tuned the link to cognition specifically to their native language<sup>8,9</sup>.

This raises a compelling question: Is infants' link to cognition, once they have tuned it specifically to language, reserved exclusively for language? This is unlikely. Across cultures, humans promiscuously appropriate otherwise non-linguistic signals, bending them to meet our communicative imperatives. For example, for centuries, we have intentionally manipulated smoke signals to transmit information over distances that exceed the reach of the human voice<sup>10</sup>. The same is true for whistling, another non-linguistic signal: at least 40 'whistle languages' have now been documented, including *Silbo Gomera*, a whistled register of Spanish that effectively communicates messages down wide ravines and across valleys in the Canary Islands<sup>11,12</sup>. A more recent example is Morse code, a system using clicks and taps created to carry precise messages across continents and oceans.

How early does this capacity develop? Can prelinguistic infants, like adults and older children, identify a novel non-linguistic signal and raise it to communicative status? Recent work suggests that they can. For example, if a novel signal is embedded in the context of a rich socio-communicative exchange between agents, infants as young as 6 to 10.5 months of age interpret that signal as communicative<sup>13,14</sup>. Moreover, the effect of this communicative exposure is robust enough to influence infants' interpretation of tone sequences later—in an entirely different object categorization task. More specifically, Ferguson and Waxman (2016)<sup>13</sup> introduced 6-month-olds to sine-wave tone sequences in a videotaped exposure vignette, embedding them in a social,

<sup>1</sup>Department of Psychology, Northwestern University, 2029 Sheridan Rd, Evanston, IL 60208, USA. <sup>2</sup>Department of Psychology, University of Southern California, 3620 McClintock Ave, Los Angeles, CA 90089, USA. <sup>3</sup>Brock Ferguson and Alexander LaTourrette share first co-authorship. ✉email: s-waxman@northwestern.edu

communicative dialogue between two young women, one speaking in English and the other ‘beeping’ in response. Next, infants participated in a classic object categorization task (See Fig. 1). During the familiarization phase, infants viewed a series of images from a single object category (e.g., DINOSAUR), each paired with the same sine-wave tone sequence. During the test phase, infants viewed two new images, presented simultaneously and in silence: a new member of the now-familiar category (e.g., a new dinosaur) and a member of a different category (e.g., a fish). The effect of this socio-communicative exposure vignette was sufficiently robust to permit infants to use tone sequences to succeed in the subsequent object categorization task—showing an above-chance preference for the novel category member at test. Moreover, this success was related specifically to the socio-communicative aspect of the exposure vignette. In a non-communicative control condition, infants heard the very same sine-wave tone sequences embedded in a similar exposure vignette, but this time, the tone sequences were uncoupled from the young women’s communicative exchange. Infants in this control condition, like those who receive no prior exposure to tone sequences, failed in the subsequent object categorization task<sup>5,6</sup>.

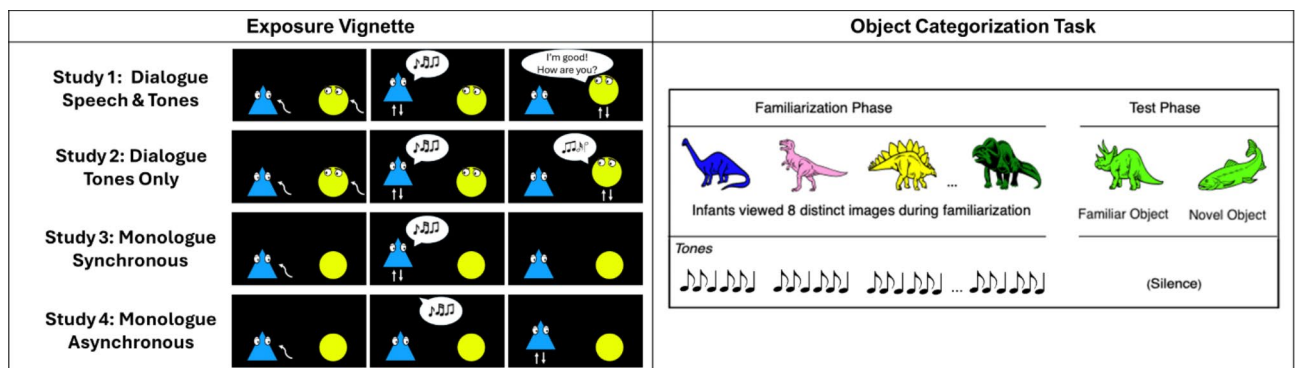
Thus, in a matter of minutes, 6-month-old infants can imbue novel sine-wave tone sequences with communicative status. But this transformation occurred only if the tones were embedded within a rich, human socio-communicative exchange. Mere exposure to the tones, uncoupled from the communicative exchange, was not enough<sup>2,13</sup>. This underscores infants’ keen sensitivity to socio-communicative cues and the benefits for subsequent learning<sup>15–20</sup>. Moreover, these results show that novel non-linguistic signals, once embedded in a rich socio-communicative exchange, are conferred sufficient communicative status to influence infants’ interpretation of that signal later – even in an entirely different task. This constitutes strong evidence that the ability to elevate a novel non-linguistic signal to communicative status does not require a fully mature linguistic system as a foundation; instead, it is a capacity that prelinguistic infants bring to bear in the process of language acquisition.

But which mechanism(s) license infants to infuse a novel signal with communicative status? To address this question, we manipulate systematically and precisely *which* socio-communicative cues are available to infants in an exposure vignette that includes sine-wave tone sequences. We then assess the impact of each cue on infants’ capacity to recruit tone sequences in a subsequent, unrelated object categorization task<sup>13</sup>. We focus on four candidate cues, asking whether the tone sequences must be produced by a human (Study 1), must be produced in concert with human speech (Study 2), must occur as part of a dyadic turn-taking exchange (Study 3), and finally, whether they must be produced synchronously with the motion of an ‘agent’ (Study 4). As we review below, infants’ sensitivity to each of these socio-communicative cues has been well-documented. At issue here is which of these cues, if any, are instrumental for 6-month-old infants to elevate sine-wave tone sequences to communicative status.

## Results

### Study 1: Must the novel signal be produced by a human?

Perhaps 6-month-old infants, who have already been immersed in rich socio-communicative experiences with other humans, will assign communicative status to novel signals only if they are produced by humans, as in Ferguson and Waxman’s communicative vignette<sup>13</sup>. Alternatively, perhaps infants will assign communicative status more broadly to signals produced by non-human agents as well, if these signals are produced in the context of a communicative exchange between the agents. After all, infants, like adults<sup>21</sup>, can attribute human-like goals, mental states or intentions to non-human agents—even those represented as simple geometric shapes with eyes<sup>17,22–27</sup>.

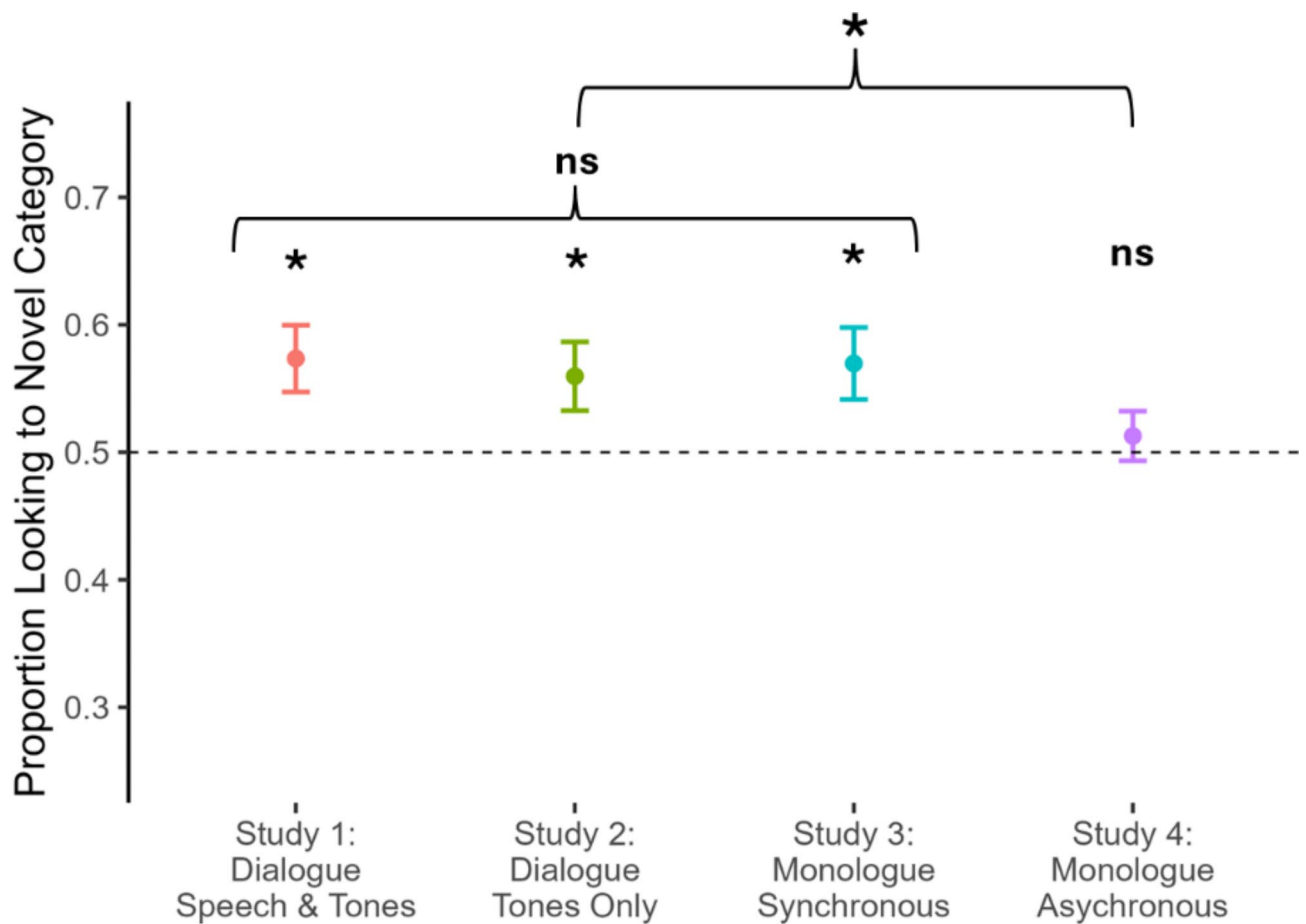


**Figure 1.** Study design. In each study, infants were first introduced to sine-wave tone sequences during an Exposure vignette. This vignette varied across studies. In Study 1, the vignette consisted of a dialogue between two agents, with one agent producing tone sequences and one producing English sentences. In Study 2, the English sentences were replaced with other, higher-pitched tone sequences. Study 3’s vignette featured a single agent moving in synchrony to the tone sequences (as if producing them), and Study 4 featured the same agent moving out of synchrony with the tone sequences. After the exposure vignette, infants participated in an object categorization task, where they were first familiarized with a series of exemplars paired with a novel tone sequence, and then presented with a novelty preference test.

To tease apart these alternatives, we retained the acoustic stream from Ferguson & Waxman's (2016) communicative vignette<sup>13</sup>, in which one human spoke and the other responded by beeping, but we changed the video stream by replacing each human with an animated geometric shape (a blue triangle and a yellow circle). Critically, each shape was endowed with two cues to its agency: eyes and an apparent capacity for self-generated motion<sup>28–31</sup>. Although what it means to be an 'agent' has long been a matter of debate, there is strong consensus that, at the most fundamental level, an agent is an individual capable of self-generated action<sup>32–35</sup>. Here, we use 'agent' in this fundamental sense, without committing to additional claims about mental states or desires.

The vignette began as each shape entered infants' visual field by moving on its own, bouncing to the center of the screen. At this point, infants listened to the same acoustic exchange as in Ferguson and Waxman's (2016)<sup>13</sup> communicative vignette, but now with the yellow circle bouncing in synchrony with the speech segments and the blue triangle with the tone sequences. Their bouncing movements coincided precisely with the acoustic input, simulating a dialogic exchange between two nonhuman agents. Immediately after watching this vignette, infants participated in a standard object categorization task (Fig. 1), in which each of the familiarization objects was accompanied by the same, consistently-applied sine-wave tone sequence—one that they had not heard in the preceding exposure vignette. Prior work has long established that communicative signals, like language, support infants in identifying object categories in this task while non-communicative signals like sine-wave tone sequences do not<sup>4,6</sup>.

Thus, if infants require that tones be produced by a human to elevate them to communicative status, they should fail in the current study's object categorization task. However, this was not the case (see Fig. 2). Instead, when the tone sequences were (apparently) produced by a blue triangle, infants nevertheless successfully formed object categories, revealing a reliable preference for the novel category member at test ( $M=0.57$ ,  $SD=0.13$ , 95% CI [0.519, 0.628]),  $t(23)=2.74$ ,  $p=.012$ ,  $d=0.57$ , all tests two-tailed). This suggests that after viewing the vignette, infants interpreted tone sequences as communicative and used them in the subsequent task to support object categorization—mirroring infants' success when the tones were produced by a human<sup>13</sup>.



**Figure 2.** Performance on categorization test in Studies 1–4. Infants showed a significant preference for the novel category member in Studies 1–3,  $ps < 0.05$ , indicating successful use of the novel tone sequences to learn the target category. However, in Study 4, infants failed to form the category, performing at chance level,  $p = .52$ . While infants' performance did not significantly differ among Studies 1, 2, and 3,  $ps > 0.70$ , infants' performance in Study 4 diverged significantly from performance in Studies 1–3 combined,  $p = .031$ . All tests are two-tailed. Error bars indicate  $\pm 1$  SEM.

### Study 2: Must the novel signal occur in concert with speech?

The exposure vignette in Study 1 included no human agents, but it did include another distinctly human cue: one shape ‘spoke’ in the infants’ native language (English) while the other beeped in response. By 6 months, infants know that speech is a communicative signal: one that is produced by, and directed toward, humans<sup>23,36–38</sup>. Perhaps, then, infants in Study 1 attributed communicative status to the tone sequences because they occurred in a ‘conversation’ that included human speech.

To test this, we modified the exposure vignette again, replacing all speech segments in the acoustic stream with sine-wave tone sequences. Thus, the yellow circle now bounced in synchrony with tone sequences (rather than segments of speech). These tones were higher-pitched than those produced by the blue triangle, to give the impression of a dialogue, with each shape beeping distinctly and in turn.

Infants who viewed this modified vignette, featuring tones produced in the absence of speech, succeeded in the subsequent categorization task. They demonstrated a reliable preference for the novel category member at test ( $M=0.56$ ,  $SD=0.13$ , 95% CI [0.504, 0.615]),  $t(23)=2.22$ ,  $p=.037$ ,  $d=0.45$ ; their preference did not differ statistically from that observed in Study 1,  $p=.68$ . Thus, 6-month-old infants do not require that a novel signal be produced in concert with human speech to grant it communicative status and recruit it for object categorization.

### Study 3: Must the novel signal be embedded in a dyadic turn-taking exchange?

The exposure vignette in Study 2 included neither human actors nor human speech, but it did include a dyadic turn-taking exchange. Perhaps this exchange licensed infants to interpret the tone sequences as communicative and thus link them to cognition. Indeed, infants’ sensitivity to turn-taking is well-documented at both behavioral<sup>14,39–41</sup> and neural levels<sup>42</sup>. Moreover, infants’ sensitivity to turn-taking has downstream effects, influencing infants’ willingness to learn from those individuals and engage with them socially<sup>14,43</sup>.

To assess whether it was the dyadic turn-taking between the two agentive geometric shapes that licensed infants’ interpretation of the tone sequences as communicative, we modified the exposure vignette again, this time removing turn-taking entirely. Specifically, we rendered the yellow circle silent and inert: eliminating its soundtrack, removing any apparent capacity for self-generated motions and also removing its eyes. Now, rather than bouncing along to reach the center of the screen, it simply dropped down from the top of the screen, resting motionless and silent for the duration of the vignette. Essentially, then, the blue triangle now appeared to be beeping, coordinated in time with its own bouncing movement, but absent any communicative partner.

Infants who observed this single animated geometric shape, bouncing and beeping entirely on its own, succeeded in the subsequent categorization task (see Fig. 2). They revealed a significant preference for the novel category member at test ( $M=0.57$ ,  $SD=0.14$ , 95% CI [0.511, 0.628]),  $t(23)=2.47$ ,  $p=.021$ ,  $d=0.50$ ). This preference did not differ statistically from those observed in either Study 1,  $p=.90$ , or Study 2,  $p=.78$ . Thus, even in the absence of dyadic turn-taking, 6-month-olds interpreted this novel signal as communicative and then linked it to an object category.

### Study 4: Must the novel signal occur in synchrony with an agent’s motion?

Why did observing a single animated shape, bouncing in synchrony with its own beeps, license infants to interpret the tones as a communicative signal? Notice that cross-modal temporal synchrony, like that displayed by the blue triangle, is abundant in infants’ social, communicative input. For example, our vocalizations occur in precise temporal synchrony with our mouth movements. Infants are sensitive to this cross-modal synchrony<sup>44–48</sup>. Moreover, synchrony engages infants’ attention<sup>49–53</sup> and supports their willingness to learn from and engage with others<sup>49–51</sup>. Thus, infants may use cross-modal temporal synchrony, especially if it is generated within an individual (here, synchrony between the tone sequences and the agent’s motions), as an especially powerful cue for identifying communicative signals.

If the cross-modal temporal synchrony of the nonhuman agent’s beeping and bouncing was instrumental in establishing the tone sequences as communicative in Study 3, then disrupting this synchrony should disrupt infants’ success in the subsequent categorization task. To test this, we modified the exposure vignette once more, this time shifting the timing so that the tone sequences *preceded* the bouncing of the nonhuman agent, rather than occurring together in synchrony. Importantly, this vignette provided the same amount of exposure to tone sequences as in all prior vignettes: only the temporal synchrony between the tone sequences and the nonhuman agent’s bounces was disrupted.

In sharp contrast to the preceding studies, infants in this asynchronous condition failed to identify the tone sequences as communicative and to recruit them in the subsequent object categorization task—performing at a level not significantly different from chance ( $M=0.51$ ,  $SD=0.10$ , 95% CI [0.464, 0.551]),  $t(23)=0.66$ ,  $p=.52$ ,  $d=0.13$ . In a subsequent analysis, designed to test the impact of synchrony directly, we compared infants’ performance in all previous synchronous conditions (Studies 1–3 combined,  $n=72$ ) to infants’ performance in the asynchronous condition (Study 4). This analysis yielded a significant, medium-sized effect of synchrony,  $t(54)=2.21$ ,  $p=.032$ ,  $d=0.44$ , Hedge’s  $g=0.44$ , indicating that infants in the synchronous conditions performed significantly better than those in the asynchronous condition (see Fig. 2). Separate, Bonferroni-corrected comparisons between Study 4 and each of the preceding studies yielded no significant differences,  $ps>0.06$ , though all differences were in the expected direction, and we note that these individual analyses have less statistical power than the combined analysis.

## Discussion

This work provides new insight into infants’ acquisition of a hallmark of human communication: our flexibility in recruiting non-linguistic, and otherwise non-communicative, signals to serve our communicative needs. Across four studies, we exposed 6-month-old infants to videotaped vignettes that included animated geometric shapes and sine-wave tone sequences; what varied was *which* socio-communicative cues were available in the vignette.

We then employed a stringent test, designed to uncover the conditions under which infants would interpret the tone sequences as communicative and go on to recruit them in a subsequent, unrelated task: an object categorization task featuring new tone sequences. Infants succeeded under a broad set of conditions, regardless of whether the tones were produced by humans (Study 1), occurred in a dialogue with human speech (Study 2), or occurred in a turn-taking exchange (Study 3). Thus, although 6-month-olds are sensitive to interactions with human interlocutors, to human speech and to turn-taking events, and although these cues are instrumental in early development, infants do not require these cues to identify a novel signal as communicative. Instead, a different cue – cross-modal temporal synchrony generated *within* a single individual – was essential. Only when this synchrony was disrupted during the vignette (i.e., the blue triangle was no longer beeping and bouncing in synchrony) did infants fail to recruit tone sequences in the subsequent categorization task (Study 4).

What is the best account for this effect of cross-modal temporal synchrony? There is considerable evidence that cross-modal synchrony enhances infant attention and supports learning<sup>7,49,51–53</sup>. One possibility is that the nonhuman agent's synchronous beeping and bouncing engaged infants' attention and enhanced their processing of tone sequences, and this, in turn, facilitated their use of tone sequences in the subsequent object categorization task. But it is also possible that this particular type of cross-modal synchrony – synchrony with the motions of an individual agent – licensed a higher-level inference: that the tones were *generated* by that agent and were therefore candidate communicative signals. That is, perhaps for infants to identify a novel signal as communicative, what matters is not only whether there is cross-modal synchrony between the sound and an object's motion, but whether that object is perceived to be an agent, capable of generating the tone sequences itself. The cross-modal temporal synchrony displayed in the vignettes of Studies 1–3 satisfied this latter condition: because it supported infants' identification of the blue triangle as the source of the tone sequences, these were therefore interpreted as candidate communicative signals. Disrupting this synchrony (Study 4) also disrupted infants' identification of the tone sequences as being produced by an agent, so infants did not interpret the tone sequences as candidate communicative signals.

This leads to a testable hypothesis: if tone sequences were introduced in vignettes that included non-agents, then infants should fail to interpret them as communicative signals. A recent study, considered in conjunction with the evidence reported here, offers *in principle* support for this speculation. Lany et al. (2022)<sup>7</sup> engaged younger, 3-month-old infants in a procedure that was structurally identical to the design implemented here. First, infants viewed a videotaped exposure vignette featuring sine-wave tone sequences; they then participated in Waxman and colleagues' object categorization task while listening to new tone sequences. But Lany et al.'s (2022) exposure vignettes differed from ours in one crucial way: the objects in their vignettes were decidedly not agents. Instead, they were artifacts (e.g., a truck) that did not move on their own: they were *caused* to move by an external mechanical source (a robotic claw) that moved them either synchronously or asynchronously with the tone sequences. After exposure to these vignettes, infants performed at chance in the subsequent categorization task, whether the motion was synchronous or asynchronous. Certainly, null effects must be interpreted with caution. Nevertheless, this null effect, coupled with the new evidence we report here, is intriguing because it is consistent with a proposal that for infants to attribute communicative status to a novel signal, what matters most is not only whether there is cross-modal synchrony between the sound and motion of an object, but whether that object is perceived to be an agent generating the tone sequences itself. If this is case, then 6-month-olds too should fail to elevate tone sequences to communicative status if they are introduced in vignettes featuring non-agents. To disentangle the effects of cross-modal synchrony and agency, it will be important to test this possibility in future research.

This work also offers insight into how infants might interpret vocalizations and other signals (e.g., gesture) as communicative, even when they are produced in the absence of any visible communicative partner (e.g., when an adult talks on the phone, or communicates via whistle language). We also suspect that communicative exposure to a signal has lasting power: once infants identify a particular signal as communicative, they continue to interpret it as communicative even when it occurs in sparser contexts without cross-modal synchrony, such as when infants listen to recorded stories or out-of-sight caregivers. The results reported in Studies 1–3 suggest that once infants were introduced to tone sequences in the context of a communicative vignette (Studies 1–3), they went on to interpret entirely new tone sequences, presented in the subsequent object categorization task, as communicative—even with no evidence of an agent as the tones' source and no evidence of cross-modal synchrony in the tones' production.

This work opens new research avenues as well. One outstanding issue is whether attributing communicative status to novel signals offers infants downstream cognitive and social advantages, beyond those documented in object categorization and abstract rule learning<sup>2,3</sup>. It will also be important to trace the contribution of infants' own experience in identifying candidate communicative signals, and the contribution of ostensive cues, including pointing and eye gaze (23). Eye-gaze serves as a case in point. Six-month-old sighted infants with sighted caregivers typically interpret eye-gaze as a communicative signal<sup>18,55–57</sup>. However, sighted infants with blind caregivers, who receive considerably less evidence that eye-gaze is communicative, are less sensitive to eye-gaze as a cue<sup>58,59</sup>. In addition, it will be important to identify how infants narrow the range of potential communicative signals to those that really are communicative. After all, many agent-produced signals (such as sneezing or coughing) are not, in fact, communicative, and prior work suggests that infants accurately reject them as communicative signals<sup>38,54</sup>. Future work might therefore examine how infants use their experience with the range of signals present in their environment to accurately identify which agent-produced signals are, in fact, communicative. For instance, we suspect variability in the signal will be instrumental in this process<sup>14</sup>. Finally, it will be important to consider whether what counts as a candidate communicative signal changes over development. Together, this work will deepen our understanding of our species' fundamental capacity to employ a range of linguistic and non-linguistic signals to communicate with each other.



## Methods

### Participants

In each study, we recruited 24 infants between 5 and 7 months of age to participate. A sample size of 24 provided over 90% power ( $\alpha=0.05$ ) for detecting a difference from chance as large as that observed in Ferguson and Waxman's Communicative condition ( $d=0.73$ ). All caregivers gave informed consent for their child's participation. All methods conform to the standards outlined in the Declaration of Helsinki and were approved by the Northwestern University Institutional Review Board.

In Study 1, participating infants had a mean age of 5.95 months ( $SD=0.29$ ; 11 females, 13 males). An additional 10 infants were excluded because they did not complete the test (6), were inattentive during familiarization ( $<25\%$  looking) (3), or experienced technical errors (1). In Study 2, we included 24 infants ( $M=5.97$  months,  $SD=0.28$ ; 11 female, 13 male). An additional 6 infants were excluded and replaced because they failed to complete the study (2), attend to the objects for more than 25% of familiarization (2), had parents interfere (1), or showed an outlier novelty preference, defined as  $>3$  MAD's and  $>2.5$  SD's from the mean (1). In Study 3, 24 infants were included in the final sample ( $M=6.09$  months,  $SD=0.33$ ; 12 female, 12 male), and an additional three infants were excluded for failing to complete the task. Finally, in Study 4, 24 infants ( $M=6.04$  months,  $SD=0.33$ ; 14 female, 10 male) participated. An additional 7 infants were excluded and replaced due to failing to complete the task (4), failing to look to the objects for at least 25% of the familiarization phase (2), or having an outlying novelty preference (1).

### Apparatus

Infants sat on their caregivers' laps approximately 110 cm from the center of a white projector screen. Caregivers wore blackout sunglasses so they could not see the screen. Sessions were recorded with a video camera positioned beneath the screen.

### Procedure

Following Ferguson and Waxman (2016)<sup>13</sup>, the study included an exposure phase followed by a categorization task (see Fig. 1).

#### Exposure

In each study, infants were first exposed to sine-wave tone sequences in a 2 min videotaped vignette. In Study 1, the vignette featured an audio stream identical to Ferguson and Waxman (2016): a turn-taking dialogue in which one speaker spoke English and the other responded in sine-wave tone sequences. However, in contrast to Ferguson & Waxman (2016), this audio stream was played in synchrony with the movement of non-human agents: a blue triangle and a yellow circle. Both agents had cartoon eyes. To give the impression of a turn-taking conversation, each agent moved in approximate synchrony with its part of the dialogue, (i.e., bouncing up and down when their tones or speech was playing). During the period of the dialogue in which the original human actors had stopped talking and waved to the infant, the shapes instead engaged in a brief chasing event before returning to their dialogue. All stimuli, data, and analyses are available on OSF: [https://osf.io/a3qz9/?view\\_only=6d9890d6717243968825d444de484ebf](https://osf.io/a3qz9/?view_only=6d9890d6717243968825d444de484ebf).

In Study 2, this vignette was modified by altering the audio stream: the English speech parts of the vignette were replaced with sine-wave tone sequences matched in length to the English sentences. These tone sequences were at a substantially higher pitch than the first set of tone sequences, giving the impression that the two agents had different "voices." In Study 3, we modified both the audio and visual stream of the vignette to remove the turn-taking aspect of the exposure vignette. Specifically, we made one of the agents (the yellow circle) into an inert object, removing the yellow circle's eyes and having it drop from the top of the screen and then remain inert for the duration of the vignette. We then removed the higher-pitched tone sequences (added in Study 2), so that the blue triangle appeared to be beeping to itself.

Finally, in Study 4, the audio and visual streams were identical to Study 3 but offset from each other, so that the tone sequences preceded the blue triangle's motions. This disrupted the impression that the tones were produced by the agent while maintaining the broader, predictive relation between the tones and the agent's motion.

#### Categorization task

Next, infants participated in the same categorization task as in previous studies<sup>4,6,8,13</sup>. During familiarization, infants saw 8 colored images of objects from a single category (either dinosaurs or fish, randomly assigned between-subjects). Each image appeared for 20 s and was paired with a single tone sequence (2.2 s in duration), presented at image onset and repeated 10 s later. This tone sequence differed in both rhythm and pitch (400 Hz, not varying) from those presented in the exposure phase.

After familiarization, infants began the test phase. Infants saw two new images: a member of the now-familiar category and a member of a novel category. These images were shown simultaneously and in silence until the infant accrued 10 s of looking. Left/right position was randomized across participants.

### Analysis

Infants' gaze locations during the familiarization and test phases were coded offline, frame-by-frame, by trained coders blind to our hypotheses. For each infant, we calculated a novelty preference score [Time Looking to Novel Category / (Time Looking to Novel Category + Time Looking to Familiar Category)]. Preliminary analyses generally yielded no significant relationship between infants' test performance and age or gender,  $ps > 0.05$ .

Finally, there was a significant effect of study on infants' attention during familiarization,  $F(3,92) = 4.08$ ,  $p = .009$ . However, Bonferroni-corrected post-hoc comparisons of the proportion of familiarization during which infants attended to the screen showed a significant difference only between Study 2 ( $M = 0.39$ ,  $SD = 0.10$ ) and Study 3 ( $M = 0.50$ ,  $SD = 0.12$ ),  $t(45) = 3.46$ ,  $p = .0012$ . Given that this difference was unanticipated and was the only difference that emerged between studies, we suspect it may be spurious. Notably, pairwise comparisons revealed no significant difference between infants' looking during familiarization in Study 4 ( $M = 0.47$ ,  $SD = 0.13$ ) compared to Study 1 ( $M = 0.42$ ,  $SD = 0.13$ ), Study 2, or Study 3,  $ps > 0.10$ . We also tested whether infants' time spent looking during familiarization was associated with their performance at test. These tests revealed no significant relationship in Studies 1, 2, or 4,  $ps > 0.10$ . In Study 3, we observed a negative association between looking during familiarization and test performance,  $\beta = -0.49$ ,  $SE = 0.22$ ,  $p = .041$ . This might reflect faster habituation by infants who learned the category more quickly. However, given that this association did not emerge in the other studies, and infants' looking at Study 4 did not differ from the other studies, the differences in infants' performance on the categorization test across studies are unlikely to be due to differences in infants' attention to the screen during familiarization.

## Data availability

All stimuli, data, and analyses supporting the findings of this manuscript are available on OSF: [https://osf.io/a3qz9/?view\\_only=6d9890d6717243968825d444de484ebf](https://osf.io/a3qz9/?view_only=6d9890d6717243968825d444de484ebf).

Received: 31 July 2024; Accepted: 4 November 2024

Published online: 13 November 2024

## References

- Perszyk, D. R. & Waxman, S. R. Linking language and cognition in infancy. *Annu. Rev. Psychol.* **69**, 231–250 (2018).
- Ferguson, B. & Lew-Williams, C. Communicative signals support abstract rule learning by 7-month-old infants. *Sci. Rep.* **6**, 25434 (2016).
- Rabagliati, H., Ferguson, B. & Lew-Williams, C. The profile of abstract rule learning in infancy: Meta-analytic and experimental evidence. *Dev. Sci.* **22**, e12704 (2019).
- Ferry, A. L., Hespos, S. J. & Waxman, S. R. Categorization in 3- and 4-month-old infants: an advantage of words over tones. *Child. Dev.* **81**, 472–479 (2010).
- Fulkerson, A. L. & Haaf, R. A. Does object naming aid 12-month-olds' formation of novel object categories? *First Lang.* **26**, 347–361 (2006).
- Fulkerson, A. L. & Waxman, S. R. Words (but not tones) facilitate object categorization: evidence from 6- and 12-month-olds. *Cognition*. **105**, 218–228 (2007).
- Lany, J., Thompson, A. & Aguero, A. What's in a name, and when can a [beep] be the same? *Dev. Psychol.* **58**, 209–221 (2022).
- Ferry, A. L., Hespos, S. J. & Waxman, S. R. Nonhuman primate vocalizations support categorization in very young human infants. *Proc. Natl. Acad. Sci. U S A.* **110**, 15231–15235 (2013).
- Perszyk, D. R. & Waxman, S. R. Infants' advances in speech perception shape their earliest links between language and cognition. *Sci. Rep.* **9**, 1–6 (2019).
- Gusinde, M. *North Wind—South Wind. Myths and Tales of the Tierra Del Fuego Indians (Translated from German)*. Translated by Kassel E. Röth. (1966).
- Busnel, R. G. & Classe, A. *Whistled Languages* vol. 13 (Springer Science & Business Media, 2013).
- Trujillo Carreño, R. *El Silbo Gomero: Análisis Lingüístico [The Gomeran Whistle: Linguistic Analysis]*. Translated by Brent, Jeff. Editorial Interinsular Canaria, Santa Cruz de Tenerife, Spain, (1978).
- Ferguson, B. & Waxman, S. R. What the [beep]? Six-month-olds link novel communicative signals to meaning. *Cognition*. **146**, 185–189 (2016).
- Tauzin, T. & Gergely, G. Variability of signal sequences in turn-taking exchanges induces agency attribution in 10.5-month-olds. *Proc. Natl. Acad. Sci.* **116**, 15441–15446 (2019).
- Gliga, T., Volein, A. & Csibra, G. Verbal labels modulate perceptual object processing in 1-year-old children. *J. Cogn. Neurosci.* **22**, 2781–2789 (2010).
- Novack, M. A. & Waxman, S. Becoming human: human infants link language and cognition, but what about the other great apes? *Philosophical Trans. Royal Soc. B.* **375**, 20180408 (2020).
- Okumura, Y., Kanakogi, Y., Kanda, T., Ishiguro, H. & Itakura, S. The power of human gaze on infant learning. *Cognition*. **128**, 127–133 (2013).
- Parise, E., Reid, V. M., Stets, M. & Striano, T. Direct eye contact influences the neural processing of objects in 5-month-old infants. *Soc. Neurosci.* **3**, 141–150 (2008).
- Yoon, J. M., Johnson, M. H. & Csibra, G. Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences* **105**, 13690–13695 (2008).
- Parise, E. & Csibra, G. Electrophysiological evidence for the understanding of maternal speech by 9-month-old infants. *Psychol. Sci.* **23**, 728–733 (2012).
- Heider, F. & Simmel, M. An experimental study of apparent behavior. *Am. J. Psychol.* **57**, 243–259 (1944).
- Choi, Y. & Luo, Y. 13-Month-Olds' understanding of social interactions. *Psychol. Sci.* **26**, 274–283 (2015).
- Csibra, G. Recognizing communicative intentions in infancy. *Mind Lang.* **25**, 141–168 (2010).
- Gergely, G., Nádasdy, Z., Csibra, G. & Biró, S. Taking the intentional stance at 12 months of age. *Cognition*. **56**, 165–193 (1995).
- Hamlin, J. K., Wynn, K. & Bloom, P. Social evaluation by preverbal infants. *Nature*. **450**, 557–559 (2007).
- Kampis, D., Duplessy, L., Askitis, H., Southgate, V. & D. & Training self-other distinction facilitates perspective taking in young children. *Child Dev.* **94**, 956–969 (2023).
- Luo, Y. Three-month-old infants attribute goals to a non-human agent. *Dev. Sci.* **14**, 453–460 (2011).
- Luo, Y., Kaufman, L. & Baillargeon, R. Young infants' reasoning about physical events involving inert and self-propelled objects. *Cogn. Psychol.* **58**, 441–486 (2009).
- Luo, Y. & Baillargeon, R. Can a self-propelled Box have a goal? *Psychol. Sci.* **16**, 601–608 (2005).
- Spelke, E. S., Phillips, A. & Woodward, A. L. Infants' knowledge of object motion and human action. in *Causal Cognition* (eds Sperber, D., Premack, D. & Premack, A. J.) 44–78 (Oxford University Press, (1995).
- Surian, L. & Caldi, S. Infants' individuation of agents and inert objects. *Dev. Sci.* **13**, 143–150 (2010).
- Ginet, C. *On Action* (Cambridge University Press, 1990).
- Lowe, E. J. *Personal Agency: The Metaphysics of Mind and Action* OUP Oxford. (2008).
- O'Connor, T. & Causality Mind, and Free Will. *Philosophical Perspect.* **14**, 105–117 (2000).

35. Schlosser, M. Agency. in *The Stanford Encyclopedia of Philosophy* (ed Zalta, E. N.) (Metaphysics Research Lab, Stanford University, (2019).
36. Legerstee, M., Barna, J. & DiAdamo, C. Precursors to the development of intention at 6 months: understanding people and their actions. *Dev. Psychol.* **36**, 627 (2000).
37. Vouloumanos, A., Druhen, M. J., Hauser, M. D. & Huizink, A. T. Five-month-old infants' identification of the sources of vocalizations. *Proc. Natl. Acad. Sci.* **106**, 18867–18872 (2009).
38. Vouloumanos, A., Martin, A. & Onishi, K. H. Do 6-month-olds understand that speech can communicate? *Dev. Sci.* **17**, 872–879 (2014).
39. Nadel, J., Carchon, I., Kervella, C. & Marcelli, D. Réserbat-Plantey, D. Expectancies for social contingency in 2-month-olds. *Dev. Sci.* **2**, 164–173 (1999).
40. Nguyen, V., Versyp, O., Cox, C. & Fusaroli, R. A systematic review and bayesian meta-analysis of the development of turn taking in adult-child vocal interactions. *Child Dev.* **93**, 1181–1200 (2022).
41. Tauzin, T. & Gergely, G. Communicative mind-reading in preverbal infants. *Sci. Rep.* **8**, 9534 (2018).
42. Forgács, B. et al. Semantic systems are mentalistically activated for and by social partners. *Sci. Rep.* **12**, 4866 (2022).
43. Wan, Y. & Zhu, L. Understanding the effects of rhythmic coordination on children's prosocial behaviours. *Infant Child. Dev.* **31**, e2282 (2022).
44. Curtindale, L. M., Bahrlick, L. E., Lickliter, R. & Colombo, J. Effects of multimodal synchrony on infant attention and heart rate during events with social and nonsocial stimuli. *J. Exp. Child Psychol.* **178**, 283–294 (2019).
45. de Hillairet, A., Tift, A. H., Minar, N. J. & Lewkowicz, D. J. Selective attention to a talker's mouth in infancy: role of audiovisual temporal synchrony and linguistic experience. *Dev. Sci.* **20**, e12381 (2017).
46. Lewkowicz, D. J., Schmuckler, M. & Agrawal, V. The multisensory cocktail party problem in children: Synchrony-based segregation of multiple talking faces improves in early childhood. *Cognition.* **228**, 105226 (2022).
47. Reynolds, G. D., Bahrlick, L. E., Lickliter, R. & Guy, M. W. Neural correlates of intersensory processing in 5-month-old infants. *Dev. Psychobiol.* **56**, 355–372 (2014).
48. Yeung, H. H. & Werker, J. F. Lip movements affect infants' Audiovisual Speech Perception. *Psychol. Sci.* **24**, 603–612 (2013).
49. Cirelli, L. K., Einarson, K. M. & Trainor, L. J. Interpersonal synchrony increases prosocial behavior in infants. *Dev. Science.* **17**, 1003–1011 (2014).
50. Cirelli, L. K., Wan, S. J. & Trainor, L. J. Fourteen-month-old infants use interpersonal synchrony as a cue to direct helpfulness. *Philosophical Trans. Royal Soc. B: Biol. Sci.* **369**, 20130400 (2014).
51. Rohlfing, K. J. & Nomikou, I. Intermodal synchrony as a form of maternal responsiveness: Association with language development. *Lang. Interact. Acquisition.* **5**, 117–136 (2014).
52. Bahrlick, L. E. & Lickliter, R. Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Dev. Psychol.* **36**, 190–201 (2000).
53. Gogate, L. J., Bolzani, L. H. & Betancourt, E. A. Attention to maternal Multimodal naming by 6- to 8-Month-Old infants and Learning of word-object relations. *Infancy.* **9**, 259–288 (2006).
54. Martin, A., Onishi, K. H. & Vouloumanos, A. Understanding the abstract role of speech in communication at 12 months. *Cognition.* **123**, 50–60 (2012).
55. Beier, J. S. & Spelke, E. S. Infants' developing understanding of Social Gaze. *Child Dev.* **83**, 486–496 (2012).
56. Farroni, T., Csibra, G., Simion, F. & Johnson, M. H. Eye contact detection in humans from birth. *PNAS.* **99**, 9602–9605 (2002).
57. Senju, A. & Csibra, G. Gaze following in Human infants depends on communicative signals. *Curr. Biol.* **18**, 668–671 (2008).
58. Senju, A. et al. Early social experience affects the development of Eye Gaze Processing. *Curr. Biol.* **25**, 3086–3091 (2015).
59. Verneti, A. et al. Infant neural sensitivity to eye gaze depends on early experience of gaze communication. *Dev. Cogn. Neurosci.* <https://doi.org/10.1016/j.dcn.2018.05.007> (2018).

## Acknowledgements

We thank the members of the Infant & Child Development Center, particularly Brooke Sprague, as well as the parents and infants who participated in this study.

## Author contributions

B.F. and S.W. designed the experiments, B.F. collected the data, A.L. and B.F. analyzed the data, B.F. and A.L. wrote the first draft of the manuscript, and all authors contributed to and reviewed the manuscript.

## Declarations

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to S.R.W.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024