Curiosity constructs communicative competence through social feedback loops

Julia A. Venditti, Emma Murrugarra, Celia R. McLean, and Michael H. Goldstein*
Department of Psychology, Cornell University, 270 Uris Hall, Ithaca, NY, United States
*Corresponding author. e-mail address: michael.goldstein@cornell.edu

Contents

Advances in Child Development and Behavior
ISSN 0065-2407, https://doi.org/10.1016/bs.acdb.2023.05.007

Abstract

One of the most important challenges for a developing infant is learning how best to allocate their attention and forage for information in the midst of a great deal of novel stimulation. We propose that infants of altricial species solve this challenge by learning selectively from events that are *contingent* on their immature behavior, such as babbling. Such a contingency filter would focus attention and learning on the behavior of social partners, because social behavior reliably fits infants' sensitivity to contingency. In this way a contingent response by a caregiver to an immature behavior becomes a source of learnable information – feedback – to the infant. Social interactions with responsive caregivers afford infants opportunities to explore the impacts of their immature behavior on their environment, which facilitates the development of socially guided learning. Furthermore, contingent interactions are opportunities to make and test predictions about the efficacy of their social behaviors and those of others. In this chapter, we will use prelinguistic vocal learning to exemplify how infants use their developing vocal abilities to elicit learnable information about language from their social partners. Specifically, we review how caregivers' contingent responses to babbling create information that facilitates infant vocal learning and drives the development of communication. Infants play an active role in this process, as their developing predictions about the consequences of their actions serve to further refine their allocation of attention and drive increases in the maturity of their vocal behavior.

1. Introduction

The world of a newborn infant provides endless sources of stimulation. Many sights, sounds, and sensations are experienced for the first time, at rates that are understandably assumed to be overwhelming (James, 1890). With so much available information, how do sensorily, socially, and motorically immature infants decide what to pay attention to and learn from? Efficiently allocating their attention is crucial to the acquisition of adaptive behaviors important for their survival as well as the development of social and cognitive abilities. We propose that infants become proficient learners in their first year of life by allocating their attention to a ubiquitously available source of information: their caregivers. In addition to provisioning nourishment, shelter and safety, caregivers provide responsive behaviors that highlight what information is important for infants to pay attention to.

By engaging in social interactions with caregivers, infants receive timely and relevant responses to their immature behaviors, such as babbling. Inherent in these interactions is contingency, or the temporal relatedness of a caregiver's response to their infant's behavior. Infants prefer stimuli that they perceive as informative (Kidd, Piantadosi, & Aslin, 2012), including those that are contingent on their own behaviors (Heyes, 2016). The preference for stimuli that they can elicit guides what and who they learn from. Most information that is contingent in infants' environments comes from social partners. Contingency drives infants' attention to socially relevant information and thus plays an important role in social learning. Furthermore, by producing immature behaviors that caregivers respond to, infants create opportunities for social learning.

Babbling is one such immature behavior that organizes social interactions and infants' learning environments. A key component of vocal learning is the production of these non-cry, prelinguistic vocalizations during vocal turn-taking interactions with caregivers. By using contingency as a filter for organizing attention and directing learning, these interactions result in infants acquiring socially relevant sounds from their caregivers. Therefore, beyond motor practice, babbling functions as a tool with which infants probe their social environments and facilitate their vocal learning.

Socially guided vocal learning proceeds from a feedback loop between caregiver responsiveness and infant attention, ultimately resulting in the ability to produce speech-like sounds. Infants' prelinguistic vocalizations elicit responses from caregivers, and infants subsequently integrate the information from those responses into their future vocalizations. Over time, infant vocalizations adopt more mature and socially relevant features and are increasingly effective at reliably eliciting future responses (Goldstein & Schwade, 2010). This process of vocal learning through social feedback demonstrates how an extended period of behavioral flexibility early in development affords infants opportunities to learn species-typical behaviors from interactions with their caregivers.

Prediction of the social consequences of vocalizing, such as receiving contingent speech from a caregiver, plays a key role in the function of the feedback loop. Vocal turn-taking interactions with caregivers allow infants to test the social efficacy of their immature vocalizations. When infants receive contingent responses from their caregivers, they update their own predictions as to how their prelinguistic vocalizations can change their environment.

Prioritizing behaviors that increase prediction accuracy, like producing speech-like sounds that have reliably elicited responses in the past, is an important strategy for vocal development. This domain-general strategy is the defining feature of curiosity-driven learning (e.g., Oudeyer & Smith, 2016). We propose that curiosity-driven learning is a process by which infants use contingency to learn within the social feedback loop.

Our chapter will highlight the connected functions of immature behavior, contingent social feedback, and infant curiosity in socially guided vocal learning. Using examples across fields and species, we will outline how gradual development and extended parental care create a developmental niche in which immature behaviors, and the contingent responses they elicit, structure early learning environments, resulting in adaptive developmental outcomes.

2. Learning within a co-developing social system

Humans are altricial. This means that, for an extended period of time after they are born, infants are sensorily and motorically immature and must rely on others for survival. This is in contrast to more precocial animals, such as deer that can get up and walk shortly after birth, who are born with more functional motor and sensory abilities. While it may seem like a disadvantage to have such a long period of immaturity, remaining reliant on parental care is actually an important feature of an infant's developmental niche.

A developmental niche describes an animal's ability to perceive and use available information in their environment in order to learn (Alberts, 2008). Fig. 1 illustrates the developmental niche of human infant vocal learners in

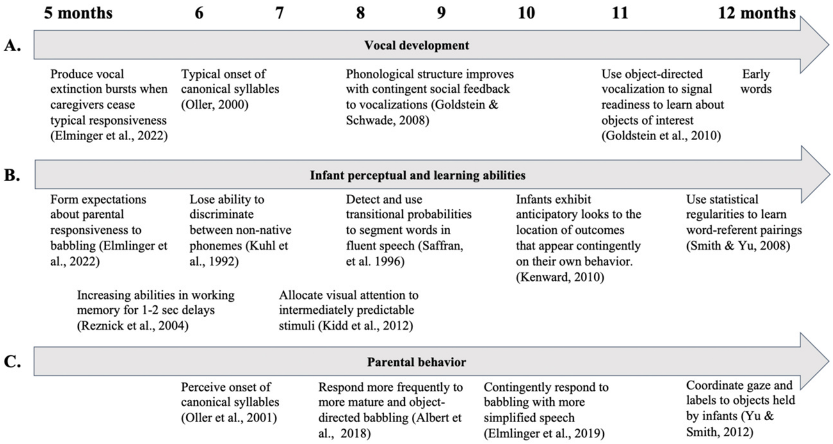

Fig. 1 The developmental niche of human infant vocal learners. Infants' developing abilities in the vocal (A), and perceptual and learning domains (B) align with the social information in their environment made available to them by responsive caregivers (C). Placement of abilities along the 5–12 month timeline relates to instances in the literature in which these abilities have been demonstrated. (Kenward (2010); Reznick, Morrow, Goldman, & Snyder (2004); Saffran, Aslin, & Newport (1996); Smith & Yu (2008); Yu & Smith (2012)).

the latter half of their first year. The niche is characterized by alignment between infants' developing vocal abilities, their changing learning capacities, and their parents' changing responsiveness. Infants are born into their niche with specific perceptual and learning abilities that shape how they interact with the world around them (Fig. 1B; Gibson, 1988).

The developmental niche of a human infant promotes social learning. When an infant acts on their social world, they can explore their affordances, which include caregivers who scaffold their continued learning (Faust, Carouso-Peck, Elson, & Goldstein, 2020). Scaffolding is made possible by caregivers' ability to attend to and learn from the information infants generate (Fig. 1C). With access to social partners who are sensitive to immaturity, infants are able to make use of their developing skills to gradually learn about the world.

The caregiver-infant dyad is a co-developing social system, in which infants and caregivers change their behaviors as a function of each other's actions. This co-development happens both in the moment during individual interactions, and over longer developmental timescales. These dynamic interactions shape how infants learn from the information parents provide in response to their immature behaviors. Likewise, these interactions also structure the complexity of the responses that parents scaffold during communication. As infants' attentional and communicative skills improve over the first two years, contingent interactions change to allow for appropriate scaffolding of maturing communicative and social capacities (Masek et al., 2021). For example, during naturalistic interactions, parents of toddlers tend to provide object labels for toys that are held in their own or their infants' hands, creating multimodal contingencies between visual and auditory information (Schroer & Yu, 2022). Experience with responsive caregivers is important for infants to become responsive social partners themselves, as caregiver contingency in interactions with infants was related to responsiveness of infants to caregiver behaviors (Kuchirko, Tafuro, & Tamis LeMonda, 2018). Social learning is thus embedded in a dynamic set of interactions with caregivers whose behaviors are emerging and changing in real time and over developmental time.

Vocal learning is a prime example of gradual social learning in the infant's developmental niche. Vocal learning is characterized as a developing feedback loop in which infants' immature vocal behavior elicits caregiver responses that guide infants' subsequent vocalizations (Fig. 2). The characteristics of prelinguistic vocalizations are constrained by infants' developing vocal apparatus as well as their limited social and learning

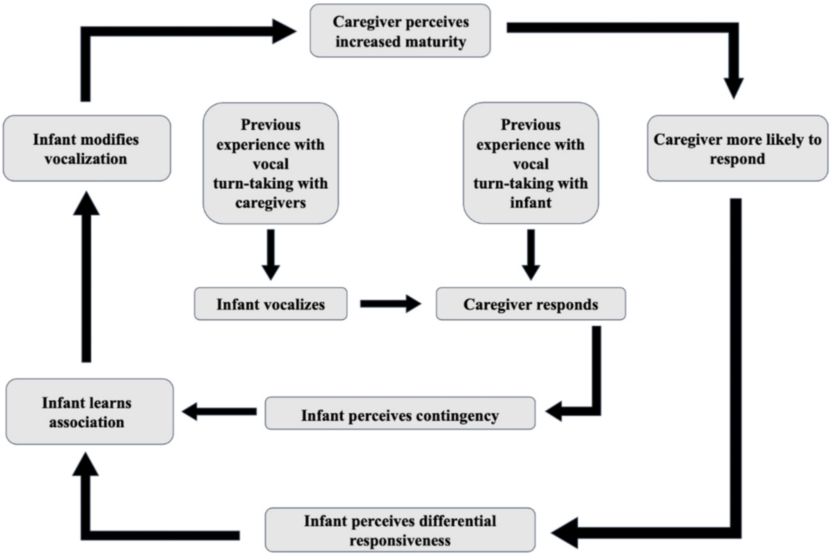

Fig. 2 Social feedback loop of socially guided vocal learning. When infants vocalize, caregivers contingently respond, which provides infants with information that their vocal behavior elicits such responses. Subsequent vocalizations integrate temporal and phonological features of contingent adult responses and produce more mature vocal forms. Parents perceive this increased maturity and begin to differentially respond to more mature vocal forms. Infants learn this association and use that to guide their vocal behavior to be more socially relevant.

capacities. When infants vocalize, their parents perceive the maturity of their infants' vocalizations (Oller et al., 2001). This guides the way caregivers interpret and respond to infants. More mature vocalizations lead parents to use those sounds as cues when responding to infants (Albert et al., 2018; Goldstein & West, 1999). When infants receive caregiver responses contingently on their own vocalizations, this serves as feedback that infants integrate into their subsequent vocalizations (Goldstein & Schwade, 2008). Infants' vocalizations begin to incorporate the phonological patterns of their caregivers' speech, as well as the features of their own behaviors that elicited responses from caregivers in the first place. This feedback loop (Fig. 2) results in mature behaviors over time as infants and caregivers engage in vocal turn-taking and infants acquire the social and vocal means to produce more mature vocalizations.

Crucially, the social feedback loop works because of the alignment between the complexity of caregivers' responsiveness and infants' developing abilities. This alignment is not a developmentally static phenomenon; it changes as a function of infant and parental development. For example, developmental changes in infant attention and memory facilitates their ability to participate in complex interactions as well as their ability to learn from them (Masek et al., 2021). In the feedback loop, immature behaviors have functional significance because caregivers respond to them, and infants learn from those responses. In the following section we outline infants' early capacities and describe their fit with caregiver behaviors.

3. Early perceptual abilities

Despite early immaturity, infants are not born completely blank slates. By the time they are born, they already display biases that shape their continued learning. For example, newborn infants show a visual preference for faces, differentially attending to top-heavy, asymmetrical, face-like shape configurations over shapes that do not resemble faces (Simion et al., 2002). This bias may also be built by early postnatal sensory experiences. Newborns' limited visual acuity constrains their visual attention to objects that are in close proximity to their face. Since they spend much of their early infancy being held, the face of the caregiver holding them is often the closest figure they can see. This early postnatal experience may also account for their attentional biases to faces (Mondloch et al., 2013). Having a bias for faces is important for social development because faces are rich sources of information.

Infants are also predisposed to pay more attention to their mother's voice than other sounds. The prenatal environment exposes infants to sounds their mothers make. Newborns show a greater preference for their mother's voice over other female voices (DeCasper & Fifer, 1980), and even show a preference for hearing familiar book passages that their mothers read aloud during the third trimester of pregnancy over unfamiliar passages (DeCasper & Spence, 1986). Research on newborn infants' preferences highlights the energy they will expend to elicit familiar, and possibly more predictable stimuli. For example, non-nutritive pacifier sucking rate increases when sucking delivers familiar auditory stimuli that infants heard while still *in utero* (DeCasper & Fifer, 1980). Infants that are only days old will display early locomotor behavior, aided by a mini skateboard, to move toward a speaker playing IDS in their native language (Hym et al., 2023), further demonstrating that their prenatal auditory experiences shape preferences and drive behaviors to elicit preferred stimuli.

Acquiring language via social learning is only an efficient strategy if individuals can allocate their attention to sources that are producing learnable linguistic input relevant to their ambient language environment. In their first 6 months of life, infants demonstrate the ability to discriminate between distinct phonemes of both native and non-native languages early in development (Doupe & Kuhl, 1999). Infants guide their exploration of their learning environments by preferentially attending to features that have learnable statistical regularities (Kidd et al., 2012). Attentional biases for socially relevant stimuli may be driven by the learnable regularities they exhibit. In addition to possessing certain attention-garnering visual and acoustic features, like familiar faces and voices, another learnable and interesting feature of social partners may be their tendency to respond contingently to infants' behaviors. The contingency of caregiver responsiveness during social interactions has been proposed to highlight what features are important for infants to attend to, resulting in perceptual learning and narrowing (Kuhl, Tsao, & Liu, 2003). In this way, contingency plays a key role in orienting infant attention to learnable and relevant information and is an important aspect of infants' communicative development (Kuhl, 2007). Thus, infants enter the second half of their first year with phonetic perception that is constrained to sounds that are relevant for learning the languages they are exposed to. This restriction likely promotes more efficient language learning later in development (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992). For contingency to impact learning, infants must attend to the temporal and semantic relatedness of contingent responses to their own behaviors (Tamis-LeMonda et al., 2014).

4. Contingency

When events occur in close proximity to one another, infants attend to the temporal relationship between those events and form expectations that those events will co-occur in the future. This sensitivity to temporally proximal events is an essential mechanism for shaping attention and driving reinforcement learning (Skinner, 1948). Crucial to contingent interactions is the size of the temporal window between events. The more time that passes between two events, the less likely they are understood to be causally related. However, events that are too closely related are perceived as overlapping or even intrusive. Both infants and caregivers are sensitive to the length of this delay during vocal turn-taking interactions (Striano et al., 2006).

Observations of caregivers' interactions with their 4 month-old infants showed that contingent responses typically occur within 2 s of preceding vocal behavior (Van Egeren et al., 2001). This contingency window is thought to coincide with the length of time that infants can attend to and hold connected events in working memory (Millar & Watson, 1979; Ross-Sheehy & Newman, 2015).

4.1 Social contingency

Contingency in social interactions entails timely response to the behavior of another individual. Social interactions are important opportunities for infants to perceive the contingency of others' actions as it relates to their own behaviors. Contingent responses to infants are characterized by their timeliness and semantic relatedness to infant behaviors (Tamis-LeMonda et al., 2014). These features of input quality play a crucial role in infant learning. Contributing to the quality of infants' language learning environments is their access to contingent social interactions with their caregivers (Golinkoff, Can, Soderstrom, & Hirsh-Pasek, 2015). For example, infants can become sensitive to different phonemes in a non-native language when they have engaged in live interactions with adult speakers using that language, but do not learn to do so from pre-recorded material (Kuhl et al., 2003). Engaging with contingent and sensitive social partners also facilitates early word learning. Infants have demonstrated the ability to learn words from both in- person and live video interactions with adults, but not from pre-recorded video with the same amount of content (Roseberry, Hirsh-Pasek, & Golinkoff, 2014). Word learning is also facilitated by semantically contingent object labels related to objects that infants are focused on (Tamis-LeMonda et al., 2014).

The nature of infants' interactions with caregivers changes as infants acquire advanced attentional and communicative skills. For example, early in infancy, social interactions between infants and caregivers occur when they are face-to-face, but throughout infancy interactions occur in a more triadic fashion with joint attention focused on objects or manual actions (Deák et al., 2014). Contingency and developing attention have been proposed to work together bidirectionally to enable this shift. Attention to contingent stimuli, likely driven by early attentional preferences to faces and eye gaze, encourages engagement in contingent social interactions, and engagement in such interactions promotes further attention and learning from them over the course of development (Masek et al., 2021).

4.2 Learning from contingency in non-human animals

Contingency is a domain- and species-general mechanism that guides learning in altricial young. This is highlighted by evidence across species that shows that immature social partners can learn from their mature social partners' contingent responses. Social interactions in particular facilitate communicative learning. In zebra finches *(Taeniopygia guttata)*, for example, visual and acoustic interactions with song tutors enable young males to learn species-typical songs (Eales, 1989). Even non-singing female songbirds guide juvenile males to learn song that is more appealing by providing rapid arousal movements contingent on immature song (Carouso-Peck & Goldstein, 2019; West & King, 1988). Juveniles can also learn to copy songs from recordings they elicit via pressing a key (Adret, 1993; Tchernichovski et al., 2001), suggesting that the contingent, temporally-locked cues they receive during social interactions facilitates learning from them. More recent research on zebra finches found that the vocalizations of deaf birds could be shaped by song-contingent visual stimuli in the absence of auditory feedback (Zai et al., 2020).

These examples support the idea that the act of eliciting contingent stimuli through interaction with the environment shapes communicative development across altricial species, including humans. The ability to elicit such stimuli, in combination with contingency perception and a preference for learning from contingent sources, are important features of the human infants' developmental niche. Paired with a social environment that includes sensitive and contingent social partners, infants are well suited to learn from social interactions.

Within the social feedback loop, infants' perception of contingency of their caregivers' responses provides information about the social impact of their vocalizations. Some sounds will be differentially responded to, such that immature sounds that used to receive a response may be less successful once infants have more mature sounds in their repertoires. Differential contingent responsiveness can be thought of as *acoustic scaffolding*, encouraging the use of more mature sounds over time as infants are able to produce them.

5. Parental scaffolding

Prioritizing learning from contingent sources is only efficient for learners if there are also mature adults that provide cues from which to learn. Parental scaffolding is characterized by modifications that caregivers make to account for their children's maturity and skill levels (Wood et al., 1976). Parents' responses to infant behavior reflect their infants' lack of knowledge and mastery in physical and social domains. The ability to perceive an immature partner's maturity and provide support for the emergence of more mature behaviors is the cornerstone of the altricial caregiver's own niche.

For caregivers' contingent responses to serve as feedback for infants, they must respond to their infants' behaviors in ways that are sensitive to their infants' immediate actions, focus, and maturity. Responsiveness that is dynamically modulated by infants' behaviors is a key feature of parental sensitivity, which is predictive of positive language learning outcomes (Baumwell, Tamis-LeMonda, & Bornstein, 1997). Like contingency, scaffolding is a domain-general mechanism that facilitates learning in altricial young across species.

5.1 Examples of scaffolding in non-human animals

Parents scaffold behaviors to promote learning relevant skills that an immature partner needs for survival. In meerkats, for example, adults facilitate learning of prey handling skills in nutritionally-dependent pups by first providing them with dead and disabled prey, specifically venomous scorpions. The adults use cues from pup calls to determine their pups' maturity and readiness for more dangerous provisions (Thornton & McAuliffe, 2006). These age-appropriate responses to pup begging calls guide the efficient learning of complex foraging skills, enabling pups to become nutritionally independent. This study highlights the sensitivity of caregivers to their pup's maturity through vocal signaling.

In other species, including bats and songbirds, researchers have found that adults adjust spectral and temporal attributes of their juvenile-directed vocalizations, such that they differ from adult-directed vocalizations (Chen, Matheson, & Sakata, 2016; Fernandez & Knörnschild, 2020). These modifications have implications for the learnability of those tailored signals. For example, pupil- directed song from adult male zebra finches was associated with higher attention and superior song learning from juveniles, as opposed to non-pupil directed song (Chen et al., 2016). These examples underscore the importance of dynamically modulated interactions between adult and offspring, especially during critical periods of development. Altriciality provides an extended period in which these fine-tuned interactions can occur.

5.2 What parents perceive

Both in humans and non-humans, caregivers actively attend to the maturity of the cues that their infants provide. This is an essential skill for being able to provide appropriate and contingent information that infants can detect. However, there is little work studying the role of parental perception in constructing informative responses, and this remains an outstanding area for future research. What we do know is that parents are sensitive to various qualities of the sounds that their infants produce. This sensitivity is thought to guide caregivers' responsiveness to infant vocalizations as they change in quality and frequency over the course of development. For example, hyperphonated cries are rated by parents as being more distressing, related to sickness, or urgent compared to less phonated cries (Crowe & Zeskind, 1992; Zeskind & Shingler, 1991). As infants get older and begin babbling, there is evidence that parents can accurately track the onset of canonical syllables in their infant's vocal repertoire (Oller et al., 2001). While there is still much work to be done on how parents perceive the cues that signal mature development, we do know that parents are shaping their communication to meet the learning demands of infants, as we describe below.

5.3 Infant-directed speech

Infant-directed speech (IDS) refers to the stereotyped way in which adults acoustically modify their speech to infants (e.g., Hilton et al., 2022). IDS is an example of scaffolding because the acoustic features of parents' speech align well with infants' abilities and communicative needs in the moment. Attending to infants' behaviors results in real-time changes in the pitch (Smith & Trainor, 2008) and hyperarticulation (Lam & Kitamura, 2012) of parents' IDS. In this way, the well-known characteristics of IDS are not fixed, but are tightly related to cues from infants during interactions. Infant behavior is thus a proximal force of change for parental behavior.

A proximal cue from infants that guides parents to produce more learnable IDS is babbling. Contingent speech in response to infant babbling is linguistically simpler than IDS that is not contingent on babbling. While prosody remains the same, linguistic differences between contingent and non-contingent speech include reductions in the proportion of single-word utterances, the number of unique words used, and the mean length of utterances. Thus, infants receive lexically simplified and shorter utterances from parents in response to their babbling (Elmlinger et al., 2019). This effect is observed cross-culturally, even in societies in which caregivers are observed to respond less to prelinguistic infants, like Tseltal Mayan caregivers (Elmlinger, Goldstein, & Casillas, 2023; Elmlinger, Schwade, Vollmer, & Goldstein, 2022). By babbling, infants elicit more learnable linguistic input from their caregivers.

In summary, the social feedback loop functions because of the fit between infants' perceptual and learning capacities, their immature behaviors, and parental responsiveness. As infants display more advanced skills, parents change their responsiveness, raising the bar for what behaviors they respond to and the complexity of their response. Extended immaturity in altricial species and the necessity of caregiving for infant survival create a niche in which infants can learn by engaging in social interactions with their caregivers. In socially guided vocal learning, prelinguistic vocalizations change as a function of maturing vocal abilities and contingent responses from caregivers. Learning the sounds relevant to the ambient language environment is enhanced by vocal turn-taking interactions, because these are opportunities to generate, test, and update predictions about one's own vocal abilities and their social efficacy.

The following section will review how prelinguistic vocalizations develop over the first year, and how their use during social interactions with caregivers impacts vocal development, and subsequent language development, over time.

6. Vocal learning in human infants

Maturation of articulators and the body constrain the types of vocalizations that infants can make. In infants' first two months of life, their sounds are typically limited to vegetative and distress vocalizations, which are produced involuntarily. They do, however, produce some vowel-like sounds, known as *quasivowels* (Oller, 2000). By approximately 4 months, infants begin to make more mature vowel-like sounds that are *fully resonant* and begin to integrate consonants into their repertoires, typically resulting in slow consonant-vowel alternations called *marginal syllables* (Oller, 2000). By 6 months of age, infants begin to produce *canonical syllables*. These syllables are defined by quick alternations between vocal tract closures and openings to produce consonant-vowel pairs with fully-resonant voicing, resembling adult speech sounds (Oller, 2000).

Vocal development depends on development in sensory, motor, social, and language domains (Kent, 2022). The influence of the maturing vocal tract on the sounds that infants are capable of making across the first year can be considered alongside internal and external factors driving vocal behavior. Machine learning models offer unique insight into how intrinsic motivation plays an important role in supporting the feedback loop of vocal development. By mastering immature behaviors first and then going on to attempt and master increasingly mature behaviors, vocal tract models demonstrate that phonetic learning can result from active exploration of acoustic space (Moulin-Frier, Nguyen, & Oudeyer, 2014). By babbling, infants can explore their vocal capacities while incidentally receiving social feedback from responsive caregivers.

Studying how social feedback is integrated into vocal behaviors aids in the understanding of how infants begin to produce sounds relevant to their ambient language environment once they have developed the physical capacity to do so.

7. Socially guided vocal learning

Infant vocal development is a socially-embedded, bidirectional process whereby infants produce immature sounds, receive responses with learnable linguistic regularities from adults, and integrate that information into subsequent sounds (Goldstein & Schwade, 2010). The timing of caregiver speech in response to infant vocal behavior makes IDS a salient signal during vocal learning. Furthermore, the changes that parents make to their contingent speech in the moment (Elmlinger et al., 2019) and as their infants mature (Albert et al., 2018) serve as scaffolding for increasingly advanced communicative behaviors. Infants play a role in creating opportunities for learning by eliciting interactions with their caregivers by vocalizing.

Beyond providing infants with attention-orienting and simplified linguistic input, contingent responses to babbling bolster infant expectations that their behaviors may elicit changes in their social partners (Bigelow & Power, 2022). What happens when infants can no longer elicit the vocal responses that they have come to expect from their caregivers? Still-face experiments have been used to assess infants' reactions to their parents when they momentarily cease typical responsiveness (Goldstein et al., 2009). When mothers stop and stare at their infant for a short period of time, ignoring all infant behaviors, infants display a burst in social bids by increasing their rate of vocalizing. This pattern is referred to as a *vocal extinction burst*, defined as a burst in vocalizing when caregiver responsiveness extinguishes. Vocalization rate decreases again when caregivers resume their normal levels of responsiveness.

Expectations around parental responsiveness to babbling are learned over the course of the first year. In a subsequent study comparing 3- and 5-month-old infants, 5-month-olds again showed the extinction burst shown in Goldstein et al. (2009), while 3-month-olds did not. Furthermore, the strength of the vocal extinction bursts was related to how responsive caregivers were in babbling during free play interactions (Elmlinger, Goldstein, et al., 2023; Elmlinger, Schwade, et al., 2022). These findings provide evidence that infants' expectations about the social influence of their vocalizations develop as infants become more experienced with the patterns of typical social interactions. These patterns become apparent when both caregiver and infant behavior follow each other in repeated bouts.

Babbling provides multiple cues for caregivers to attend and respond to. For example, in addition to the maturity of a given vocalization, the directedness provides clues about what infants are interested in. The salience of vocal maturity and directedness have been experimentally tested in playback paradigms in which caregivers watch and respond to naturalistic videos of infants vocalizing while playing (Albert et al., 2018). Parents selectively responded more frequently to more mature vocalizations, and to vocalizations that were object- directed. Additionally, parents used different types of responses to different types of vocalizations. For example, canonical syllables received more imitative responses than less mature marginal syllables. Object-directed vocalizations received more responses that were sensitive to what the infant was vocalizing toward, while undirected vocalizations received more affirmations. Caregivers' responses to the playback stimuli were comparable to their responses to their own infants during free play (Albert et al., 2018). Therefore, maturity and directedness of vocalizations play an influential role in caregiver responsiveness to babbling, creating more opportunities for scaffolding vocal, social, and cognitive development.

As infants learn about their ability to create learning opportunities by babbling, they begin learning new phonological forms through social interactions with their caregivers. Contingent caregiver responsiveness to infant vocalizations has been demonstrated to play a role in the emergence of syllabic structure of vocalizations (Goldstein & Schwade, 2008). In an experiment manipulating the timing and content of caregiver responses to babbling, infants either received fully-resonant vowels or monosyllabic consonant-vowel sounds from their caregivers as contingent responses to their prelinguistic vocalizations. Infants in the yoked-control group received the same amount of input, but on the schedule of the experimental group infants, so that it was not temporally linked to their babbling. As such, they heard the same amount of input.

Infants who received contingent responses integrated the phonological patterns of their caregivers' speech, while the infants in the control group did not show changes to their vocalizations (Goldstein & Schwade, 2008). Contingent responses from caregivers helped infants learn the phonological patterns of their caregivers' speech beyond mere imitation of the sounds they heard. Vocal turn-taking is thus a context in which mature vocal forms develop after having been shaped by caregiver responsiveness.

In summary, contingent social responding is the component of the feedback loop that facilitates vocal change over time because they can be used by infants to guide the production of future vocalizations. Early in infancy, contingent responsiveness helps infants learn that vocal behavior has social consequences (Goldstein et al., 2009). Contingent caregiver speech is simpler than non-contingent IDS (Elmlinger et al., 2019), making it more learnable. As infants' prelinguistic vocalizations mature, they begin to receive differential responsiveness contingently to their more mature and speech-like vocalizations (Albert et al., 2018), which provides infants with information about how their new sounds impact the social environment in comparison to less mature sounds. Across species, socially guided vocal learning is an efficient strategy for altricial vocal learners who have extended access to responsive caregivers during their periods of behavioral plasticity. Communicative development across species can thus proceed from similarly-structured social feedback loops in which immature behaviors develop as a function of timely and tailored responses from caregivers.

7.1 Vocal learning in non-human animals

Much of what is known about vocal production learning is informed by research on songbirds, who use their song for mate attraction and territory defense. Zebra finches serve as a popular model species that provides insight on the neurobiological and behavioral mechanisms underlying vocal learning. Zebra finch vocal development first entails a sensitive period of sensory learning in which the young male birds memorize the song of a tutor. This period overlaps with the later emergence of vocal plasticity, during which time juveniles produce immature sounds known as *subsong* and developmentally later *plastic song* (Liu, Gardner, & Nottebohm, 2004). At around 90 days of age zebra finches are sexually mature and song reaches a stereotyped state, known as *crystallized song*, after which point vocalizations undergo little developmental change (Immelmann, 1969).

Social interactions guide this development. Exposure to mature social partners is crucial for vocal development, as zebra finches raised in isolation do not develop typical song (Eales, 1985, 1987), and juvenile males choose their song tutor based on their prior interactions with them (Clayton, 1987). Quality of learning can be predicted by the pupil's attentiveness to their tutor (Baran, Peck, Kim, Goldstein, & Adkins-Regan, 2017; Chen et al., 2016). The modality of the tutoring experience also affects the pupil bird's crystallized song. Juvenile zebra finches who have live, multimodal interactions with their tutor learn a song that highly resembles that of their tutor, as compared to pupils who could not interact with the tutor, or who could only hear them (Chen et al., 2016; Varkevisser, Mendoza, et al., 2022).

Young male zebra finches also learn from behavior of non-singing social partners (Carouso-Peck & Goldstein, 2019). In a playback experiment, researchers manipulated the timing of visual feedback presented to juveniles. When a juvenile bird in the experimental group vocalized, he received visual feedback consisting of a video of a female fluffing up her feathers (a typical response to attractive song; Vyas, Harding, Borg, & Bogdan, 2008). Each time the bird in the contingent condition elicited a playback, the video was also played to his genetic brother in another chamber, regardless of the brother's vocal activity (yoked-control condition). At sexual maturity, birds who received contingent visual feedback outperformed their yoked counterparts, despite having vocalized at similar rates and having received the same amount of visual stimulation. The social contingency of the visual stimulus was key to creating learnable feedback that led to more mature song at sexual maturity (Carouso-Peck & Goldstein, 2019).

The findings from the playback experiment were supported by a naturalistic study that found a correlation between the amount of maternal contingent fluff-ups and song outcome measures (Carouso-Peck, Goldstein, & Fitch, 2021). This study further determined that paternal contingent behavior also facilitated song development, as juveniles whose fathers consistently sang immediately after them developed a more accurate song than juveniles with less responsive fathers (Carouso-Peck, Menyhart, DeVoogd, & Goldstein, 2020). This social learning mechanism is not unique to zebra finches, however; it is likely to be found in other species in which the sensory and sensorimotor learning periods overlap (Carouso-Peck et al., 2021). That is, species that can concurrently alter their memorized song representation as well as their vocal output can use this flexibility to modify their vocalizations based on the feedback of social partners. For example, brown-headed cowbirds (*Molothrus ater*) also use contingent social feedback to guide their song development (West & King, 1988).

The social feedback loop applies well to both human and non-human vocal learning because of parallels in developmental niches. Juvenile songbirds' early vocal plasticity is an advantage because they are surrounded by adults that can provide feedback to immature vocal productions. Adults alter their feedback to account for immaturity (e.g., Chen et al., 2016) and juveniles' preference for contingent learning sources may work to filter what to learn from (e.g., Zai et al., 2020). Much like humans, other vocal learning species with access to responsive caregivers can develop socially relevant sounds by producing immature vocalizations in social interactions and using contingent feedback to guide subsequent vocal productions.

In summary, the social feedback loop describes the adaptive links between immature vocalizing and caregiver responses. Caregivers provide acoustic scaffolding to immature vocalizations, and infants are capable of using contingent caregiver behavior to guide their future vocalizations and learn about language. Contingency works to orient infant attention to their caregivers' responses (Masek et al., 2021), and more specifically to more learnable forms of caregiver speech (Elmlinger et al., 2019).

7.2 Benefits of learning from the social feedback loop

Vocal learning from the social feedback loop not only results in infants' learning to produce mature vocal forms, but also learning that their vocal productions can elicit more learnable information from their caregivers. By babbling, infants can use their vocal behavior to organize their word learning environments. Infants' object-directed vocalizations are an important signal to caregivers about their interest and readiness to receive an object label. Contingent responses during instances when infants are vocalizing at specific objects promote word-object recognition (Goldstein et al., 2010). These interactions predict increases in vocabulary size by 15 months of age (Goldstein & Schwade, 2010).

Why does contingency play such a central role in the feedback loop? Contingency is a temporal regularity that affords building associations between immature vocalizing and parental behavior. Attention has been proposed as a mechanism by which contingency facilitates infants' developing temporal, semantic, and pragmatic understanding of language. We posit that attention to contingency is an outcome of a broader mechanism, *curiosity*, that drives the feedback loop.

Contingent social interactions are opportunities in which infants can form predictions about the social consequences of their vocalizing. An intrinsic motivation to increase prediction accuracy in social interactions may drive infants' attention to contingent responses from their caregivers. Thus, we turn to a framework of prediction formation and testing, *curiosity-driven learning*, as a cognitive mechanism of motivation and learning within the social feedback loop.

8. Curiosity as a mechanism for vocal learning

The social feedback loop works because caregivers respond contingently to immature behaviors in a way that scaffolds vocal maturity. Specifically, as infants begin to produce more mature vocalizations, caregivers differentially respond to sounds that display more vocal maturity (Albert et al., 2018). Such differential responding is data that informs infants' predictions about the efficacy of their vocalizations. Predictions that infants have formed around their caregivers' responsiveness to their vocal behaviors are updated when they perceive the differential contingency of their parents' responses to their new behaviors. Socially guided vocal development is thus made possible by infants' ability to refine their predictions about the social consequences of vocalizing.

Across domains, infants are able to make use of their experiences to generate predictions about incoming input from their environments. As infants gain more experience, their predictions increase in accuracy. In this way, infants behave as "scientists in the crib" (Gopnik, Meltzoff, & Kuhl, 1999), experimenting and comparing their predictions about the world to new experiences. According to Bayesian theories of learning, infants organize their learning using priors that are based on their past experiences (Gopnik & Bonawitz, 2015). Learning from active exploration is guided by knowledge from previous experiences. It has been demonstrated within a Bayesian framework that infants generate priors about event probabilities based on past observations and integrate new information to update their priors for subsequent events (Téglás et al., 2011). In other words, infants track the probabilistic likelihood of observing certain features in their environment. Infants' sensitivity to unexpected outcomes relies on predictions based on their past experiences. For example, when trained on "impossible" stimuli that violate natural physics, such as a solid object passing through another solid object, infants will show greater surprisal when they are then shown a stimulus that violates that pattern (a solid object colliding when in contact with another object; Cashon & Cohen, 2000). Infants' understanding of the rules that govern their physical and social worlds can be built through repeated experiences with the learnable statistical regularities, experiences which they actively seek out over others. When infants guide their exploration of the ambient environment by attending to features that have such learnable regularities, they create opportunities for learning.

9. Selecting what to learn

Predictability plays a role in curiosity-driven learning (CDL) because it provides a signal for learnability that guides attention to or from stimuli depending on a learner's goals, prior knowledge, and prediction accuracy (Kidd et al., 2012). Sensitivity to intermediately-surprising events generates attention to information that can then be learned from. Sensitivity to stimuli that are intermediately predictable allows infants to home in on the features of the world that offer the greatest opportunities for learning (i.e., increasing prediction accuracy). CDL is predicated on the assumptions that individuals are intrinsically motivated to forage for information that can be used to increase prediction accuracy, and that increasing prediction accuracy is rewarding. CDL should thus drive individuals to attend to highly informative stimuli (Oudeyer & Smith, 2016).

When presented with stimuli with varying levels of predictability, infants' visual attention was maintained the most by intermediately predictable stimuli. In contrast, overly predictable or very surprising stimuli were more likely to lose infants' visual attention (Kidd et al., 2012). This demonstrates that infants actively track statistics about event probabilities and preferentially attend to events with learnable statistics.

9.1 CDL in non-human animals

We should expect to see that non-human animals will also work to obtain information and will prefer information that is appropriate to their learning goals. Indeed, non-human animals will consistently choose to receive information about upcoming events, even though they are unable to change the outcome, or if doing so results in a sacrifice (Blanchard, Hayden, & Bromberg-Martin, 2015; Bromberg-Martin, and Hikosaka, 2009; Wang & Hayden, 2019; Wyckoff, 1952). As in humans, the motivation to gain new information drives macaque attention toward intermediately predictable events. Adult rhesus macaques display the same pattern of visual attention as a function of surprisal as the 7–8 month old infants in Kidd et al. (2012), Wu et al. (2022). Predictability is thus a driving force for shaping how both humans and non-human primates attend to statistically relevant cues. For altricial animals, and social animals more broadly, social partners are often the richest source of contingent and relevant information to benefit learning across domains. This pattern of attention allocation may allow learners to focus on stimuli that maximizes learning progress (Wu et al., 2022). Importantly, predictability is a moving target as learners improve prediction accuracy and begin to seek more complexity over time.

9.2 Motivation to increase prediction accuracy

Encountering novel information with both familiar and relatively unexpected features attracts infants' attention because infants experience a state of uncertainty about that information. In other words, they become curious. Curiosity guides how infants explore the environment and allows them to integrate new information to validate or update their currently held beliefs.

Infants and children prioritize information that they are uncertain about (Pelz & Kidd, 2020). Curiosity-driven learners are intrinsically motivated to increase the accuracy of their predictions and reduce uncertainty (Oudeyer & Smith, 2016). In this framework, individuals should preferentially attend to information that is neither unpredictable nor too predictable (Kidd et al., 2012). By attending to stimuli with intermediate predictability, they have the potential to increase prediction accuracy. What is predictable and unpredictable is determined by experience. Thus, once prediction accuracy stops improving for a given stimulus, attention to that stimulus decreases (Fig. 3).

10. CDL as a framework for socially guided vocal learning

While not typically studied in the social domain, we propose that infants' curiosity also motivates them to seek out predictable yet dynamic sources of information, which in most cases are their caregivers or other social partners. Furthermore, we propose that socially guided vocal learning

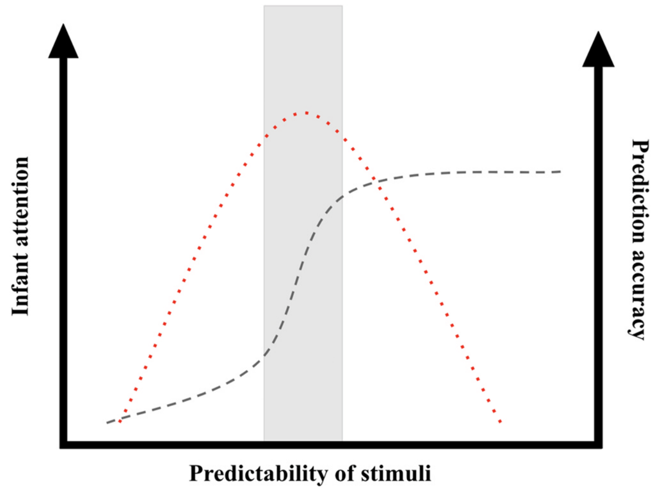

Fig. 3 Attention to stimuli depends on its predictability and infants' improving prediction accuracy. Kidd et al. (2012) have demonstrated that infants allocate their visual attention depending on the predictability of stimuli. When stimuli have low predictability, the ability to improve prediction accuracy (gray dashed curve) is low, and infant attention (red dotted line) is also low. When stimuli are highly predictable, prediction accuracy stops improving, and infant attention is low. When stimuli are intermediately predictable, there is potential to improve prediction accuracy, and infants pay more attention as a result.

is driven by CDL mechanisms. In a CDL framework of vocal learning, the social feedback loop is effective because it allows infants make predictions about their own behaviors and those of others, test predictions by producing immature behaviors, and use contingent responses to update their prior predictions with new social information from caregivers (Fig. 4).

By 5 months, infants have gained experience with caregivers who typically respond to their babbling (Elmlinger, Goldstein, et al., 2023; Elmlinger, Schwade, et al., 2022). This regularity contributes to infants' predictions about the likelihood that a caregiver will contingently respond to their vocalizations. After exposure to social contingency, violations of infants' expectations will prompt them to work to make their worlds align with their predictions. This is exemplified by the vocal extinction burst produced during still-face paradigms (Elmlinger, Goldstein, et al., 2023; Elmlinger, Schwade, et al., 2022). Over the course of their first year, infants begin to produce vocalizations that vary in vocal quality and maturity, and their caregivers begin to respond in different ways to infants' different types of vocal forms. Thus, contingent caregiver responsiveness to infants'

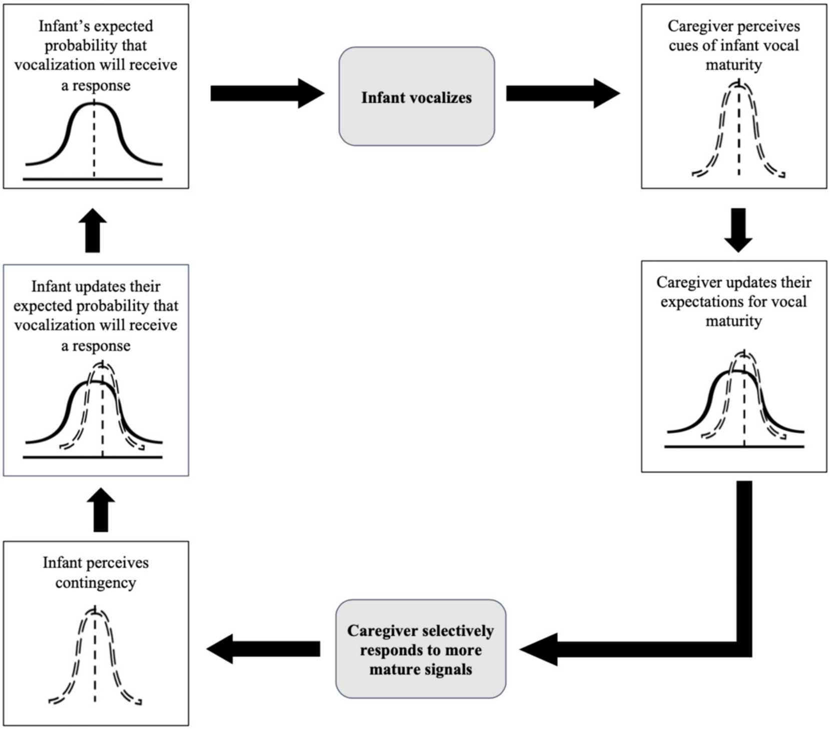

Fig. 4 Curiosity-driven learning as a mechanism for socially guided vocal learning. Based on prior experience with vocal turn-taking interactions with their caregivers, infants have expectations about the probability that their vocalizations will receive a timely response. The solid black curve on the left represents this expectation as distribution of probabilities of receiving a response to their next vocalization from their caregiver. They can vocalize to test their predictions about how caregivers will respond. Caregivers' previous interactions with their infants inform their expectations about the maturity of their infants' behaviors. New input from their infant (displayed by the dashed distribution on the right) leads them to update their expectations (solid distribution on the right) for what their infant is capable of. Selective parental contingency (displayed by the dashed distribution on the left) can be perceived by infants and is used to update infants' expectations to align with the information that their changing behavior elicits. This updated expectation guides their future behavior. Infants are motivated to engage in these interactions because they involve opportunities to increase prediction accuracy about caregiver responsiveness.

vocalizations can be used by infants to create and test predictions about the social efficacy of vocal forms that differ in maturity. Early experiences with contingent behaviors from caregivers shape infants' predictions about their own behaviors and the behaviors of social partners in vocal turn-taking interactions.

The social feedback loop is supported by a mechanistic feedback loop in which (1) infants' prediction accuracy about the effects of their vocalizing is continuously updated by social behavior, and (2) caregivers' predictions about the maturity of their infants are updated by changes in immature vocal behavior (Fig. 4). These predictions can be represented as distributions of probabilities. Over time, these distributions would hypothetically become narrower as they center around the most-encountered probability. Infant's beliefs about when a caregiver will respond to their next vocalization and are based on typical levels of caregiver contingent responsiveness experienced in past interactions. When a caregiver responds, an infant perceives the contingency of this response and compares it to their initial prediction for a contingent response. If their caregivers' response aligns with the predictions, this is a signal of prediction accuracy. Infants can make the social world more predictable by vocalizing because their caregivers respond to vocalizations with statistically regular, contingent patterns of behavior. When caregivers contingently respond to infants' immature vocalizations in ways that align with infants' predictions about caregiver responsiveness, this response is rewarding because it signals improving prediction accuracy (Gottlieb, Hayhoe, Hikosaka, & Rangel, 2014; Gottlieb, Oudeyer, Lopes, & Baranes, 2013).

How does the motivation to improve prediction accuracy facilitate the development of more mature vocal forms? Early in the first year, infants' immature vocalizations elicit highly predictable patterns of responsiveness from their caregivers (Elmlinger, Goldstein, et al., 2023; Elmlinger, Schwade, et al., 2022). This high level of predictability, however, leaves little room for improvement in prediction accuracy, so infants become less rewarded by continuing to produce those immature sounds. As infants become better able to produce more mature vocalizations, such as speech-like, canonical syllables, parents respond in less predictable (and thus more rewarding) ways. Producing more mature vocalizations is prioritized by infants because doing so creates opportunities to improve prediction accuracy around increasing caregiver responsiveness to new sound forms in their developing vocal repertoire. Caregivers scaffold their infants' vocal development by modifying their responsiveness as they perceive their infants' increasing maturity. Such scaffolding works together with infants' developing motor and social skills as well as infants' intrinsic motivation to increase prediction accuracy.

Caregivers' predictions for their infants' vocal maturity play an important role in this feedback loop. Contingent responsiveness differs as a function of infants' developing vocalizations. When infants become capable of producing sounds that more closely resemble the ambient language, such as consonant-vowel clusters and speech-like canonical syllables, caregivers change their responsiveness. They become more likely to respond to their infants' vocalizations, and they increase their use of verbal acknowledgments (Albert et al., 2018; Goldstein et al., 2003; Gros-Louis, West, Goldstein, & King, 2006). Caregivers' predictions are changed by interactions with their infants when infants begin to use new vocal forms. As such, caregivers will come away from those interactions with higher expectations for what their infants are capable of in future interactions. Caregivers' contingent responses to new vocal forms are informative to their infants. Infants will produce those more mature vocalizations to improve their own prediction accuracy around caregiver responsiveness to them. Within a CDL framework, infants will prioritize the production of vocalizations that create opportunities to increase prediction accuracy by eliciting differential contingent responsiveness from caregivers. Over time, this facilitates vocal learning because increasingly mature vocal forms are used and practiced.

In our view, vocal development is likely determined both by curiosity-driven exploration of one's own vocal abilities (Moulin-Frier et al., 2014) and the reinforcement by differentially contingent caregiver responsiveness that depends on vocal maturity. Babbling in interactions with caregivers is an opportunity to put predictions to the test about the social efficacy of vocal behaviors (Fig. 4). As caregivers change patterns of responsiveness to the increasing maturity of their infants' vocalizations (Albert et al., 2018), infants produce more advanced forms in an attempt to probe the more uncertain spaces their new abilities afford them. In summary, infants' curiosity results from uncertainty about caregiver responsiveness to infants' own developing vocal behaviors. This curiosity drives them to engage in social interactions with caregivers whose responsiveness can be used as a signal of prediction accuracy.

10.1 Selecting who to learn from

Dynamic caregiver responsiveness, specifically responsiveness that depends on the infants' vocal maturity, is engaging to infants, as it provides opportunities to learn what types of prelinguistic vocalizations are the most socially effective. Importantly, in a CDL framework, this varied and intermediate level of social contingency works better for learning than perfect contingency to infant vocal behavior. For example, if an infant receives a response every time they vocalize, then vocal turn-taking interactions would not be opportunities to improve the accuracy of their predictions about the likelihood of contingent responsiveness. Intermediate contingency of engaged, yet selectively responsive caregivers is sufficient to orient infant attention to informative communicative input. In summary, social contingency that occurs in a dynamic fashion that depends on infants' changing behavior should be engaging to infants.

While a dynamic level of caregiver contingency works well for infant learning, infants with caregivers who respond in extremely unpredictable ways display impairments in their ability to learn from their caregivers. For example, mothers with postpartum depression have been found to display longer and more variable lengths of pauses between vocal turns with their infants, which may impact the synchrony between infants and depressed mothers in vocal turn- taking interactions (Zlochower & Cohn, 1996). Unpredictable contingency may impair infants' ability to learn from their caregivers, since such a low level of predictability would impact their ability to increase prediction accuracy about contingent responsiveness by vocalizing. In a social CDL framework, levels of caregiver contingency must change based on infants' behavior to create social feedback that is an optimal signal of learnability for infants.

In addition to unpredictability of social partners or environments, disorders associated with social deficits, as in autism spectrum disorder (ASD), may change how children interact and learn from their environments (Ellis Weismer & Saffran, 2022). ASD may be associated with difficulties with making predictions, resulting in hyperplasticity and impaired learning (Sinha et al., 2014). Impairments in properly weighting prediction error when stimuli do not meet infants' prior expectations can lead to social and learning deficits (Van de Cruys et al., 2014).

Impairments in forming and updating predictions may explain differential developmental outcomes in this population in domains where learning relies on prediction. ASD may present as a deficit in social understanding because the domain-general ability to make predictions about imperfectly contingent behavior is crucial to social learning.

11. Future directions

By operationalizing predictability in social interactions, researchers can test how the contingency of social partners impacts infant attention and vocal learning. Kuhl (2007) posited that speech learning is "gated" by social interactions, and that contingency may play an important role in this. It is possible that predictability, in the form of contingent caregiver responsiveness to infant vocal behavior, creates the gateway that allows social interactions to facilitate speech and language learning.

11.1 History matters

Challenges in manipulating predictability of social interactions in a lab setting include the fact that predictability is dependent on prior experience. In other domains, subjects' priors can be built at the beginning of an experiment by presenting participants with novel stimuli and then testing the impact of the stimuli's predictability on behavior (e.g. Kidd et al., 2012). In the social domain, researchers cannot discount the role of infants' prior social experience with caregivers, who exhibit individual differences in contingency. Infants' experience may set a baseline for their expectations about the predictability of new social partners and impact the level of contingency infants can learn from or find interesting. A better starting point may be to assess infant vocal learning from non-human contingent agents, for which they would have no priors (Goldstein & Schwade, 2016). Assessing how the predictability of contingent responses interacts with infant attention to the agent and how infants learn from them would inform subsequent research on how predictability impacts learning from biological social partners.

11.2 Comparative models

Investigating these concepts in laboratory studies with animal models offer many advantages, as developmental social experiences can be much more tightly controlled by the researcher. Zebra finches, whose song development process is commonly studied as an analog of human speech development (Carouso-Peck et al., 2020; Doupe & Kuhl, 1999), learn their songs after only a few months of training. Automated recording software can record every vocalization made by the birds during this time, enabling researchers to examine an individual bird's vocal trajectory in fine detail over various timescales (Tchernichovski et al., 2000).

Juveniles can learn through operant control paradigms (Adret, 1993; Derégnaucourt, Poirier, Kant, Linden, & Gahr, 2013; Tchernichovski et al., 2001) as well as from multimodal and unimodal access to tutors (Chen et al., 2016; Varkevisser, Mendoza, et al., 2022) and audiovisual playback of tutors (Varkevisser, Simon, et al., 2022) and non-singing females (Carouso-Peck & Goldstein, 2019). Furthermore, several automated processes for delivering feedback contingent on a bird's behavior have been developed (Araguas et al., 2022; Sober & Brainard, 2009; Benichov et al., 2016).

These features and resources make the zebra finch a tractable model for studying the role of the social feedback loop in the acquisition of complex communicative skills. Through manipulation of the predictability of adult zebra finch response, juvenile preference for reliable sources of social information could be investigated. Juvenile males best learn the songs that they prefer to listen to (Rodríguez-Saltos et al., 2023). It is presently unknown whether this preference arises from the predictability of past social interactions with one tutor as compared to the other. Quantifying and manipulating the behavior of both agents within the social feedback loop would help determine how the microstructure of social interactions over a small scale (e.g. seconds, hours, days) accumulate to facilitate learning over a larger scale (e.g. development).

11.3 Implications for caregivers

The role of the caregiver is critical to the social feedback loop. Within this loop, caregivers are not static entities but are actively learning and adapting to the behaviors of their infants. We know that caregivers scaffold increasingly complex skills for their infants, but much less is known about how they change their behavior to accommodate infants' developing skills. This is crucial because a key reason why contingent responses drive socially guided vocal learning is because they change in the moment and over time as a function of infants' maturing vocal behaviors. Responses become feedback for infants only when they are aligned with infants' moment-to-moment abilities and behaviors. Caregivers' changing contingency creates adaptive unpredictability that serves as a signal for learnability. Intermediate unpredictability of social behavior motivates infants to produce increasingly advanced vocal forms (as they become able to do so) to make and test predictions about their new vocal behaviors' social efficacy. Furthermore, this signal of learnability changes from the infants' point of view as infants acquire increasingly advanced means of communication. Future research should investigate real-time changes in caregiver contingency as a function of their infants' vocal behavior, as this is a major feature of parental sensitivity.

Responsiveness alone does not predict how well infants learn from caregivers. Rather, simplification of adult behaviors that are delivered contingently on infants' behaviors is what scaffolds learning and connects parental sensitivity and responsiveness to later social and cognitive development (Rodrigues et al., 2021). The prediction-based CDL feedback loop is robust to specific levels of contingency and provides a viable explanatory framework for understanding how early language learning is possible across the diverse caregiving styles that have been observed across cultures (Casillas, Brown, & Levinson, 2020; Elmlinger, Goldstein, et al., 2023; Elmlinger, Schwade, et al., 2022). It may also guide future interventions for at-risk families to improve caregiver-infant vocal interactions, for example, in families of low socioeconomic status, who have been observed to use less IDS (Dailey & Bergelson, 2022; Hart & Risley, 1995). The CDL feedback loop may also be useful for understanding the development of communication in infants with an elevated likelihood for ASD, as their caregivers have been observed to be less likely to modify their own vocal responses to babbling as a function of their infants' vocal maturity (Warlaumont et al., 2014). By better understanding the impact of early adversity on both infants' perception of contingency and the predictability parents' feedback, interventions can be designed that help parents repair gaps in the social feedback loop.

12. Conclusion

Infant vocal learning via integration of social feedback results not only in the acquisition of mature vocal forms, but also the understanding of how to learn from social feedback (Elmlinger, Goldstein, et al., 2023; Elmlinger, Schwade, et al., 2022). This is arguably important for language development, as infants enter their second year having gained an understanding that their behaviors influence their social environment. Socially guided vocal learning occurs against the backdrop of developing motor, cognitive and social skills, all of which influence the developmental trajectory of communication. Nonetheless, social interactions in which infants elicit social feedback via their immature behaviors are key moments for learning, and result in the production of more mature forms over time (Goldstein & Schwade, 2010). Furthermore, infant vocal development may be driven by intrinsic motivation to increase prediction accuracy, and curiosity-driven learning may be a foundational mechanism of the feedback loop in which socially guided vocal learning occurs.

Altricial infants can learn from a social feedback loop because they remain in the care of adults who are able to scaffold their learning with developmentally tailored input. A key feature of the altricial vocal learner's developmental niche is extended reliance on adults who are sensitive to features of infants' behaviors that signal maturity. The alignment between infants' developing perceptual, cognitive, and communicative skills and the feedback they receive from caregivers allows infants to learn from social interactions with their caregivers.

Finally, the social feedback loop is dynamic. Infants learn about the social world by acting on it. In doing so they can test their predictions, which are based on previous experience, against new information they encounter. Curiosity-driven learning allows for more efficient and expansive exploration of patterns in social responses. Their maturing ability to perceive statistical regularities in their social environments allows them to learn patterns and relationships crucial to their continued learning. In social interactions, their explorations are guided by both their own curiosity and the responses they receive from caregivers. However, many aspects of CDL as a mechanism of learning in social contexts remain unknown. Future research should focus on better understanding the interactions between infants and their caregivers during this co-developed and dynamic feedback loop.

References

Adret, P. (1993). Operant conditioning, song learning and imprinting to taped song in the zebra finch. *Animal Behaviour, 46*, 149–159.

Albert, R. R., Schwade, J. A., & Goldstein, M. H. (2018). The social functions of babbling: Acoustic and contextual characteristics that facilitate maternal responsiveness. *Developmental Science, 21*(5), e12641. https://doi.org/10.1111/desc.12641

Alberts, J. R. (2008). The nature of nurturant niches in ontogeny. *Philosophical Psychology, 21*(3), 295–303. https://doi.org/10.1080/09515080802169814

Araguas, A., Guellaï, B., Gauthier, P., Richer, F., Montone, G., Chopin, A., & Derégnaucourt, S. (2022). Design of a robotic zebra finch for experimental studies on developmental song learning. *Journal of Experimental Biology, 225*(3), jeb242949.

Baran, N. M., Peck, S. C., Kim, T. H., Goldstein, M. H., & Adkins-Regan, E. (2017). Early life manipulations of vasopressin-family peptides alter vocal learning. *Proceedings of the Royal Society B: Biological Sciences, 284*, 20171114.

Baumwell, L., Tamis-LeMonda, C. S., & Bornstein, M. H. (1997). Maternal verbal sensitivity and child language comprehension. *Infant behavior and Development, 20*(2), 247–258.

Benichov, J. I., Benezra, S. E., Vallentin, D., Globerson, E., Long, M. A., & Tchernichovski, O. (2016). The forebrain song system mediates predictive call timing in female and male zebra finches. *Current Biology, 26*(3), 309–318.

Bigelow, A. E., & Power, M. (2022). Influences of infants' and mothers' contingent vocal responsiveness on young infants' vocal social bids in the Still Face Task. *Infant Behavior and Development, 69*, 101776. https://doi.org/10.1016/j.infbeh.2022.101776

Blanchard, T. C., Hayden, B. Y., & Bromberg-Martin, E. S. (2015). Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron, 85*(3), 602–614. https://doi.org/10.1016/j.neuron.2014.12.050

Bromberg-Martin, E. S., & Hikosaka, O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron, 63*(1), 119–126. https://doi.org/10.1016/j.neuron.2009.06.009

Carouso-Peck, S., & Goldstein, M. H. (2019). Female social feedback reveals non-imitative mechanisms of vocal learning in zebra finches. *Current Biology, 29*(4), 631–636.e3. https://doi.org/10.1016/j.cub.2018.12.026

Carouso-Peck, S., Goldstein, M. H., & Fitch, W. T. (2021). The many functions of vocal learning. *Philosophical Transactions of the Royal Society B: Biological Sciences, 376*(1836), 20200235. https://doi.org/10.1098/rstb.2020.0235

Carouso-Peck, S., Menyhart, O., DeVoogd, T. J., & Goldstein, M. H. (2020). Contingent parental responses are naturally associated with zebra finch song learning. *Animal Behaviour, 165*, 123–132. https://doi.org/10.1016/j.anbehav.2020.04.019

Cashon, C. H., & Cohen, L. B. (2000). Eight-month-old infants' perception of possible and impossible events. *Infancy: The Official Journal of the International Society on Infant Studies, 1*(4), 429–446.

Casillas, M., Brown, P., & Levinson, S. C. (2020). Early language experience in a Tseltal Mayan village. *Child Development, 91*(5), 1819–1835.

Chen, Y., Matheson, L. E., & Sakata, J. T. (2016). Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proceedings of the National Academy of Sciences of the United States of America, 113*(24), 6641–6646. https://doi.org/10.1073/pnas.1522306113

Clayton, N. S. (1987). Song tutor choice in zebra finches. *Animal Behaviour, 35*(3), 714–721. https://doi.org/10.1016/S0003-3472(87)80107-0

Crowe, H. P., & Zeskind, P. S. (1992). Psychophysiological and perceptual responses to infant cries varying in pitch: Comparison of adults with low and high scores on the child abuse potential inventory. *Child Abuse & Neglect, 16*(1), 19–29. https://doi.org/10.1016/0145-2134(92)90005-C

Dailey, S., & Bergelson, E. (2022). Language input to infants of different socioeconomic statuses: A quantitative meta-analysis. *Developmental Science, 25*, e13192. https://doi.org/10.1111/desc.13192

Deák, G. O., Krasno, A. M., Triesch, J., Lewis, J., & Sepeta, L. (2014). Watch the hands: Infants can learn to follow gaze by seeing adults manipulate objects. *Developmental science, 17*(2), 270–281.

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science (New York, N. Y.), 208*(4448), 1174–1176. https://doi.org/10.1126/science.7375928

DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior & Development, 9*(2), 133–150.

Derégnaucourt, S., Poirier, C., Kant, A. V., Linden, A. V., & Gahr, M. (2013). Comparisons of different methods to train a young zebra finch (*Taeniopygia guttata*) to learn a song. *Journal of Physiology-Paris, 107*(3), 210–218. https://doi.org/10.1016/j.jphysparis.2012.08.003

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience, 22*(1), 567–631.

Eales, L. A. (1985). Song learning in zebra finches: Some effects of song model availability on what is learnt and when. *Animal Behaviour, 33*(4), 1293–1300. https://doi.org/10.1016/S0003-3472(85)80189-5

Eales, L. A. (1987). Song learning in female-raised zebra finches: Another look at the sensitive phase. *Animal Behaviour, 35*(5), 1356–1365. https://doi.org/10.1016/S0003-3472(87)80008-8

Eales, L. A. (1989). The influences of visual and vocal interaction on song learning in zebra finches. *Animal Behaviour, 37*(3), 507–508. https://doi.org/10.1016/0003-3472(89)90097-3

Ellis Weismer, S., & Saffran, J. R. (2022). Differences in prediction may underlie language disorder in autism. *Frontiers in Psychology, 13*, 3212.
Elmlinger, S. L., Goldstein, M. H., & Casillas, M. (2023). Immature vocalizations simplify the speech of Tseltal Mayan and U.S. Caregivers. *Topics in Cognitive Science, 15*(2), 315–328. https://doi.org/10.1111/tops.12632
Elmlinger, S. L., Schwade, J. A., & Goldstein, M. H. (2019). The ecology of prelinguistic vocal learning: Parents simplify the structure of their speech in response to babbling. *Journal of Child Language, 46*(5), 998–1011. https://doi.org/10.1017/S0305000919000291
Elmlinger, S. L., Schwade, J. A., Vollmer, L., & Goldstein, M. H. (2022). Learning how to learn from social feedback: The origins of early vocal development. *Developmental Science*e13296. https://doi.org/10.1111/desc.13296
Faust, K. M., Carouso-Peck, S., Elson, M. R., & Goldstein, M. H. (2020). The origins of social knowledge in altricial species. *Annual Review of Developmental Psychology, 2*(1), 225–246. https://doi.org/10.1146/annurev-devpsych-051820-121446
Fernandez, A. A., & Knörnschild, M. (2020). Pup directed vocalizations of adult females and males in a vocal learning bat. *Frontiers in Ecology and Evolution, 8*, 265. https://doi.org/10.3389/fevo.2020.00265
Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology, 39*(1), 1–42.
Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences of the United States of America, 100*(13), 8030–8035. https://doi.org/10.1073/pnas.1332441100
Goldstein, M. H., Schwade, J., Briesch, J., & Syal, S. (2010). Learning while babbling: Prelinguistic object-directed vocalizations indicate a readiness to learn. *Infancy: the Official Journal of the International Society on Infant Studies, 15*(4), 362–391. https://doi.org/10.1111/j.1532-7078.2009.00020.x
Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science, 19*(5), 515–523. https://doi.org/10.1111/j.1467-9280.2008.02117.x
Goldstein, M. H., & Schwade, J. A. (2010). *From birds to words: Perception of structure in social interactions guides vocal development and language learning. Oxford handbook of developmental behavioral neuroscience.* Oxford University Press 708–729.
Goldstein, M. H., & Schwade, J. A. (2016). *The role of contingency in vocal learning from biological and non-biological interaction partners [Poster presentation]. International congress of infant development*. USA: New Orleans. ⟨https://infantstudies.org/wp-content/uploads/2016/06/2016-ICIS-Program-Book-1.pdf⟩.
Goldstein, M. H., Schwade, J. A., & Bornstein, M. H. (2009). The value of vocalizing: Five-month-old infants associate their own noncry vocalizations with responses from caregivers. *Child Development, 80*(3), 636–644. https://doi.org/10.1111/j.1467-8624.2009.01287.x
Goldstein, M. H., & West, M. J. (1999). Consistent responses of human mothers to prelinguistic infants: the effect of prelinguistic repertoire size. *Journal of Comparative Psychology, 113*(1) 52.
Golinkoff, R. M., Can, D. D., Soderstrom, M., & Hirsh-Pasek, K. (2015). (Baby)Talk to me: The social context of infant-directed speech and its effects on early language acquisition. *Current Directions in Psychological Science, 24*(5), 339–344. https://doi.org/10.1177/0963721415595345
Gopnik, A., & Bonawitz, E. (2015). Bayesian models of child development. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*(2), 75–86.
Gopnik, A., Meltzoff, A. N., & Kuhl, P. K. (1999). *The scientist in the crib: Minds, brains, and how children learn.* William Morrow & Co.

Gottlieb, J., Hayhoe, M., Hikosaka, O., & Rangel, A. (2014). Attention, reward, and information seeking. *The Journal of Neuroscience, 34*(46), 15497–15504.
Gottlieb, J., Oudeyer, P. Y., Lopes, M., & Baranes, A. (2013). Information-seeking, curiosity, and attention: computational and neural mechanisms. *Trends in Cognitive Sciences, 17*(11), 585–593.
Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development, 30*(6), 509–516.
Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Paul H Brookes Publishing.
Heyes, C. (2016). Born pupils? Natural pedagogy and cultural pedagogy. *Perspectives on Psychological Science, 11*(2), 280–295. https://doi.org/10.1177/1745691615621276
Hilton, C. B., Moser, C. J., Bertolo, M., Lee-Rubin, H., Amir, D., Bainbridge, C. M., & Mehr, S. A. (2022). Acoustic regularities in infant-directed speech and song across cultures. *Nature Human Behaviour, 6*(11), 1545–1556.
Hym, C., Dumuids, M. V., Anderson, D. I., Forma, V., Provasi, J., Brière-Dollat, C., ... Barbu-Roth, M. (2023). Newborns modulate their crawling in response to their native language but not another language. *Developmental Science, 26*(1), e13248.
Immelmann, K. (1969). Song development in the zebra finch and other estrildid finches. *Bird vocalizations,* 61–77.
James, W. (1890). *The principles of psychology, 2.* Henry Holt and Company.
Kent, R. D. (2022). The maturational gradient of infant vocalizations: Developmental stages and functional modules. *Infant Behavior and Development, 66.* https://doi.org/10.1016/j.infbeh.2021.101682 101682.
Kenward, B. (2010). 10-Month-olds visually anticipate an outcome contingent on their own action. *Infancy: The Official Journal of the International Society on Infant Studies, 15,* 337–361. https://doi.org/10.1111/j.1532–7078.2009.00018.x
Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS One,* 7(5), e36399. https://doi.org/10.1371/journal.pone.0036399
Kuchirko, Y., Tafuro, L., & Tamis LeMonda, C. S. (2018). Becoming a communicative partner: Infant contingent responsiveness to maternal language and gestures. *Infancy: The Official Journal of the International Society on Infant Studies, 23*(4), 558–576. https://doi.org/10.1111/infa.12222
Kuhl, P. K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science, 10*(1), 110–120. https://doi.org/10.1111/j.1467-7687.2007.00572.x
Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, 100*(15), 9096–9101. https://doi.org/10.1073/pnas.1532872100
Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science (New York, N.Y.), 255*(5044), 606–608. https://doi.org/10.1126/science.1736364
Lam, C., & Kitamura, C. (2012). Mommy, speak clearly: Induced hearing loss shapes vowel hyperarticulation. *Developmental Science, 15*(2), 212–221. https://doi.org/10.1111/j.1467-7687.2011.01118.x
Liu, W., Gardner, T. J., & Nottebohm, F. (2004). Juvenile zebra finches can use multiple strategies to learn the same song. *Proceedings of the National Academy of Sciences, 101*(52), 18177–18182. https://doi.org/10.1073/pnas.0408065101
Masek, L. R., McMillan, B. T. M., Paterson, S. J., Tamis-LeMonda, C. S., Golinkoff, R. M., & Hirsh- Pasek, K. (2021). Where language meets attention: How contingent interactions promote learning. *Developmental Review, 60*, 100961. https://doi.org/10.1016/j.dr.2021.100961

Millar, W. S., & Watson, J. S. (1979). The effect of delayed feedback on infant learning reexamined. *Child Development, 50*, 747–751.

Mondloch, C. J., Segalowitz, S. J., Lewis, T. L., Dywan, J., Le Grand, R., & Maurer, D. (2013). The effect of early visual deprivation on the development of face detection. *Developmental Science, 16*, 728–742. https://doi.org/10.1111/desc.12065

Moulin-Frier, C., Nguyen, S. M., & Oudeyer, P. Y. (2014). Self-organization of early vocal development in infants and machines: The role of intrinsic motivation. *Frontiers in Psychology, 4*, 1006. https://doi.org/10.3389/fpsyg.2013.01006

Oller, D. K. (2000). *Infraphonology: Overview and central results The Emergence of the Speech Capacity* (1st ed.). Psychology Press60–74 Chapter 4.

Oller, D. K., Eilers, R. E., & Basinger, D. (2001). Intuitive identification of infant vocal sounds by parents. *Developmental Science, 4*(1), 49–60. https://doi.org/10.1111/1467-7687.00148

Oudeyer, P.-Y., & Smith, L. B. (2016). How evolution may work through curiosity-driven developmental process. *Topics in Cognitive Science, 8*, 492–502.

Pelz, M., & Kidd, C. (2020). The elaboration of exploratory play. *Philosophical Transactions of the Royal Society B, 375*(1803) 20190503.

Reznick, J. S., Morrow, J. D., Goldman, B. D., & Snyder, J. (2004). The onset of working memory in infants. *Infancy: The Official Journal of the International Society on Infant Studies, 6*, 145–154. https://doi.org/10.1207/s15327078in0601_7

Rodrigues, M., Sokolovic, N., Madigan, S., Luo, Y., Silva, V., Misra, S., & Jenkins, J. (2021). Paternal sensitivity and children's cognitive and socioemotional outcomes: A meta-analytic review. *Child Development, 92*(2), 554–577.

Rodríguez-Saltos, C. A., Bhise, A., Karur, P., Khan, R. N., Lee, S., Ramsay, G., & Maney, D. L. (2023). Song preferences predict the quality of vocal learning in zebra finches. *Scientific Reports, 13*(1), 1–10.

Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2014). Skype Me! Socially contingent interactions help toddlers learn language. *Child Development, 85*(3), 956–970. https://doi.org/10.1111/cdev.12166

Ross-Sheehy, S., & Newman, R. S. (2015). Infant auditory short-term memory for non-linguistic sounds. *Journal of Experimental Child Psychology, 132*, 51–64.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science (New York, N. Y.), 274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Schroer, S. E., & Yu, C. (2022). Looking is not enough: Multimodal attention supports the real-time learning of new words. *Developmental Science, 26*(2), e13290.

Simion, F., Valenza, E., Macchi Cassia, V., Turati, C., & Umiltà, C. (2002). Newborns' preference for up–down asymmetrical configurations. *Developmental Science, 5*(4), 427–434. https://doi.org/10.1111/1467-7687.00237

Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., ... Held, R. M. (2014). Autism as a disorder of prediction. *Proceedings of the National Academy of Sciences, 111*(42), 15220–15225. https://doi.org/10.1073/pnas.1416797111

Skinner, B. F. (1948). Superstition' in the pigeon. *Journal of Experimental Psychology, 38*(2), 168.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition, 106*(3), 1558–1568.

Smith, N. A., & Trainor, L. J. (2008). Infant-directed speech is modulated by infant feedback. *Infancy: The Official Journal of the International Society on Infant Studies, 13*(4), 410–420. https://doi.org/10.1080/15250000802188719

Sober, S. J., & Brainard, M. S. (2009). Adult birdsong is actively maintained by error correction. *Nature neuroscience, 12*(7), 927–931.

Striano, T., Henning, A., & Stahl, D. (2006). Sensitivity to interpersonal timing at 3 and 6 months of age. *Interaction Studies, 7*(2), 251–271.

Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why is infant language learning facilitated by parental responsiveness? *Current Directions in Psychological Science, 23*(2), 121–126. https://doi.org/10.1177/0963721414522813

Tchernichovski, O., Mitra, P. P., Lints, T., & Nottebohm, F. (2001). Dynamics of the vocal imitation process: How a zebra finch learns its song. *Science (New York, N. Y.), 291*(5513), 2564–2569. https://doi.org/10.1126/science.1058522

Tchernichovski, O., Nottebohm, F., Ho, C. E., Pesaran, B., & Mitra, P. P. (2000). A procedure for an automated measurement of song similarity. *Animal Behavior, 59*(6), 1167–1176. https://doi.org/10.1006/anbe.1999.1416

Téglás, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B., & Bonatti, L. L. (2011). Pure reasoning in 12-month-old infants as probabilistic inference. *Science (New York, N. Y.), 332*(6033), 1054–1059.

Thornton, A., & McAuliffe, K. (2006). Teaching in wild meerkats. *Science (New York, N. Y.), 313*(5784), 227–229. https://doi.org/10.1126/science.1128727

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., De-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: predictive coding in autism. *Psychological Review, 121*(4) 649.

Van Egeren, L. A., Barratt, M. S., & Roach, M. A. (2001). Mother–infant responsiveness: Timing, mutual regulation, and interactional context. *Developmental Psychology, 37*(5), 684–697. https://doi.org/10.1037/0012-1649.37.5.684

Varkevisser, J. M., Mendoza, E., Simon, R., Manet, M., Halfwerk, W., Scharff, C., & Riebel, K. (2022). Multimodality during live tutoring is relevant for vocal learning in zebra finches. *Animal Behaviour, 187*, 263–280. https://doi.org/10.1016/j.anbehav.2022.03.013

Varkevisser, J. M., Simon, R., Mendoza, E., How, M., Hiklkema, I., Jin, R., ... Riebel, K. (2022). Adding colour-realistic video images to audio playbacks increases stimulus engagement but does not enhance vocal learning in zebra finches. *Animal Cognition, 25*, 249–274. https://doi.org/10.1007/s10071-021-01547-8

Vyas, A., Harding, C., Borg, L., & Bogdan, D. (2008). Acoustic characteristics, early experience, and endocrine status interact to modulate female zebra finches' behavioral responses to songs. *Hormones and Behavior, 55*(1), 50–59. https://doi.org/10.1016/j.yhbeh.2008.08.005

Wang, M. Z., & Hayden, B. Y. (2019). Monkeys are curious about counterfactual outcomes. *Cognition, 189*, 1–10. https://doi.org/10.1016/j.cognition.2019.03.009

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science, 25*(7), 1314–1324. https://doi.org/10.1177/0956797614531023

West, M., & King, A. (1988). Female visual displays affect the development of male song in the cowbird. *Nature, 334*, 244–246.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Child Psychology & Psychiatry & Allied Disciplines, 17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Wu, S., Blanchard, T., Meschke, E., Aslin, R. N., Hayden, B. Y., & Kidd, C. (2022). Macaques preferentially attend to intermediately surprising information. *Biology Letters, 18*, 20220144. https://doi.org/10.1098/rsbl.2022.0144

Wyckoff, L. B., Jr. (1952). The role of observing responses in discrimination learning. Part 1. *Psychological Review, 59*(6), 431–442. https://doi.org/10.1037/h0053932

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

Zai, A. T., Cavé-Lopez, S., Rolland, M., et al. (2020). Sensory substitution reveals a manipulation bias. *Nature Communications, 11*, 5940. https://doi.org/10.1038/s41467-020-19686-w

Zeskind, P. S., & Shingler, E. A. (1991). Child abusers' perceptual responses to newborn infant cries varying in pitch. *Infant Behavior and Development, 14*(3), 335–347. https://doi.org/10.1016/0163-6383(91)90026-O

Zlochower, A. J., & Cohn, J. F. (1996). Vocal timing in face-to-face interaction of clinically depressed and nondepressed mothers and their 4-month-old infants. *Infant Behavior and Development, 19*(3), 371–374. https://doi.org/10.1016/S0163-6383(96)90035-1

Further reading

Kidd, C., & Hayden, B. Y. (2015). The psychology and neuroscience of curiosity. *Neuron, 88*(3), 449–460.