Acquiring Complex Communicative Systems: Statistical Learning of Language and Emotion

Ashley L. Ruba, Seth D. Pollak, Jenny R. Saffran

Department of Psychology, University of Wisconsin – Madison

Abstract

During the early postnatal years, most infants rapidly learn to understand two naturally evolved communication systems: language and emotion. While these two domains include different types of content knowledge, it is possible that similar learning processes subserve their acquisition. In this review, we compare the learnable statistical regularities in language and emotion input. We then consider how domain-general learning abilities may underly the acquisition of language and emotion, and how this process may be constrained in each domain. This comparative developmental approach can advance our understanding of how humans learn to communicate with others.

Keywords: Emotion; Language; Statistical learning; Development; Infancy

1. Introduction

Infants are faced with a staggering problem: they must learn to effectively communicate with other humans in their world, without explicit instructions about how to do so. To communicate with others, infants must learn to understand the complex signals sent by others and begin to generate interpretable signals of their own. Despite the scope of this challenge, neurotypical infants rapidly learn two naturally evolved communication systems-language and emotion-during the first few years of postnatal life. While these two domains include

different types of content knowledge, it is possible that similar learning processes underlie their acquisition. In this review, we consider how domain-general statistical learning abilities may support the acquisition of language and emotion. We focus on how infants come to *understand* input in each domain, rather than how infants learn to *produce* communicative signals. Specifically, we focus on how infants learn to understand foundational and early-acquired aspects of emotion and language (such as the meanings of facial configurations and words). This initial knowledge provides infants with new ways of representing the world, setting the stage for acquiring more complex and sophisticated aspects of language and emotion (e.g., pragmatics, complex syntax, social display rules, theory of mind). Our comparative developmental approach has the potential to further illuminate the nature of human learning.

1.1. Input to learners of complex communicative systems

As adults, we tend to think of our linguistic and emotional communication systems as objective and consistent, but it is unlikely that either system appears that way to an infant. Why should we preferentially attend to faces rather than fingers when discerning another person's emotional state? When listening to someone talk, why attend to vocal prosody and not eye blinks? The developing child learns to ignore a vast amount of available, but irrelevant, perceptual information, in order to focus on the most relevant regularities (e.g., infants learn that a sneer conveys emotional information, while a sneeze does not). Infants also learn to generalize and detect these relevant regularities over vast individual differences—in people's voices, faces, personalities, genders, ages, and other person-specific features. In these regard, infants' linguistic and emotional environments can be characterized, simultaneously, by both a richness and poverty of available input. The input is *rich* in that it contains massive amounts of linguistic and emotion-relevant cues. Yet the input is also *impoverished* in that it is largely unlabeled and incredibly noisy (both statistically and perceptually). From this complex assemblage of input, infants learn to formulate abstract meanings, categories, inferences, and generalizations about language and emotion. Below, we consider specific characteristics of the input available to young learners in each domain.

1.2. Input to infant language learners

From decades of research in linguistics and other areas of cognitive science, we know that all natural languages consist of structured sequences of sounds or signs, organized in a limited number of ways. These include sounds or sign features (e.g., phonemes), phonology (patterns of phonemes and other "musical" aspects of a language, like rhythm), morphemes (the smallest units of meanings), words (consisting of one or more morphemes), lexical categories (e.g., nouns and verbs), meanings (semantic representations), and syntax (patterns or rules built over lexical categories and other such elements). Starting from the simplest and most perceptually available components of this input (i.e., sounds in spoken languages; handshapes and movements in sign languages), infants arrive at the richer and more abstract components of language (e.g., grammars), which permit generalization beyond the input that has been received. Yet, there is still substantial debate regarding innate knowledge about

language (for a range of recent perspectives, see Christiansen & Chater, 2015; Lasnik & Lidz, 2016; Linzen & Baroni, 2021; Pearl, 2021; Perfors, Tenenbaum, & Regier, 2011; Shi, Legrand, & Brandenberger, 2020). Regardless of one's stance in this debate, infants are faced with complex learning problems, including learning idiosyncrasies of a specific language. Moreover, these problems change over the course of language learning. For example, an infant cannot begin to learn their native language's syntax until they have learned some of its words (Saffran & Wilson, 2003). The challenges facing language learners are thus underscored by the simultaneous richness and poverty of linguistic stimuli.

From the perspective of the richness of the stimuli, part of infants' challenge is to sift and winnow language input to discern relevant cues to native language structure. An example of this process comes from the varied ways in which pitch is used across languages. An English word spoken by two individuals (e.g., a female child vs. an adult male) may differ widely in pitch but have the same meaning. However, tonal languages like Mandarin or Hmong incorporate systematic pitch differences that completely alter the meanings of words. Thus, the structure of pitch in linguistic input varies from language to language, and infants must either learn to ignore this variation (if learning a non-tonal language) or track it (if learning a tonal language) (Hay, Graf Estes, Wang, & Saffran, 2015; Quam & Swingley, 2010). A similar example is found with phonemes. Although individuals pronounce phonemes differently across languages, infants are initially able to differentiate all phonemes regardless of their native language(s). With experience, however, infants have difficulty distinguishing phonemes in their non-native languages (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker & Tees, 2002). This "perceptual narrowing" process (also observed in other domains; Hannon & Trehub, 2005; Maurer & Werker, 2014; Pascalis et al., 2014) allows infants to expertly perceive relevant sound distinctions while ignoring other, less relevant distinctions (Kuhl et al., 2006; Maye, Werker, & Gerken, 2002).

Language input is also impoverished in many ways. Infant language learning is largely unsupervised, in that most input is unlabeled. Caregivers do not point out word boundaries in continuous speech or the meanings of many words that infants hear (consider the implausibility of a caregiver explaining the meanings of "the" or "it"). Caregivers also do not label lexical categories for infants (e.g., whether "book" is a noun or verb), and these categories are not clearly signaled by observable semantic information (e.g., how does "love" map onto the visible world?). Even in situations that seem clear, there is still extensive referential ambiguity. When an infant hears a given word (e.g., "dog") and sees an adult point at the animal, the word could refer to any number of observable features of an object—its color, its ears, the whole animal, the action in which the animal is engaged, etc. (Quine, 1960). Given that infants' language production lags months or even years behind their language comprehension, much of their knowledge is unobservable to caregivers. It is difficult to determine what words an infant understands and whether infants have correctly construed meanings and grammatical regularities. Errors are sometimes visible in children's language production, but caregivers seldom provide explicit negative feedback (i.e., caregivers do not correct infants' grammatical errors; Brown & Hanlon, 1970; Morgan & Travis, 1989). Infants learn their native language(s) largely in the absence of any explicit instruction. Many additional limitations in the input also exist, as highlighted in classic linguistic arguments about the "poverty

of the stimulus" (Chomsky, 1965). Yet, despite the complexity of the linguistic environment, infants nevertheless uncover key structures in language input with impressive speed.

1.3. Input to infant emotion learners

While language researchers have benefited from the field of linguistics, emotion researchers do not have an analogous discipline. For this reason, much less is known about the range, frequency, types, and extent of cultural variation in emotion cues, or the regularity of cues that infants encounter. Emotions are a social construction that refer to instances of feelings. There are many ways that humans can feel things, and there is no shared agreement among scholars about which feelings ought to be considered as "emotions" versus something distinct from emotions, such as reflexes, sensations, or sentiments. For this reason, some definitions of emotion include hunger, warmth or coldness, physical pain, desire, fatigue, thirst, interest, and the experience of touching something, whereas other definitions exclude these feelings from the construct of emotion. Nonetheless, there is shared agreement that an instance of emotion is when we categorize, organize, or otherwise make sense of the unfolding of our feelings.

Generally, the input for emotion learning includes various types of cues—such as paralanguage (e.g., sighs, grunts, giggles; Friend, 2000; Woodard, Plate, Morningstar, Wood, & Pollak, 2021), facial movements (e.g., activation of one or more facial muscles; Barrett, Adolphs, Marsella, Martinez, & Pollak, 2019), skin coloration (e.g., changes in blood flow and oxygenation contributing to blushing or pallor; Thorstenson, 2018; Thorstenson, McPhetres, Pazda, & Young, 2021), non-facial body movements (e.g., tightening or raising of fists, crossed arms, slumped shoulders; Aviezer et al., 2008; Witkower, Hill, Koster, & Tracy, 2021), behavioral action tendencies (e.g., approach, withdrawal; Adams, Ambady, Macrae, & Kleck, 2006), emotion language (words and phrases that denote affective states; Hoemann, Xu, & Barrett, 2019; Lakoff, 2016; Ruba, Meltzoff, & Repacholi, 2020b)—and the context (i.e., the broader social situation) in which any of the aforementioned actions occur (Ruba & Pollak, 2020; K. E. Smith, Leitzke, & Pollak, 2020). Starting from relatively simple and perceptually available components of this input (e.g., a sharp or gentle voice, an open or closed mouth), infants must arrive at abstract components of emotion (e.g., subjective, internal states and predictions about others' actions). As with research on language, there are still ongoing debates regarding potential innate specifications of emotion, and whether certain emotions are universally experienced and signaled (Barrett et al., 2019; Cowen, Sauter, Tracy, & Keltner, 2019). Nevertheless, there is agreement that displays of emotion vary across cultures and subcultures, and infants must learn complex rules depending on the idiosyncrasies of their social and cultural group (e.g., whether to mask/suppress a negative emotional response; Crivelli, Russell, Jarillo, & Fernández-Dols, 2016; Elfenbein & Ambady, 2002; Matsumoto, 1993).

Our account is not limited to any particular type of emotion. Rather, our concern is how children learn to organize their feelings into whatever meaningful categories of emotion exist in their social world to make sense of their experiences and facilitate communication. Contrary to popular belief, emotion categories such as sadness, happiness, anger, fear, disgust, and surprise are not "basic" emotions that are highly similar across humans. These

categories are merely those that English-speaking scientists have defined and studied the most frequently. A child exposed to English might use a label of sadness to organize an experience of loss, a speaker of Tagalog might use *gigil* to make sense of their irresistible urge to pinch or squeeze something cute (Lomas, 2016), and the Ilongot might categorize their feeling of an exuberant surge or burst of energy as *liget* (Wierzbicka, 1992). Rather than focusing on any subset of emotion linguistic categories, our aim is to understand how children develop the capacity to infer emotional meaning from observable cues in their environments.

Like language input, emotional input is characterized by vast richness. Other people's feelings cannot be directly observed, and no single muscle movement, behavior, or utterance consistently indicates a specific emotional state (Barrett et al., 2019). We can smile when we feel happy, but also when we feel nervous, embarrassed, manipulative, proud, satiated, superior, pained, subservient, irritated, or when we are attempting to mask another emotion (Martin, Rychlowska, Wood, & Niedenthal, 2017; Seyfarth & Cheney, 2003). Similarly, when we feel happy, we may smile, cry "tears of joy," or try to hide our emotion, particularly if it comes after another person's misfortune. Given the indeterminate nature of these emotion cues, they always need to be interpreted within the larger context in which they are expressed–smiling at a birthday party is quite different than smiling when being scolded. Emotions are also probabilistic even within a given context: puppies usually spark joy, but perhaps not if one has recently lost their pet or was bitten by a dog. Infants also need to appreciate that cues sometimes used to convey emotional states can also occur for other physical sensations or cognitive states. Crossed arms may indicate that someone feels physically cold, while a furrowed brow may reflect that someone is intensely focused. Moreover, the cause of an emotion could be an observable event in the environment, or an emotion may originate from thoughts or memories, with no observable referent whatsoever. Emotional input is thus characterized by both referential ambiguity and variability. At the same time, emotions reflect some systematicity; they are not random occurrences (although the degree of this systematicity remains hotly debated, see Barrett, 2021; Cowen et al., 2021; Le Mau et al., 2021).

Also similar to language, the processes through which infants learn to understand emotion cues are largely unsupervised (Ruba & Repacholi, 2020b). Parents infrequently label emotions in their interactions with children. One content analysis found that, in conversations with 2- to 5-year-old children, parents frequently use terms for general attitudes ("like", "love"), but rarely use labels for specific emotions like "happy" (Lagattuta & Wellman, 2002). There is little evidence that caregivers engage in explicit instruction about emotions: "I am smiling because I feel happy" or "I feel happy because it's my birthday" (Brownell, Svetlova, Anderson, Nichols, & Drummond, 2013). Even words that are used as labels for emotions are not clearly signaled by observable information (e.g., it is not readily apparent how "love," "pride," "shame," or "hurt feelings" map onto the visible world). Similarly, a child may hear "leave me alone" or "go away" from a sibling, without reference to any emotion category (e.g., does such a statement reflect that the other person is busy, sad, angry, hungry, tired, or something else?). Still, long before they can produce emotion labels, infants attend to and use cues about others' feelings to guide their own behavior (for recent reviews, see Ruba & Pollak, 2020; Ruba & Repacholi, 2020a).

1.4. Comparing language and emotion input

As highlighted above, the problems facing language and emotion learners are similar in their overall structure. Infants must sort through myriad relevant and irrelevant cues to detect meaningful patterns, with virtually no supervised instruction. Yet, the input in these two domains differs in important ways. On the most surface level, language input is temporal (in the case of spoken languages) or both temporal and spatial (sign languages use visual space in a structured way), whereas emotion input is both temporal and spatial. Language input is either largely auditory (as in spoken languages, with additional information provided by the visual world, including speech cues on the face) or entirely visual (as in sign languages), while emotion input is communicated through an ever-changing mix of sensory modalities. These input differences influence the learning problems that infants face.

Aspects of language and emotion input also differ with respect to their abstractness, and, as a result, their predictability. For instance, there are agreed-upon "correct" labels for referents within languages, such as what should be labeled a "dog" versus a "cat." Many object labels are reliably used to refer to aspects of the environment. An adult native-English speaker is unlikely to label a dog as a "cat," and when they do, children lose trust in the speaker (for a recent review, see Tong, Wang, & Danovitch, 2020). In contrast, while there are cultural norms regarding the appropriateness of certain emotions (e.g., laughing at another person's pain is unkind), there is not a single "correct" way for emotions to be felt and expressed, and there is not a consistent pairing between any discrete event and any specific emotion or set of emotions. In this way, the co-occurrences of expressive behaviors and internal affective states may be less predictable than the co-occurrences of labels and referents. Some aspects of language are similarly abstract (e.g., the meaning of "the" or "truth") or inconsistent (e.g., what is described as "yummy" may vary across individuals). Yet, they are nevertheless predictable once the child has learned about them (e.g., "the" precedes nouns in English, providing a clue to the syntactic category of "truth" even without knowing its meaning).

Despite the complexity of the input in each of these two domains, most infants learn to understand aspects of others' linguistic and emotion cues during the first 2 years of postnatal life. In the next section, we consider how infants learn these rich communicative systems.

2. Learning complex communicative systems

Traditionally, theories about language and emotion development emphasize either (a) evolutionarily preserved, universal aspects of each domain, or (b) functional and environmental adaptations of each domain—and, to be clear, there is compelling evidence for both. According to the former, *nativist*, perspectives, the human brain possesses rudimentary features or biases that direct infant learning to salient aspects of the environment. Nativist perspectives offer an account of similarities across individuals and cultures, and a plausible explanation for how infants understand language and emotion so early in development. In contrast, *empiricist* views hold that, prior to sensory learning, the human brain does not include content, biases, or packages of skills. Instead, empiricist approaches emphasize the role of sensory

experience as the basis of knowledge. This approach accounts for variability across individuals and similarities across domains of cognition. Each of these perspectives assumes a different initial state of knowledge about language and emotion in the human brain.

These questions about the initial state of the human brain are often confounded with issues about how knowledge is acquired or transformed. One way in which developmental change might occur is that specialized capacities support infant learning of specific skills or information. This view, *domain-specificity*, holds that many aspects of cognition, including language and emotion, are supported by distinct, evolutionarily-specified learning processes or biases that guide acquisition. These processes are tailored to, and prepared for, specific input in a particular domain. In contrast, *domain-general* theories argue that a common set of computational principles drive learning across multiple domains. In this view, domains develop differently not because of their initial state, but because of differences in the sensory or structural properties of their inputs. Domain specificity is typically associated with nativism, whereas domain generality is often paired with empiricist approaches. But these types of theories need not be paired. In principle, the nativist/empiricist dimension is orthogonal from the domain-specific/domain-general dimension. Below, we highlight how a nativist, domain-general approach can be applied to language and emotion learning.

2.1. *Statistical learning in human infants*

Humans are born with an array of perceptual and cognitive abilities used to interact with and learn from the world. One such ability is sensitivity to statistical regularities in the environment. Infants can track a range of statistical regularities, including exemplar frequency, forwards and backward transitional probabilities, non-adjacent co-occurrences, and category-level patterns (for recent reviews, see Frost, Armstrong, & Christiansen, 2019; Saffran & Kirkham, 2018). This sensitivity to statistical regularities does not appear to be acquired via experience: it is evident days after birth (Bulf, Johnson, & Valenza, 2011; Fló et al., 2019; Teinonen, Fellman, Näätänen, Alku, & Huotilainen, 2009) and in many non-human animal species (Boros et al., 2021; Santolin & Saffran, 2018; Wilson et al., 2020). While infants detect statistical regularities in the auditory and visual modalities, infants are generally more proficient at tracking auditory sequences than visual sequences (Emberson, Misyak, Schwade, Christiansen, & Goldstein, 2019; Krogh, Vlach, & Johnson, 2013)—likely because the auditory world is fleeting and highly sequential, while many aspects of the visual environment are more stable (Conway & Christiansen, 2009; Saffran, 2002). Infants can also track cross-modal correlations, such as those between labels and objects (L. B. Smith & Yu, 2008; Vouloumanos & Werker, 2009) or between facial movements and vocalizations (Grossmann, Striano, & Friederici, 2006; Kahana-Kalman & Walker-Andrews, 2001; Vaillant-Molina, Bahrick, & Flom, 2013). This sensitivity to regularities in the environment, including distributions, probabilities of co-occurrence, and correlations, is known as *statistical learning*.

Below, we suggest that statistical learning approaches can provide explanations for how infants learn to comprehend key aspects of language and emotion. Statistical learning abilities appear to be domain-general, in that their computational underpinnings are not designed for a specific domain (i.e., language or emotion), but are available for learning information drawn

from numerous domains. Despite being domain-general, our statistical learning approach is also nativist, in the sense that learning is constrained by genetically endowed computational abilities and constraints on perception and processing. This perspective is also inherently developmental, in that infants' learning is constrained by emerging cognitive abilities (e.g., attention, working memory), prior sensory experiences, and learning in other domains (e.g., social cognition). Below, we briefly overview evidence concerning infant statistical learning of language and emotion, focusing on how and why infants may track statistics across each of these domains.

2.2. *Statistical regularities in linguistic and emotional input*

For statistical learning to occur, language and emotional input must contain learnable regularities. Early statistical learning studies focused on word segmentation: how infants detect where words begin and end in the absence of clear perceptual boundaries, like pauses (Aslin, Saffran, & Newport, 1998; Goodsitt, Kuhl, & Morgan, 1993; Saffran, Aslin, & Newport, 1996). This body of research suggests that infants are sensitive to probabilities of syllable co-occurrence—important cues to word boundaries—in both simplified artificial languages and natural speech (for recent theoretical reviews, see Erickson & Thiessen, 2015; Saffran, 2020). Infants are also sensitive to statistics at different levels of analysis, including distributional statistics for phoneme categories (Maye et al., 2002), cross-situational statistics for label-referent pairs (L. B. Smith & Yu, 2008; Vouloumanos & Werker, 2009), and category-based regularities in grammatical structures (Gomez & Gerken, 1999; Lany & Saffran, 2010; Saffran et al., 2008). In each case, infants' statistical learning abilities appear well-suited to linguistic input, at least in simplified laboratory tasks. Indeed, it is plausible that linguistic input itself has been shaped by human statistical learning abilities: only learnable structures should persist in the languages of the world (Christiansen & Chater, 2008; Saffran, 2001).

Emotion learning has not traditionally been framed in terms of statistical learning, but more recent accounts are beginning to examine the role of regularities in children's emotional environments (Doan, Friedman, & Denison, 2018; Plate, Wood, Woodard, & Pollak, 2019; Woodard et al., 2021). Similar to sequences of syllables, infants are sensitive to transitional probabilities in artificial sequences of facial configurations (Mermier, Quadrelli, Turati, & Bulf, 2022). Infants are also sensitive to category-based regularities in expressive behaviors. As an example, they perceive a group of individuals who are smiling to belong to the same category, in contrast to individuals who are not smiling (Ruba & Repacholi, 2020a). Infants also expect that laughter will co-occur with smiles, and that a person will smile rather than frown after receiving a gift (Grossmann et al., 2006; Kahana-Kalman & Walker-Andrews, 2001; Ruba, Meltzoff, & Repacholi, 2019; Vaillant-Molina et al., 2013). Infants use these regularities to make predictions about other people's behavior, expecting that someone who appears to be angry will continue to display anger in new social contexts (Doan, Friedman, & Denison, 2020; Plate et al., 2019; Repacholi, Meltzoff, Hennings, & Ruba, 2016; Repacholi, Meltzoff, Toub, & Ruba, 2016). Taken together, these data indicate that infants and young children are sensitive to various statistical regularities in emotion input.

2.3. Primitives and constraints for statistical learning

There are an infinite number of statistics that infants could compute over the input that they experience in their environments. How do infants determine which regularities to track? One important factor is that infants track statistics over a limited set of perceptual primitives, which necessarily vary across domains. In the domain of language, infants learning spoken language initially track information about distributions of phonemes or syllables, whereas infants learning sign languages track information about the distribution of handshapes and movement trajectories. Learners acquiring languages in both modalities attend to regularities in the face, though the relevant information differs—the mouth provides speech cues for spoken languages (including information about rhythm and phonemic content), whereas sign languages use a wide range of facial cues including both the eyes and mouth. The primitives over which infants compute language statistics are also affected by prior learning. As novice learners, infants are confronted with a stream of sounds (or signs). Before they can learn word meanings, they must first figure out where words begin and end (a complex endeavor given that there are no reliable acoustic markers of word boundaries in fluent speech). Similarly, infants cannot track relationships among lexical categories (like nouns and verbs) until they have figured out which words belong to each category. Languages also contain statistical distributions relevant for some learners but not others. As previously mentioned, lexical tones (i.e., pitches and pitch contours associated with words) are crucial for acquiring vocabulary in tonal languages like Mandarin or Hmong, but not for languages like English, where the tone is uncorrelated with word meanings. Thus, the primitives that are used for language learning change dynamically as learning unfolds. What infants have already learned alters what becomes available to learn in the future.

Far less research has examined the primitives for emotion learning. However, extant research provides clues as to the information that may be most relevant to infant learners. For instance, infants may initially track information about transitional probabilities and co-occurrences in expressive behaviors across modalities, particularly between faces and voices (newborns have attentional biases for faces and primate vocalizations; Johnson, Dziurawiec, Ellis, & Morton, 1991; Mermier et al., 2022; Morton & Johnson, 1991; Shackman & Pollak, 2005; Vouloumanos, Hauser, Werker, & Martin, 2010). Infants may also track the situational contexts in which these expressive behaviors occur, such as those contexts associated with alleviation of distress (Moriceau & Sullivan, 2006). As with language, the primitives over which infants compute statistics are likely impacted by prior learning. Around 7 months of age, infants preferentially attend to (and perhaps track statistics for) negative expressive behaviors (for a review, see Vaish, Grossmann, & Woodward, 2008). These changing attentional biases may reflect an increase in caregivers' negative expressive behaviors, coinciding with the onset of infant self-produced locomotion (Campos et al., 2000). Additionally, emotions likely contain statistical distributions relevant for some learners but not others. Infants with depressed or maltreating caregivers may primarily observe negative expressive behaviors (Plate et al., 2019), and thus, may preferentially (and adaptively) track these behaviors across different situations. As with language, the primitives used for emotion learning are likely dynamic and changing across the first few years of life alongside other aspects of

cognitive, motor, and social development (Herzberg, Fletcher, Schatz, Adolph, & Tamis-LeMonda, 2021; Hoemann et al., 2020; Ruba & Pollak, 2020).

Another key constraint on statistical learning lies in the computations themselves. Infants do not track every statistical regularity available to them, and many questions remain about the limits of statistical learning (for additional discussion, see Saffran & Kirkham, 2018). One hint comes from research on structures in language and emotion input. For example, infants better learn auditory units that contain predictive patterns in both linguistic and nonlinguistic materials (Saffran et al., 2008; Santolin & Saffran, 2019). The structures that infants find more learnable are precisely those that tend to occur in the languages of the world. It is plausible that similar connections occur in emotion learning. For instance, caregivers may naturally express their emotions in ways that maximize learning possibilities (e.g., "emotionese"; Benders, 2013; Ruba & Repacholi, 2020b; Schachner & Hannon, 2011; Trainor, Austin, & Desjardins, 2000; Wu, Schulz, Frank, & Gweon, 2021). This includes displaying emotions in congruent, multimodal ways across predictable situations—typically smiling and laughing during playtime, frowning and raising their voice when mediating a tantrum, or exclaiming "ew" with a scrunched nose when changing a diaper. Like language, these predictive patterns may help infants learn about when and how other people tend to display their emotions.

Finally, statistical learning is constrained by the bodies and brains in which it occurs—infants are limited in the types of information they perceive and in the ways that they can act on the world. For instance, newborns' relatively poor visual acuity and motor abilities bias their attention to proximally close objects, such as faces. The faces that dominate infants' early visual environments are persistent in the visual field and tend to be of fewer than three individuals (Jayaraman & Smith, 2019; Jayaraman, Fausey, & Smith, 2017). With these early, frequent, and close-up experiences with faces, infants may preferentially attend to facial movements that provide information about their caregivers' emotional state and the language environment (e.g., rhythm regularities). With advances in visual and motor development, older infants turn their attention to hands and objects (Fausey, Jayaraman, & Smith, 2016). This provides new visual access to labeled referents in the environments as well as the situational contexts in which emotionally expressive behaviors occur. Memory development also constrains statistical learning. The limited working abilities of infants may, paradoxically, help learning in domains where input is noisy. In noisy environments, infants can hone in on generalizations rather than being sidetracked by specific exemplars (Newport, 1990). Together, these examples illustrate how various developmental processes work to constrain and focus learning (Ruba & Pollak, 2020).

2.4. Why do infants track statistical regularities?

Language and emotion are evolutionarily advantageous communicative systems. These systems not only allow infants to learn about their environment but also to share their desires, goals, and internal states with other humans. Infants are likely intrinsically motivated to learn about language and emotion to communicate and navigate their social worlds. However, it is an open question whether the drive to communicate is a key factor very early in development:

as noted previously, even neonates track statistical regularities (Bulf et al., 2011; Fló et al., 2019; Teinonen et al., 2009).

Infants may also track statistics because statistical regularities facilitate predictive processing (Köster, Kayhan, Langeloh, & Hoehl, 2020; Romberg & Saffran, 2013; Saffran, 2020). That is, rather than possessing isolated "statistical learning mechanisms," infants may exploit statistical regularities to detect and learn from prediction errors (Elman, 1990), to generate faster and more accurate predictions about upcoming input, and as part of the process of organizing information in memory (e.g., Thiessen, 2017). Predictive information is highly valuable for encoding and processing both language and emotion. Real-time language processing is extremely challenging, given the rapidity with which linguistic signals unfold. Statistical regularities in words and word combinations may speed up infants' ability to process this information, map it to meanings, and deal with challenging listening conditions (Graf Estes, Evans, Alibali, & Saffran, 2007; Lany, Shoaib, Thompson, & Estes, 2018; McMillan & Saffran, 2016). Relatedly, words presented in predictable contexts may be easier to learn than those presented in less predictable contexts (Benitez & Saffran, 2018; Benitez & Smith, 2012; Eiteljoerge, Adam, Elsner, & Mani, 2019). Real-time emotion processing is also challenging, given the many manners and situations in which emotions are expressed. Infants may use statistical regularities in a person's emotional behavior to predict how that person will behave in novel situations (Repacholi, Meltzoff, Hennings, et al., 2016; Repacholi, Meltzoff, Toub, et al., 2016).

Finally, statistical information may help infants direct their learning and attentional resources toward the most informative data in their environment. In other words, tracking statistics can help infants reduce uncertainty. During word learning, children and adults preferentially sample items about which they are more uncertain (Zettersten & Saffran, 2021). In the affective domain, infants often engage in social referencing (i.e., seeking out emotional cues from an adult) during novel and ambiguous situations (Kim & Kwak, 2011; Sorce, Emde, Campos, & Klinnert, 1985). Thus, rather than being passive "sponges" soaking up regularities, infants are active learners, using statistics to inform the data that they sample. Consideration of the infant as an active learner may explain how infants learn so much so quickly (Mani & Ackermann, 2018; Raz & Saxe, 2020). Infants may direct their prodigious learning abilities toward the sources that maximize information gain, making learning more efficient.

3. Outstanding questions

In this review, we have highlighted how domain-general statistical learning abilities may explain how humans learn about key aspects of language and emotion. This approach provides potential explanations for differences in what and when infants learn about these domains. Our approach suggests that input that is structured, consistent, predictable, and engaging to infants should be learned more rapidly than input that is riddled with exceptions and noise, or that is not particularly interesting to young learners. In other words, differences in the input to learning contribute to differences in infants' ability to learn from this input.

This approach opens the door for exciting future research directions. First, researchers—particularly emotion researchers—need to precisely describe the structure of learning input. What range, frequency, and types of input do infants observe in their daily lives? What are the learnable statistical regularities in this input? From this information, researchers can begin to specify the primitives over which infants track statistical information. Are these primitives apparent from birth, or are they learned via exposure to language and emotion input? For instance, it is possible that infants initially attend to emotion information on faces since faces are evolutionary-adaptive emotion signals (Ekman, 1994), or since faces are salient and persistent in newborns' visual fields (L. B. Smith, Jayaraman, Clerkin, & Yu, 2018). Similarly, communicative sounds—even those from other species—seem to help young infants organize their experiences into categories that will eventually become words (for a recent review, see Perszyk & Waxman, 2018). It is also imperative to specify how infant learning is constrained by other developing abilities, such as memory, attention, and language. Comparing children at multiple ages will be critical to this work (for recent examples, see Raviv & Arnon, 2018; Ruba, Meltzoff, & Repacholi, 2020a)

It will also be important to move beyond "classic" statistical learning tasks (i.e., a learning phase followed by a separate test phase). Statistical learning researchers have made great strides in developing novel, continuous measures of real-time learning in adults and children (Arnon, 2020; Frost et al., 2019; Kidd et al., 2020). Infancy researchers are beginning to do so as well, with both neuroimaging methods (e.g., Functional magnetic resonance imaging [fMRI], Functional near-infrared spectroscopy [fNIRS], Electroencephalography[EEG]; Choi, Batterink, Black, Paller, & Werker, 2020; Ellis et al., 2021; Emberson, Richards, & Aslin, 2015) and behavioral measures (e.g., pupillometry, anticipatory eye movement tasks) (Havron, de Carvalho, Fiévet, & Christophe, 2019; Reuter, Borovsky, & Lew-Williams, 2019; Romberg & Saffran, 2013; Zhang & Emberson, 2020). By moving away from passive measures of learning, researchers can discover how infants actively solve learning problems. Further, many paradigms used to assess statistical learning use artificial stimuli (static faces, disembodied voices) and/or clearly non-communicative stimuli like geometric shapes or computer alert sounds. Although infants learn in these paradigms, it seems unlikely that infants perceive this learning as advantageous for social communication. Future research can help to tease apart myriad motivations for learning, from uncertainty reduction to a drive to communicate.

It will also be important to better understand the computational underpinnings of learning in these two domains. Other approaches, such as Bayesian cognitive models, have also been applied to both language learning and, more recently emotion learning (e.g., Wu et al., 2021; Xu & Tenenbaum, 2007). There are many important differences between, for example, Bayesian and connectionist models. However, researchers are converging on approaches to understanding development in which environmental input is treated as information that the child learns by generating predictions and updating those predictions based on errors. The goal of such learning is to reduce uncertainty. Perhaps the types of learning outlined in the current paper—starting from primitives available to infants very early in life—provide some of the priors that both constrain the acquisition of more complex aspects of language and emotion and serve as the basis for the development of increasingly complex behaviors.

In sum, we have highlighted how a comparative developmental approach—intersecting language, emotion, and potentially other domains (e.g., social category development, multimodal perceptual development)—provides a useful lens through which to consider the learning problems faced by fledgling communicators. As researchers, we tend to focus on our own domain of expertise (e.g., language, emotion). However, considering how development unfolds through the lenses of multiple domains can generate useful insights that may not be obvious when focusing on a single domain (e.g., Hoemann et al., 2020; Maurer & Werker, 2014). By looking beyond a single domain and, ideally, by developing collaborative research programs, researchers may come closer to understanding how humans rapidly learn to understand and communicate with others.

Acknowledgments

Note

1 Here, we do not address the ontogeny of brain specialization—that is, whether brain areas that perform statistical computations are innately specialized for different inputs or whether specialization is a developmental process, constrained by computational properties and neural connectivity to different inputs (e.g., Behrmann & Plaut, 2020; Frost, Armstrong, Siegelman, & Christiansen, 2015; Schapiro, Turk-Browne, Botvinick, & Norman, 2017).

References

Adams, R. B., Ambady, N., Macrae, C. N., & Kleck, R. E. (2006). Emotional expressions forecast approach-avoidance behavior. *Motivation and Emotion*, *30*(2), 179–188. https://doi.org/10.1007/s11031-006-9020-2

Arnon, I. (2020). Do current statistical learning tasks capture stable individual differences in children? An investigation of task reliability across modality. *Behavior Research Methods*, *52*(1), 68–81. https://doi.org/10.3758/s13428-019-01205-5

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *94*, 321–324. https://doi.org/10.1111/1467-9280.00063

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A. K., … Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science*, *19*(7), 724–732. https://doi.org/10.1111/j.1467-9280.2008.02148.x

Barrett, L. F. (2021). Debate about universal facial expressions goes big. *Nature*, *589*, 202–203.

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional Expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, *20*(1), 1–68. https://doi.org/10.1177/1529100619832930

Behrmann, M., & Plaut, D. C. (2020). Hemispheric organization for visual object recognition: A theoretical account and empirical evidence. *Perception*, *49*(4), 373–404

Benders, T. (2013). Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behavior and Development*, *36*(4), 847–862. https://doi.org/10.1016/j.infbeh.2013.09.001

Benitez, V. L., & Saffran, J. R. (2018). Predictable events enhance word learning in toddlers. *Current Biology*, *28*(17), 2787–2793.e4. https://doi.org/10.1016/j.cub.2018.06.017

Benitez, V. L., & Smith, L. B. (2012). Predictable locations aid early object name learning. *Cognition*, *125*(3), 339–352. https://doi.org/10.1016/j.cognition.2012.08.006

Boros, M., Magyari, L., Török, D., Bozsik, A., Deme, A., & Andics, A. (2021). Neural processes underlying statistical learning for speech segmentation in dogs. *Current Biology*, *31*(24), 5512–5521.

Brown, R., & Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 11–53). New York: Wiley.

Brownell, C. A., Svetlova, M., Anderson, R., Nichols, S. R., & Drummond, J. (2013). Socialization of early prosocial behavior: Parents' talk about emotions is associated with sharing and helping in toddlers. *Infancy*, *18*(1), 91–119. https://doi.org/10.1111/j.1532-7078.2012.00125.x

Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, *121*(1), 127–132. https://doi.org/10.1016/j.cognition.2011.06.010

Campos, J. J., Anderson, D. I., Barbu-Roth, M. A., Hubbard, E. M., Hertenstein, M. J., & Witherington, D. C. (2000). Travel broadens the mind. *Infancy*, *1*(2), 149–219. https://doi.org/10.1207/S15327078IN0102_1

Choi, D., Batterink, L. J., Black, A. K., Paller, K. A., & Werker, J. F. (2020). Preverbal infants discover statistical word patterns at similar rates as adults: Evidence from neural entrainment. *Psychological Science*, *31*(9), 1161–1173. https://doi.org/10.1177/0956797620933237

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT press

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*(5), 489–509. https://doi.org/10.1017/S0140525/08004998

Christiansen, M. H., & Chater, N. (2015). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*, e62. https://doi.org/10.1017/S0140525/1500031X

Conway, C. M., & Christiansen, M. H. (2009). Seeing and hearing in space and time: Effects of modality and presentation rate on implicit statistical learning. *European Journal of Cognitive Psychology*, *21*(4), 561–580. https://doi.org/10.1080/09541440802097951

Cowen, A. S., Keltner, D., Schroff, F., Jou, B., Adam, H., & Prasad, G. (2021). Sixteen facial expressions occur in similar contexts worldwide. *Nature*, *589*(7841), 251–257. https://doi.org/10.1038/s41586-020-3037-7

Cowen, A. S., Sauter, D. A., Tracy, J. L., & Keltner, D. (2019). Mapping the passions: Toward a high-dimensional taxonomy of emotional experience and expression. *Psychological Science in the Public Interest*, *20*(1), 69–90. https://doi.org/10.1177/1529100619850176

Crivelli, C., Russell, J. A., Jarillo, S., & Fernández-Dols, J.-M. (2016). The fear gasping face as a threat display in a Melanesian society. *Proceedings of the National Academy of Sciences*, *113*(44), 12403–12407. https://doi.org/10.1073/pnas.1611622113

Doan, T., Friedman, O., & Denison, S. (2018). Beyond belief: The probability-based notion of surprise in children. *Emotion*, *18*(8), 1163–1173. https://doi.org/10.1037/emo0000394

Doan, T., Friedman, O., & Denison, S. (2020). Young children use probability to infer happiness and the quality of outcomes. *Psychological Science*, *31*(2), 149–159. https://doi.org/10.1177/0956797619895282

Eiteljoerge, S. F. V., Adam, M., Elsner, B., & Mani, N. (2019). Consistency of co-occurring actions influences young children's word learning. *Royal Society Open Science*, *6*(8), 190097. https://doi.org/10.1098/rsos.190097

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin*, *115*(2), 268–287. https://doi.org/10.1037/0033-2909.115.2.268

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, *128*(2), 203–235. https://doi.org/10.1037/0033-2909.128.2.203

Ellis, C. T., Skalaban, L. J., Yates, T. S., Bejjanki, V. R., Córdova, N. I., & Turk-Browne, N. B. (2021). Evidence of hippocampal learning in human infants. *Current Biology*, *31*(15), 1–7. https://doi.org/10.1016/j.cub.2021.04.072

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

Emberson, L. L., Misyak, J. B., Schwade, J. A., Christiansen, M. H., & Goldstein, M. H. (2019). Comparing statistical learning across perceptual modalities in infancy: An investigation of underlying learning mechanism(s). *Developmental Science*, *22*(6), 1–10. https://doi.org/10.1111/desc.12847

Emberson, L. L., Richards, J. E., & Aslin, R. N. (2015). Top-down modulation in the infant brain: Learning-induced expectations rapidly affect the sensory cortex at 6 months. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(31), 9585–9590. https://doi.org/10.1073/pnas.1510343112

Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language: Theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review*, *37*, 66–108.

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

Fló, A., Brusini, P., Macagno, F., Nespor, M., Mehler, J., & Ferry, A. L. (2019). Newborns are sensitive to multiple cues for word segmentation in continuous speech. *Developmental Science*, *22*(4), 1–17. https://doi.org/10.1111/desc.12802

Friend, M. (2000). Developmental changes in sensitivity to vocal paralanguage. *Developmental Science*, *32*, 148–162. https://doi.org/10.1111/1467-7687.00108

Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin*, *145*(12), 1128–1153. https://doi.org/10.1037/bul0000210

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, *193*, 117–125.

Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*(2), 109–135. https://doi.org/10.1016/S0010-0277(99)00003-7

Goodsitt, J. V., Kuhl, P. K., & Morgan, J. L. (1993). Perceptual strategies in prelingual speech segmentation. *Journal of Child Language*, *20*(2), 229–252. https://doi.org/10.1017/S0305000900008266

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, *18*(3), 254–260. https://doi.org/10.1111/j.1467-9280.2007.01885.x

Grossmann, T., Striano, T., & Friederici, A. D. (2006). Crossmodal integration of emotional information from face and voice in the infant brain. *Developmental Science*, *93*, 309–315. https://doi.org/10.1111/j.1467-7687.2006.00494.x

Hannon, E. E., & Trehub, S. E. (2005). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences*, *102*(35), 12639–12643.

Havron, N., de Carvalho, A., Fiévet, A. C., & Christophe, A. (2019). Three- to four-year-old children rapidly adapt their predictions and use them to learn novel word meanings. *Child Development*, *90*(1), 82–90. https://doi.org/10.1111/cdev.13113

Hay, J. F., Graf Estes, K., Wang, T., & Saffran, J. R. (2015). From flexibility to constraint: The contrastive use of lexical tone in early word learning. *Child Development*, *86*(1), 10–22. https://doi.org/10.1111/cdev.12269

Herzberg, O., Fletcher, K. K., Schatz, J. L., Adolph, K. E., & Tamis-LeMonda, C. S. (2021). Infant exuberant object play at home: Immense amounts of time-distributed, variable practice. *Child Development*, *93*(1), 150-164. https://doi.org/10.1111/cdev.13669

Hoemann, K., Wu, R., LoBue, V., Oakes, L. M., Xu, F., & Barrett, L. F. (2020). Developing an understanding of emotion categories: Lessons from objects. *Trends in Cognitive Sciences*, *24*(1), 39–51. https://doi.org/10.1016/j.tics.2019.10.010

Hoemann, K., Xu, F., & Barrett, L. F. (2019). Emotion words, emotion concepts, and emotional development in children: A constructionist hypothesis. *Developmental Psychology*, *55*(9), 1830–1849. https://doi.org/10.1037/dev0000686

Jayaraman, S., Fausey, C. M., & Smith, L. B. (2017). Why are faces denser in the visual experiences of younger than older infants? *Developmental Psychology*, *53*(1), 38–49. https://doi.org/10.1037/dev0000230

Jayaraman, S., & Smith, L. B. (2019). Faces in early visual environments are persistent not just frequent. *Vision Research*, *157*, 213–221. https://doi.org/10.1016/j.visres.2018.05.005

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1–2), 1–19. https://doi.org/10.1016/0010-0277(91)90045-6

Kahana-Kalman, R., & Walker-Andrews, A. S. (2001). The role of person familiarity in young infants' perception of emotional expressions. *Child Development*, *72*(2), 352–369. https://doi.org/10.1111/1467-8624.00283

Kidd, E., Arciuli, J., Christiansen, M. H., Isbilen, E. S., Revius, K., & Smithson, M. (2020). Measuring children's auditory statistical learning via serial recall. *Journal of Experimental Child Psychology*, *200*, 104964. https://doi.org/10.1016/j.jecp.2020.104964

Kim, G., & Kwak, K. (2011). Uncertainty matters: Impact of stimulus ambiguity on infant social referencing. *Infant and Child Development*, *20*(5), 449–463. https://doi.org/10.1002/icd.708

Köster, M., Kayhan, E., Langeloh, M., & Hoehl, S. (2020). Making sense of the world: Infant learning from a predictive processing perspective. *Perspectives on Psychological Science*, *15*(3), 1–10. https://doi.org/10.1177/1745691619895071

Krogh, L., Vlach, H. A., & Johnson, S. P. (2013). Statistical learning across development: Flexible yet constrained. *Frontiers in Psychology*, *3*, 1–11. https://doi.org/10.3389/fpsyg.2012.00598

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*(2), 12–21. https://doi.org/10.1111/j.1467-7687.2006.00468.x

Kuhl, P. K., Williams, K., Lacerda, F., Stevens, K., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*(5044), 606–608. https://doi.org/10.1126/science.1736364

Lagattuta, K. H., & Wellman, H. M. (2002). Differences in early parent-child conversations about negative versus positive emotions: Implications for the development of psychological understanding. *Developmental Psychology*, *38*(4), 564–580. https://doi.org/10.1037/0012-1649.38.4.564

Lakoff, G. (2016). Language and emotion. *Emotion Review*, *8*(3), 269–273. https://doi.org/10.1177/1754073915595097

Lany, J., & Saffran, J. R. (2010). From statistics to meaning: Infants' acquisition of lexical categories. *Psychological Science*, *2*(12), 284–291. https://doi.org/10.1177/0956797609358570

Lany, J., Shoaib, A., Thompson, A., & Estes, K. G. (2018). Infant statistical-learning ability is related to real-Time language processing. *Journal of Child Language*, *45*(2), 368–391. https://doi.org/10.1017/S0305000917000253

Lasnik, H. & Lidz, J. L. (2016). The argument from the poverty of the stimulus. In I. Robert (Ed.), *The Oxford Handbook of Universal Grammar*. (pp. 221–248) Oxford, UK: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199573776.013.10

Le Mau, T., Hoemann, K., Lyons, S. H., Fugate, J. M. B., Brown, E. N., Gendron, M., & Barrett, L. F. (2021). Professional actors demonstrate variability, not stereotypical expressions, when portraying emotional states in photographs. *Nature Communications*, *12*(1), 1–13. https://doi.org/10.1038/s41467-021-25352-6

Linzen, T., & Baroni, M. (2021). Syntactic Structure from Deep Learning. *Annual Review of Linguistics*, *7*, 195–212. https://doi.org/10.1146/annurev-linguistics-032020-051035

Lomas, T. (2016). Towards a positive cross-cultural lexicography: Enriching our emotional landscape through 216 'untranslatable' words pertaining to well-being. *The Journal of Positive Psychology*, *11*(5), 546–558.

Mani, N., & Ackermann, L. (2018). Why do children learn the words they do? *Child Development Perspectives*, *12*(4), 253–257. https://doi.org/10.1111/cdep.12295

Martin, J., Rychlowska, M., Wood, A., & Niedenthal, P. M. (2017). Smiles as multipurpose social signals. *Trends in Cognitive Sciences*, *21*(11), 864–877. https://doi.org/10.1016/j.tics.2017.08.007

Matsumoto, D. (1993). Ethnic differences in affect intensity, emotion judgments, display rule attitudes, and self-reported emotional expression in an American sample. *Motivation and Emotion*, *17*(2), 107–123. https://doi.org/10.1007/BF00995188

Maurer, D., & Werker, J. F. (2014). Perceptual narrowing during infancy: A comparison of language and faces. *Developmental Psychobiology*, *56*(2), 154–178. https://doi.org/10.1002/dev.21177

Maye, J., Werker, J. F., & Gerken, L. A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), 101–111. https://doi.org/10.1016/S0010-0277(01)00157-3

McMillan, B. T. M., & Saffran, J. R. (2016). Learning in complex environments: The effects of background speech on early word learning. *Child Development*, *87*(6), 1841–1855. https://doi.org/10.1111/cdev.12559

Mermier, J., Quadrelli, E., Turati, C., & Bulf, H. (2022). Sequential learning of emotional faces is statistical at 12 months of age. *Infancy*. Advance online publication. https://doi.org/10.1111/infa.12463

Morgan, J. L., & Travis, L. L. (1989). Limits on negative information in language input. *Journal of Child Language*, *16*(3), 531–552. https://doi.org/10.1017/S0305000900010709

Moriceau, S., & Sullivan, R. M. (2006). Maternal presence serves as a switch between learning fear and attraction in infancy. *Nature Neuroscience*, *9*(8), 1004–1006. https://doi.org/10.1038/nn1733

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review*, *98*(2), 164–181. https://doi.org/10.1037/0033-295X.98.2.164

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, *14*(1), 11–28. https://doi.org/10.1016/0364-0213(90)90024-Q

Pascalis, O., Loevenbruck, H., Quinn, P. C., Kandel, S., Tanaka, J. W., & Lee, K. (2014). On the links among face processing, language processing, and narrowing during development. *Child Development Perspectives*, *8*(2), 65–70. https://doi.org/10.1111/cdep.12064

Pearl, L. (2021). Theory and predictions for the development of morphology and syntax: A Universal Grammar + statistics approach. *Journal of Child Language*, *48*(5), 907–936. https://doi.org/10.1017/S0305000920000665

Perfors, A., Tenenbaum, J. B., & Regier, T. (2011). The learnability of abstract syntactic principles. *Cognition*, *118*(3), 306–338. https://doi.org/10.1016/j.cognition.2010.11.001

Perszyk, D. R., & Waxman, S. R. (2018). Linking language and cognition in infancy. *Annual Review of Psychology*, *69*(1), 231–250. https://doi.org/10.1146/annurev-psych-122216-011701

Plate, R. C., Bloomberg, Z., Bolt, D. M., Bechner, A. M., Roeber, B. J., & Pollak, S. D. (2019). Abused children experience high anger exposure. *Frontiers in Psychology*, *10*, 440. https://doi.org/10.3389/fpsyg.2019.00440

Plate, R. C., Wood, A., Woodard, K., & Pollak, S. D. (2019). Probabilistic learning of emotion categories. *Journal of Experimental Psychology: General*, *148*(10), 1814–1827. https://doi.org/10.1037/xge0000529

Quam, C., & Swingley, D. (2010). Phonological knowledge guides 2-year-olds' and adults' interpretation of salient pitch contours in word learning. *Journal of Memory and Language*, *62*(2), 135–150. https://doi.org/10.1016/j.jml.2009.09.003

Quine, W. V. O. (1960). *Word* and object. Cambridge, MA: MIT Press.

Raviv, L., & Arnon, I. (2018). The developmental trajectory of children's auditory and visual statistical learning abilities: Modality-based differences in the effect of age. *Developmental Science*, *21*(4), 1–13. https://doi.org/10.1111/desc.12593

Raz, G., & Saxe, R. (2020). Learning in Infancy Is Active, Endogenously Motivated, and Depends on the Prefrontal Cortices. *Annual Review of Developmental Psychology*, *2*(1), 247–268. https://doi.org/10.1146/annurev-devpsych-121318-084841

Repacholi, B. M., Meltzoff, A. N., Hennings, T. M., & Ruba, A. L. (2016). Transfer of social learning across contexts: Exploring infants' attribution of trait-like emotions to adults. *Infancy*, *21*(6), 785–806. https://doi.org/10.1111/infa.12136

Repacholi, B. M., Meltzoff, A. N., Toub, T. S., & Ruba, A. L. (2016). Infants' generalizations about other people's emotions: Foundations for trait-like attributions. *Developmental Psychology*, *52*(3), 364–378. https://doi.org/10.1037/dev0000097

Reuter, T., Borovsky, A., & Lew-Williams, C. (2019). Predict and redirect: Prediction errors support children's word learning. *Developmental Psychology*, *55*(8), 1656–1665. https://doi.org/10.1037/dev0000754

Romberg, A. R., & Saffran, J. R. (2013). Expectancy learning from probabilistic input by infants. *Frontiers in Psychology*, *3*, 1–16. https://doi.org/10.3389/fpsyg.2012.00610

Ruba, A. L., Meltzoff, A. N., & Repacholi, B. M. (2019). How do you feel? Preverbal infants match negative emotions to events. *Developmental Psychology*, *55*(6), 1138–1149. https://doi.org/10.1037/dev0000711

Ruba, A. L., Meltzoff, A. N., & Repacholi, B. M. (2020a). The development of negative event-emotion matching in infancy: Implications for theories in affective science. *Affective Science*, *11*, 4–19. https://doi.org/10.1007/s42761-020-00005-x

Ruba, A. L., Meltzoff, A. N., & Repacholi, B. M. (2020b). Superordinate categorization of negative facial expressions in infancy: The influence of labels. *Developmental Psychology*, *56*(4), 671–685. https://doi.org/10.1037/dev0000892

Ruba, A. L., & Pollak, S. D. (2020). The development of emotion reasoning in infancy and early childhood. *Annual Review of Developmental Psychology*, *2*, 503–531. https://doi.org/10.1146/annurev-devpsych-060320-102556

Ruba, A. L., & Repacholi, B. M. (2020a). Do preverbal infants understand discrete facial expressions of emotion? *Emotion Review*, *12*(4), 235–250. https://doi.org/10.1177/1754073919871098

Ruba, A. L., & Repacholi, B. M. (2020b). Beyond language in infant emotion concept development. *Emotion Review*, *12*(4), 255–258. https://doi.org/10.1177/1754073920931574

Saffran, J. R. (2001). The use of predictive dependencies in language learning. *Journal of Memory and Language*, *44*(4), 493–515. https://doi.org/10.1006/jmla.2000.2759

Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, *47*(1), 172–196. https://doi.org/10.1006/jmla.2001.2839

Saffran, J. R. (2020). Statistical language learning in infancy. *Child Development Perspectives*, *14*(1), 49–54. https://doi.org/10.1111/cdep.12355

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

Saffran, J. R., Hauser, M., Seibel, R., Kapfhamer, J., Tsao, F., & Cushman, F. (2008). Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition*, *107*(2), 479–500. https://doi.org/10.1016/j.cognition.2007.10.010

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, *69*(1), 181–203. https://doi.org/10.1146/annurev-psych-122216-011805

Saffran, J. R., & Wilson, D. P. (2003). From syllables to syntax: Multilevel statistical learning by 12-month-old infants. *Infancy*, *4*(2), 273–284. https://doi.org/10.1207/S15327078IN0402_07

Santolin, C., & Saffran, J. R. (2018). Constraints on statistical learning across species. *Trends in Cognitive Sciences*, *22*(1), 52–63. https://doi.org/10.1016/j.tics.2017.10.003

Santolin, C., & Saffran, J. R. (2019). Non-Linguistic Grammar Learning by 12-Month-Old Infants: Evidence for Constraints on Learning. *Journal of Cognition and Development*, *20*(3), 433–441. https://doi.org/10.1080/15248372.2019.1604525

Schachner, A., & Hannon, E. E. (2011). Infant-directed speech drives social preferences in 5-month-old infants. *Developmental Psychology*, *47*(1), 19–25. https://doi.org/10.1037/a0020740

Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017). Complementary learning systems within the hippocampus: A neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160049.

Seyfarth, R. M., & Cheney, D. L. (2003). Signalers and receivers in animal communication. *Annual Review of Psychology*, *54*, 145–173. https://doi.org/10.1146/annurev.psych.54.101601.145121

Shackman, J. E., & Pollak, S. D. (2005). Experiential influences on multimodal perception of emotion. *Child Development*, *76*(5), 1116–1126. https://doi.org/10.1111/j.1467-8624.2005.00901.x

Shi, R., Legrand, C., & Brandenberger, A. (2020). Toddlers track hierarchical structure dependence. *Language Acquisition*, *27*(4), 397–409. https://doi.org/10.1080/10489223.2020.1776010

Smith, K. E., Leitzke, B. T., & Pollak, S. D. (2020). Youths' processing of emotion information: Responses to chronic and video-based laboratory stress. *Psychoneuroendocrinology*, *122*(February), 104873. https://doi.org/10.1016/j.psyneuen.2020.104873

Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences*, *22*(4), 325–336. https://doi.org/10.1016/j.tics.2018.02.004

Smith, L. B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

Sorce, J. F., Emde, R. N., Campos, J. J., & Klinnert, M. D. (1985). Maternal emotional signaling: Its effect on the visual cliff behavior of 1-year-olds. *Developmental Psychology*, *21*(1), 195–200. https://doi.org/10.1037/0012-1649.21.1.195

Teinonen, T., Fellman, V., Näätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience*, *10*. https://doi.org/10.1186/1471-2202-10-21

Thiessen, E. D. (2017). What's statistical about learning? Insights from modelling statistical learning as a set of memory processes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160056.

Thorstenson, C. A. (2018). The Social Psychophysics of Human Face Color: Review and Recommendations. *Social Cognition*, *36*(2), 247–273. https://doi.org/10.1521/soco.2018.36.2.247

Thorstenson, C. A., McPhetres, J., Pazda, A. D., & Young, S. G. (2021). The role of facial coloration in emotion disambiguation. *Emotion*. Advance online publication. https://doi.org/10.1037/emo0000900

Tong, Y., Wang, F., & Danovitch, J. (2020). The role of epistemic and social characteristics in children's selective trust: Three meta-analyses. *Developmental Science*, *23*(2), 1–18. https://doi.org/10.1111/desc.12895

Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science*, *11*(3), 188–195. https://doi.org/10.1111/1467-9280.00240

Vaillant-Molina, M., Bahrick, L. E., & Flom, R. (2013). Young infants match facial and vocal emotional expressions of other infants. *Infancy*, *18*, 97–111. https://doi.org/10.1111/infa.12017

Vaish, A., Grossmann, T., & Woodward, A. L. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, *134*(3), 383–403. https://doi.org/10.1037/0033-2909.134.3.383

Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The tuning of human neonates' preference for speech. *Child Development*, *81*(2), 517–527. https://doi.org/10.1111/j.1467-8624.2009.01412.x

Vouloumanos, A., & Werker, J. F. (2009). Infants' learning of novel words in a stochastic environment. *Developmental Psychology*, *45*(6), 1611–1617. https://doi.org/10.1037/a0016134

Werker, J. F., & Tees, R. C. (2002). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *25*(1), 121–133. https://doi.org/10.1016/S0163-6383(02)00093-0

Wierzbicka, A. (1992). Talking about emotions: Semantics, culture, and cognition. *Cognition & Emotion*, *6*(3–4), 285–319.

Wilson, B., Spierings, M., Ravignani, A., Mueller, J. L., Mintz, T. H., Wijnen, F., … Rey, A. (2020). Non-adjacent dependency learning in humans and other animals. *Topics in Cognitive Science*, *12*(3), 843–858. https://doi.org/10.1111/tops.12381

Witkower, Z., Hill, A. K., Koster, J., & Tracy, J. L. (2021). Beyond face value: Evidence for the universality of bodily expressions of emotion. *Affective Science*, *2*, 221–229. https://doi.org/10.1007/s42761-021-00052-y

Woodard, K., Plate, R. C., Morningstar, M., Wood, A., & Pollak, S. D. (2021). Categorization of vocal emotion cues depends on distributions of input. *Affective Science*, *2*, 301–310. https://doi.org/10.1007/s42761-021-00038-w

Wu, Y., Schulz, L., Frank, M., & Gweon, H. (2021). Emotion as information in early social learning. *Current Directions in Psychological Science*, *30*(6), 468–475 1–23.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, *114*(2), 245

Zettersten, M., & Saffran, J. R. (2021). Sampling to learn words: Adults and children sample words that reduce referential ambiguity. *Developmental Science*, *24*(3), 1–12. https://doi.org/10.1111/desc.13064

Zhang, F., & Emberson, L. L. (2020). Using pupillometry to investigate predictive processes in infancy. *Infancy*, *25*(6), 758–780. https://doi.org/10.1111/infa.12358