CHAPTER 15: INFANT COMMUNICATION

Abstract

From birth, infants experience the bidirectional influences of social and cognitive processing. **A**mple evidence reveals our evolutionarily endowed urge to interact: **b**abies want to be "in touch" with other people, and this honed instinct has consequences for how they develop. Indeed, acquisition of a range of human abilities is dependent on extensive and extended early social interaction. In this chapter, we examine the early biases that contribute to the development of socio-cognitive functions, including social attention, verbal and nonverbal communication, and full-blown language. A critical component of healthy child development is regular communication with others. Thus, infants' engagement in communication has implications for their acquisition of a range of socio-cognitive abilities. Not surprisingly, typical (and atypical) infant communication is also predictive of developmental outcomes. Thus, the goal of probing how infants communicate is to understand how they perceive and interpret the actions of others in relation to themselves, processes that have enormous implications for future developmental outcomes.

Infant Communication

What are the building blocks underlying language comprehension and production? Some would say that the answer is infants' receptivity to language input; others would focus on the receptivity of their neural tissue. The reality is: The act of processing input shapes the tissue itself. From birth, social and cognitive processes influence one another. Nowhere is this more evident than in the case of language learning. Humans are social animals, and the emergence of our cognitive abilities are dependent on the language we learn through extensive social interaction. The language skills and abilities that infants accrue contribute to how they will interact with people as they continue to develop.

Here, we focus on infant communication and how it lays the groundwork for language learning, including the early social biases that guide this learning. Evidence from a range of sources reveals our evolutionarily endowed urge to interact with others. Babies and young

Bortfeld, H., & Gabouer, A. (2021). Infant communication (chapter 15). In A. Vangelisti (Ed.), *Routledge Handbook of Family Communication* (3rd Edition). Routledge.

children are motivated to be "in touch" with other people, and this honed instinct has major developmental consequences. Infants elicit input and feedback from caregivers, who provide the raw material for early language development, including speech sound discrimination and initial sound production. Healthy social development helps infants and young children communicate with others, which in turn helps them learn to communicate more and better. Not surprisingly then, typical and atypical early social cognition are predictive of later learning and development, including of language.

Building Blocks of Language Comprehension and Production

Language shapes our social interaction, engaging almost every area of the brain. In the case of spoken language, producing speech requires specialized neural circuits for controlling the muscles of the larynx, tongue, mouth, and lips; understanding speech requires circuits for parsing a dense incoming stream of auditory input and assigning meaning to its component parts. On top of that, we identify and interpret facial expressions, body language, and tone of voice, distinguishing between sentiments as distinct as sincerity and sarcasm. These pragmatic processes must be integrated with the syntactic and semantic processes in some way as well. In short, communicating via language requires a brain that allows formulation of ideas well enough to express them in such a way that others will understand. The journey from being a preverbal infant to being a mature speaker/listener seems should be a long one.

Auditory Input Interacts with Developing Brain Structure

A key question to ask about language development is to what degree the neonatal brain is preconfigured to learn language, and to what degree early—even prenatal—input prepares it for the learning process? A purely non-nativist view favors bottom-up influences on language development, according to which sensory input shapes a plastic brain; a purely nativist view favors a black-box, modular characterization of how infants come to handle the complex, combinatorial features of language as quickly as they do. Although the 'nature versus nurture' debate presents a false dichotomy, it remains informative as we try pull apart the intertwined threads.

The first issue is to understand that infants' acoustic exposure begins in the womb. The tissue and liquid barriers protecting the developing fetus serve as low-pass filters, absorbing frequencies greater than 5000 Hz (Jardri et al., 2008) much like apartment walls absorb high frequency sounds and transmit low frequency sounds. This means that phoneme-specific information, such as that distinguishing one speech-specific sound from another, is not experienced by the developing fetus. Prosody, on the other hand, conveys the overall shape of speech to the developing fetus. Researchers have used both behavioral and physiological measures (i.e., fetal movement, heart rate) to extrapolate information about when fetal auditory processing begins and what sorts of auditory distinctions the developing fetus can make. For example, researchers have determined that sound is processed as early as 22 gestational weeks (Hepper & Shahidullah, 1994; Lecanuet et al., 1995). Fetal habituation, a metric of encoding or memory, takes place in response to repeated sounds as early as 32 gestational weeks (Morokuma et al., 2004). Closer to term, infants' sensitivity to more complex sounds increases, allowing them to distinguish between variations in music (Kisilevsky et al., 2004) and to differentiate the prosodic structures of familiar and novel rhymes (Decasper et al., 1994).

Remarkably, these fetal sensitivities were established while the child is still *in utero*! For example, DeCasper and colleagues (1994) used measures of fetal heartrate changes to familiar and unfamiliar rhymes to determine what the fetus was hearing. But neonates likewise demonstrate what they have learned while in utero by showing auditory preferences immediately following birth. In a seminal study, newborns learned to elicit playback of their mother's voice or the voice of an unfamiliar female by sucking on a nonnutritive nipple in different ways. They demonstrated a clear preference for their mother's voice over the unfamiliar voice, despite having limited postnatal experience (DeCasper & Fifer, 1980). Likewise, newborns demonstrated a preference for a story to which they were exposed repeatedly—Dr. Seuss's *The Cat in the Hat*—over a novel story in the same language (DeCasper & Spence, 1986).

Additional findings abound: Newborns can discriminate speech from non-speech when it is played forwards, but not backwards (Ramus et al., 2000); they prefer the language of their own environment over another, unfamiliar language (Moon et al., 1993); they can distinguish between stress patterns of different multisyllabic words (Moon et al., 1993); and they can categorically discriminate lexical versus grammatical words (Shi et al., 1999). Newborns can distinguish between two rhythmically dissimilar languages (Mehler et al., 1988); are sensitive to word boundaries (Christophe et al., 1994); and can differentiate between good and poor syllable forms (Bertoncini & Mehler, 1981).

Of course, when we say that newborns prefer a stimulus like their native language over another with a different rhythmic structure, this is probably not *only* due to prenatal exposure, as they have likely experienced some early postnatal exposure. The constraints of doing research with newborns means that they may be a few days old by the time they are tested. Nonetheless, early perceptual abilities are not due to any "genetic" transmission of the language the caregivers speak. Rather, infants are primed for the full-blown version of the language they've been hearing in low fidelity form for the second half of their prenatal period. Consistent with this, infants in bilingual households who have been exposed to two languages prenatally do not show differential preference for them when tested as newborns, despite being capable of discriminating between the two (Byers-Heinlein et al., 2010).

Emerging Sensitivity to Language-Specific Structure

Language learning involves learning a language's phonemes, from which an infant can begin to develop the phonological representations critical to recognizing the sound structure of individual words. What phonemes are used in any given language has been influenced by the more general way that acoustic continua produced by the vocal apparatus can be broken up. For example, the voicing continuum (i.e., timing of vibrations of the vocal folds, voice onset time VOT) can be divided into specific perceptual categories, such as the English minimal pairs /b/ and /p/. Such sound pairs are perceived categorically (i.e., as one sound or the other with no in between) by adult native speakers, and are used contrastively to denote different meanings (e.g., bat vs. pat). Different languages parcellate voicing (and other) continua differently (Cho et al., 2019).

How sensitive are newborns to their own and other languages' phonetic categories? Many of the acoustic differences are minute (e.g., 20 ms difference in VOT turns /b/ into /p/), yet young infants are exquisitely sensitive to the acoustic changes that mark phonetic boundaries. In a ground breaking study, Peter Eimas and colleagues (Eimas et al., 1971) found that when 1-to 4-month-old English exposed infants were habituated to one of two synthetic English syllables, /ba/ or /pa/, differing only on VOT, they dishabituated upon exposure to the other one. From this, one might conclude that language-specific auditory sensitivities are innate to humans. However, the ability to categorically perceive such sounds was shown to not be specific to humans, as it was observed in crickets (Wyttenbach et al., 1996), chinchillas (Kuhl & Miller, 1975) and macaques (Kuhl & Padden, 1983). In other words, human phonetic categories evolved from constraints imposed by general characteristics of the sensory system rather than ones unique to human ears and brains (Kuhl, 1986; Smith & Lewicki, 2006).

In another set of classic studies, Janet Werker and Richard Tees (Werker et al., 1981; Werker & Tees, 1984) demonstrated that 6-month-old infants were able to categorically perceive phonemes in languages they'd never before been exposed to, while adult monolingual English speakers were oblivious to these contrasts. Based on this remarkable finding, it has since been demonstrated that infants transition from a broad ability to distinguish many/most minimally contrastive phoneme pairs to an increasingly specialized capacity to distinguish primarily the phoneme pairs relevant to their ambient language(s). This perceptual tuning takes place in the first postnatal year, meaning that by between 10 and 12 months of age, infants are more adept at discriminating native compared with nonnative contrasts. In other words, the ability to discriminate nonnative phonetic contrasts declines across the second half of the first year of life. Thus, while both monolingual English- and Japanese-exposed infants are able to make the English /r/ and /l/ distinction prior to 6 months of age, by the end of their first year, only the monolingual English-exposed infants will continue to differentiate these sounds (Kuhl et al., 2006), because the /r/ vs. /l/ distinction is not used in the Japanese language and the neural process of perceptual tuning renders it increasingly imperceptible to a Japanese-exposed infant.

How does this tuning happen? Although perceptual learning is a lifelong ability, developmental critical periods like the one for native language minimal pairs have a constrained timeline, whereby functional brain circuits are shaped in an experience-dependent manner. Other examples of the emergence or maintenance of specific abilities abound. For example, studies with people acquiring an initial language post-puberty (e.g., deaf children raised without visual language, so-called feral children) show that acquisition of native syntax, among other language specific abilities, is constrained and that lack of stimulation during the early childhood period has lasting effects on the both the brain and behavior. Thus, deprivation during the early developmental time frame leads to lifelong deficits in language—and likely other—abilities. The importance of early exposure to language, whether spoken or signed, to the brain development of congenitally deaf children is a topic of recent research (Hall et al., 2018a, 2018b), as is making sure that pediatric cochlear implant users are receiving appropriate levels of stimulation (Sevy et al., 2010).

In sum, infants younger than six months can discriminate a wide range of speech sound contrasts, both native and nonnative (see Saffran et al., 2006 for a review). Their sound-specific sensitivity increases with continued exposure to the phonetic contrasts in their environment, while they experience a corresponding loss of sensitivity to sounds from other, nonnative languages (Werker & Desjardin, 1995). Thus, perceptual tuning is best conceptualized as an experience-dependent sharpening of the boundaries between two closely related native speech sounds (Aslin & Pisoni, 1980; Kuhl et al., 2001; Polka et al., 2001) relative to those between nonnative sounds. For nonnative sounds, maturation and lack of experience further blurs the nonnative speech sound boundaries until they longer exist. This matters: children's sensitivity to native relative to nonnative contrasts at the end of the first year positively predicts their language abilities in the second year, whereas maintenance of sensitivity to nonnative contrasts shows a negative relationship with language ability measures in the second year (Kuhl et al., 2005). Finally, these nonnative boundaries can be resharpened is exposure is reintroduced (as when a child begins learning a new language before puberty), but the window is a closing one that becomes permanently shut post-puberty.

Language Engages the Infant Brain beyond the Language Network

Across languages, caregivers exaggerate their speech to infants (Fernald et al., 1989; Fernald & Morikawa, 1993), increasing the pitch of their voice and expanding its range and variability over time. Moreover, caregivers also repeat themselves, a lot. Infants are attracted to these features, attend to speech directed at them, and thereby gain experience from their interactions with people who produce such sing-song and repetitive speech. While the acoustic properties of infant-directed speech appear to underlie its effectiveness in attracting infants' attention, the particular components that drive infants' extended preference are less clear. It would seem that infants develop their preference for this infant-directed speech postnatally (assuming parents only produce it once the child is born), meaning they are not biologically predisposed to exhibit such a preference. Yet at birth infants discriminate between positive and negative emotions (Mastropieri & Turkewitz, 1999), and respond differently to positive and negative emotions as conveyed by tone of voice (Fernald, 1992; M. Papoušek et al., 1990). Of course, people are generally happy when they address infants, so the issues of pitch and affect are tightly intertwined (Kitamura & Burnham, 1998). As it turns out, when affect and pitch are manipulated independently, the positive affective properties override pitch in attracting infants' attention! In other words, infants show no preference for infant- over adult-directed speech when the affect is held consistently positive across both (Singh et al., 2002). Not surprisingly, since "happy talk" draws infants' attention in a positive way, caregivers (and doting others) are more inclined to manipulate their vocal acoustics to elicit this response. Indeed, and also unsurprisingly, adults rate infants' facial responses to infant-directed speech as more "attractive" than their facial responses to adult-directed speech (Werker & McLeod, 1989). Thus, infants' preference for positive emotion, along with adults' inclination to produce happy talk when speaking to infants, is a critical component of their early preference for infant-directed speech.

The combination of infants' early exposure to speech and their precocious preference for faces—the source of spoken language—positions them to learn the relationship between auditory and visual speech (i.e., the visual information provided by the articulating mouth). As a result, perceptual tuning extends to visual speech, as demonstrated by Weikum et al. (2007), who exposed 4-, 6-, and 8-month-old monolingual English infants to either English or French faces producing visual speech (i.e., the faces articulating sentences with the sound turned off). Four- and 6-month-old, but not 8-month-old, monolinguals detected the difference between English and French from visual cues alone. In contrast, 8-month-old bilingual (English-French) infants were able to detect the difference between the two languages.

Infants quickly become attuned to the temporal relationship between auditory and visual speech as well. In an influential early study, Kuhl and Meltzoff (1982) observed that when hearing /a/ vowels, 4-month-old infants looked longer at a face articulating /a/ than at a simultaneously articulating face producing /i/, a finding that was later replicated and extended to infants as young as 2 months of age (Patterson & Werker, 1999, 2003). Similarly, 8-month-olds prefer to look at a gender congruent visual display relative to an incongruent one when hearing a gender specific voice (Patterson & Werker, 2002). Five- to 15-month-olds look longer at a visual speech stream that matches a three syllable nonword with degraded phonetic detail (Baart et al., 2014), and 4- and 6-month-olds even prefer to look at a cooing monkey face over a monkey face producing a grunt if they are simultaneously hearing the cooing sound (Lewkowicz & Ghazanfar, 2006).

Are infants sensitive to more extended temporal relationships between auditory and visual speech? Lewkowicz and Pons (2013) demonstrated that they are, finding that by between 10 and 12 months of age, infants' representation of speech entails multiple modalities. First, the researchers familiarized infants with either their native (English) or a nonnative (Spanish) continuous auditory speech stream, after which the infants were presented with two simultaneous (silent) videos of a face producing visual speech. In one video, infants saw the face articulating the previously heard string of speech while in the other, a semantically matched version (i.e., a translation) of the same string of speech was articulated in the opposite language. Ten- to 12-month-olds looked less at their native visual speech when they were previously familiarized to auditory speech in that language, most likely because they had detected the correspondence between sight and sound and preferred the novel visual speech (i.e., they showed a novelty effect). By the end of the first year, infants have associated the auditory and visual forms with their ambient language. A similar study found the same effect in six-month-olds (Kubicek et al., 2014). Thus, from birth, infants use multiple sources of information to learn about communication, and they do so in increasingly complex ways as they consolidate experience across the first year of life.

Is it Imitation, or an Urge to Interact?

Thus far, we have focused on speech-specific aspects of infant language learning. However, several other factors contribute to the process. Fundamentally, infants are wired to connect with other people. In particular face recognition is a foundational component of typically developing infants' communication toolkit. Faces convey information of profound social significance, including a person's identity and emotional state, and perceiving faces quickly becomes one of children's most developed skills. A host of findings have converged on the relevant perceptual mechanisms supporting face recognition and the neural mechanisms underlying this ability. While infants may initially engage in social interactions for seemingly simple reasons, such as maintaining proximity to their caregiver (Gergely, 2001; 2010), a range of mechanisms are in play early on to ensure that infants continue to elicit attention from others. This is largely focused on others' faces, with infants striving to imitate others facial expressions. The process of imitation involves the gradual coordination of several emerging systems—social, cognitive and motor—across the first two years of life and contributes enormously to an infant's ability to communicate.

In a now classic study, newborn infants imitated an adult model's facial gesturing, including both tongue protrusion and mouth opening (Meltzoff & Moore, 1983). Although methodological aspects of these demonstrations are nuanced, there has been sufficient replication over the years (Butterworth, 1999; Meltzoff, 2005) to suggest that infants model others' gestures early in life: they imitate others. Although questions remain about whether this is truly imitative or rather reflects the triggering of a motor schema or action pattern available in a preexisting motion repertoire (Jones, 2017), imitative-like behavior appears to be a critical entry point for infants into the social world around them.

Language Learning is Social Learning

In an influential study, Patricia Kuhl and her colleagues (2003) demonstrated that human's capacity to learn language is influenced by interacting with other people. This was achieved by comparing the relative loss/maintenance of categorical perception in 9-month-old English learning infants as a product of their immediate social environment. Across two experiments, infants were trained on a Mandarin Chinese contrast that is undetectable by native English speakers. In the first study the infants were trained on these contrasts during play sessions with adult Mandarin speakers; control infants experienced the same play sessions, but in English. At the end of the training, the experimental group of infants were able to distinguish the speech contrast, performing at a level comparable to Mandarin learning infants in Taiwan. The control group did not. In a second study, the same training sessions took place with a new group of infants, but this time this was presented by video. Again, half got the training in Mandarin and half in English. Despite showing high engagement with the videos, there was no discrimination after training by either the experimental or control groups, highlighting the importance of engaged and contingent interaction for language learning to take place. In short, language is social, and language learning takes place in responsive social exchanges.

Caregivers elicit speech from their infants and young children by being sensitive to their abilities, including responding to their speaking attempts in a supportive and contingent manner. Indeed, learning is most efficient at promoting language development when caregivers calibrate their own speech to be just challenging enough for their child. For adults to maintain speech in this range, they must be in touch with a child's rapidly changing abilities. Thus, a caregiver's own frequent exposure to the child's language (e.g., through active conversation) helps guide appropriate tuning to the child's specific developmental level (Zimmerman et al., 2009). Prime opportunities for language learning occur when adults focus on and talk about things that are relevant to an infant's own attentional focus. Caregivers who are responsive to what their infants are engaged by may specifically support advances in language development by providing labels for objects and events when they are receiving attention from both the caregiver and the child. This joint attention eases the challenge for the child to match linguistic symbols to their referents, a process that reinforces the social-communicative function of language itself (Tomasello & Farrar, 1986).

Contingencies Model Coordinated Communication

Contingency-seeking in infants was demonstrated in a seminal study (Watson & Ramey, 1972) in which 2-month-olds increased their leg kicking when it resulted in a highly contingent stimulus, such as making a mobile rotate, but not when the event was noncontingent on leg-kicking. Infants prefer a certain level of contingency (Bahrick & Watson, 1985); they also produce a particular amount of contingency in their own behavior. Indeed, parent-infant dyads develop contingency in their interactions, with one member of the dyad closely following the other's behavior with similar behavior. By the second month, infants convey positive affect during live interactions with their mothers that contain variable contingency, whereas they actually show *negative* affect when they are shown replays of their mothers from these interactions in which their mother necessarily is acting noncontingently (Murray & Trevarthen, 1985). This is in spite of these replays having just as much emotion and social stimulation as the live interaction!

In a three-part (live-replay-live) paradigm, mothers interacted directly and contingently with their 9-week-old infants both before and after infants were exposed to a replay of the mother. In the replay, the mother similarly engaging as in the two live sessions, but because the replay condition was prerecorded, she was not acting contingently in response to her infant. Based on measures of affect and engagement (i.e., smiling, gazing at mother, mouth closed, frowning, infants were found to react more negatively in the replay condition than in either of the live conditions (Nadel et al., 1999). Indeed, in the second live condition, infants' positivity returned to levels observed in the first live condition, demonstrating that the change in behavior in the replay condition was not due to a natural decline in infant engagement, but rather to the lack of contingency between the mother and infant. In short, contingency helps infants establish correspondence between themselves and others, thereby encouraging development of the representation of oneself as an intentional agent in the world.

This contingency-based coordination process starts early. In a clever experiment, Michael Goldstein and his colleagues tested the importance of contingency by telling caregivers precisely *when* to respond to infant vocalizations (Goldstein et al., 2003). In half of the infant-caregiver pairs, mothers were trained to respond contingently to their infants' vocalizations by producing nonvocal social responses (i.e., smiling, touching). Mothers in the other half of the infant-caregiver pairs were instructed to respond to their infants based on the response schedules of the mothers in the contingent group, thus noncontingently. Infants who received social feedback that was contingent on their vocalizations produced more developmentally advanced vocalizations—both during the manipulation and after maternal responding was no longer being manipulated—relative to those infants who received feedback independent of when they vocalized. These results thus demonstrate that social interaction is a critical link between early speech perception and subsequent production. Infants who learn the contingency between their own vocalizations and the responses of their caregivers thus learn to influence the behavior of social partners, an important step forward in early communicative development. When caregivers respond contingently to infants' vocalizations with their own speech, infants structure their sounds to match those phonological patterns (e.g., vowel sounds elicit more vowel sounds from infants and words elicit more consonant-vowel combinations) (Goldstein & Schwade, 2008).

Goldstein and colleagues (Elmlinger et al., 2019) recently observed that caregivers simplify the statistical and syntactic structure of their speech in response to their infants' babbling. Their contingent speech contained fewer unique words and contained both shorter utterances and more single-word utterances than non-contingent speech. Thus, infants' immature vocalizations create opportunities for them to learn by eliciting simplified responses from parents that contain simplified, more learnable information. Moreover, caregivers who use more linguistically diverse contingent speech with their infants have infants with more advanced vocalizations. Thus, a functional perspective has emerged whereby infant language learning regulates and is regulated by social interactions with adults. This early communication is foundational to the development of complex language going forward.

The Importance of Both Quantity and Quality in Speech Input

We have argued that infants' fundamentally social motivation to communicate provides them with the opportunity for language learning. While this is by now supported by extensive research, it is also true that language exposure is not uniform across all children. In a seminal study, Hart and Risley (1995) demonstrated that the raw number of words children hear varies enormously as a function of a family's socioeconomic status (SES), with average income families producing up to double the number of words as is produced by lower income families. These researchers made the (then provocative) suggestion that such differences in frequency of exposure might underlie the reliable differences in language outcomes observed as children from these families enter and proceed through formal education. There is now substantial evidence that early differences in both the amount and quality of speech children hear influences their emerging language abilities. The amount of language that children experience before the age of three is positively correlated with their language production skills and cognitive development more generally (e.g., Arterberry et al., 2007; Bornstein & Haynes, 1998; Huttenlocher, 1998; Huttenlocher et al., 1991; Rowe et al., 2005; Shonkoff & Phillips, 2000). The quality of that language exposure, including both linguistic (i.e., vocabulary diversity and sophistication, grammatical complexity, and narrative use (Rowe, 2012), as well as interactional features (i.e., contiguous (time-locked), contingent (topically similar), back-and-forth conversation (Hirsh-Pasek et al., 2015) also matters. Fortunately, caregiver-child conversational turn-taking involves a rich experience of high-quality linguistic, attentional, and social features, but the quality and quantity of speech input is often mediated by socioeconomic status (Hoff, 2003).

There is a broad relationship between SES and young children's brain structure and function (for review, see Farah, 2017). More recently, there has been increasing focus on establishing mechanistic knowledge about specific environmental factors associated with specific variation in brain structure. There is now some evidence that certain aspects of children's language environments relate to functional brain responses in prefrontal cortical regions (Garcia-Sierra et al., 2016; Sheridan et al., 2012). We know there is substantial variation in the amount of language children experience, in part due to variable SES (Gilkerson et al., 2017). Do such differences in language exposure relate to differences in neuroanatomy? By relating young children's real world language exposure between the ages of 4 and 6 years to their neuroanatomy, Romeo and colleagues (2018) were able to confirm that greater adult-child conversational experience is related to stronger, more coherent white matter connectivity in white matter tracts relevant to language. In other words, how the children's brains were developing was a product of their prior language experiences. Interestingly, these findings were independent of SES and of the sheer quantity of adult speech. Rather, microstructural analyses of the white matter were found to mediate the relationship between conversational turns and children's language skills, indicating a neuroanatomical mechanism underlying the so-called language gap. In short, quality matters more than quantity to children's emerging language abilities, and is reflected neuroanatomically as well.

Conclusion

Communication is inherently social. At the earliest stages of development, infants are shaped by the sounds around them. Subsequently, caregivers' biases to communicate in particular ways help infants focus their attention specifically on speech sounds. The inherent structure of the speech signal together with the contingent structure of the infant-caregiver interaction serve to highlight regularities in speech and in interactive form; infants respond to this, as reflected in their subsequent productions of new vocal forms. Particular maternal responses, such as imitations and expansions, correlate positively with language development. Through these responses, infants learn the association between the production of particular sounds and their outcomes. Finally, caregivers' input during social interactions and early "conversations" scaffold language learning by providing information about activities and objects that are the focus of infants' attention in the first place. In sum, socially guided communication is fundamental to infants' initial language development, and lays the foundation for subsequent advances in language learning and social development alike.

References

Adamson, L. B., Bakeman, R., & Deckner, D. F. (2004). The development of symbol-infused joint engagement. *Child Development*, *75*(4), 1171–1187. https://doi.org/10.1111/j.1467-8624.2004.00732.x

Aktar, E., Qu, J., Lawrence, P. J., Tollenaar, M. S., Elzinga, B. M., & Bögels, S. M. (2019). Fetal and infant outcomes in the offspring of parents with perinatal mental disorders: Earliest influences. *Frontiers in Psychiatry*, *10*, 391. https://doi.org/10.3389/fpsyt.2019.00391

Arterberry, M. E., Cain, K. M., & Chopko, S. A. (2007). Collaborative problem solving in five-year-old children: Evidence of social facilitation and social loafing. *Educational Psychology*, *27*(5), 577–596. https://doi.org/10.1080/01443410701308755

Aslin, R. N., & Pisoni, D. B. (1980). Effects of early linguistic experience on speech discrimination by infants: A critique of Eiler, Gavin, and Wilson (1979). *Child Development*, *51*(1), 107–112. https://doi.org/10.1111/j.1467-8624.1980.tb02516.x

Baart, M., Vroomen, J., Shaw, K., & Bortfeld, H. (2014). Degrading phonetic information affects matching of audiovisual speech in adults, but not in infants. *Cognition*, *130*(1), 31–43. https://doi.org/10.1016/j.cognition.2013.09.006

Bahrick, L. E., & Watson, J. S. (1985). Detection of intermodal proprioceptive-visual contingency as a potential basis of self-perception in infancy. *Developmental Psychology*, *21*(6), 963–973. https://doi.org/10.1037/0012-1649.21.6.963

Bertoncini, J., & Mehler, J. (1981). Syllables as units in infant speech perception. *Infant Behavior and Development*, *4*(1), 247–260. https://doi.org/10.1016/S0163-6383(81)80027-6

Bornstein, M. H., & Haynes, O. M. (1998). Vocabulary competence in early childhood: Measurement, latent construct, and predictive validity. *Child Development*, *69*(3), 654–671. https://doi.org/10.2307/1132196

Butterfield, E. C., & Siperstein, G. N. (1970). Influence of contingent auditory stimulation upon non-nutritional suckle. *Third Symposium on Oral Sensation and Perception: The Mouth of the Infant*, 313–334.

Butterworth, G. (1999). Neonatal imitation: Existence, mechanisms and motives. In J. Nadel & G. Butterworth (Eds.), *Imitation in infancy*. Cambridge University Press.

Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science*, *21*(3), 343–348. https://doi.org/10.1177/0956797609360758

Campos, J. J., & Stenberg, C. R. (1981). Perception, appraisal, and emotion: The onset of social referencing. In M. E. Lamb & L. R. Sherrod (Eds.), *Infants social cognition: Empirical and social considerations* (pp. 273–314). Lawrence Erlbaum.

Carpenter, M., Nagell, K., Tomasello, M., & Butterworth, G. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*, 1–174. https://doi.org/10.2307/1166214

Cho, T., Whalen, D. H., & Docherty, G. (2019). Voice onset time and beyond: Exploring laryngeal contrast in 19 languages. *Journal of Phonetics*, *72*, 52–65.

https://doi.org/10.1016/j.wocn.2018.11.002

Chomsky, N. (1966/2009). *Cartesian linguistics: A chapter in the history of rationalist thought*. Cambridge University Press.

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America*, *95*(3), 1570–1580. https://doi.org/10.1121/1.408544

Colombo, J., & Bundy, R. S. (1981). A method for the measurement of infant auditory selectivity. *Infant Behavior and Development*, *4*(1), 219–223. https://doi.org/10.1016/S0163-6383(81)80025-2

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, *208*(4448), 1174–1176.

DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development*, *9*(2), 133–150. https://doi.org/10.1016/0163-6383(86)90025-1

DeCasper, A., Lecanuet, J., Busnel, M., Granier-Deferre, C., & Maugeais, R. (1994). Fetal reactions to recurrent maternal speech. *Infant Behavior and Development*, *17*(2), 159–164. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.693.1071&rep=rep1&type=pdf

Delgado, C. E. F., Mundy, P., Crowson, M., Markus, J., Yale, M., & Schwartz, H. (2002). Responding to joint attention and language development: A comparison of target locations. *Journal of Speech, Language, and Hearing Research*, *45*(4), 715–719.

https://doi.org/10.1044/1092-4388(2002/057)

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*(3968), 303–306.

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

Farah, M. J. (2017). The neuroscience of socioeconomic status: Correlates, causes, and consequences. *Neuron*, *96*, 56–71. https://doi.org/10.1016/j.neuron.2017.08.034

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

Fernald, A. (1992). Meaningful melodies in mothers' speech to infants. In H. Papoušek, U. Jurgens, & M. Papoušek (Eds.), *Nonverbal vocal communication: Comparative and developmental aspects* (pp. 262–282). Cambridge University Press.

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development*, *64*(3), 637–656.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., De Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*(3), 477–501. https://doi.org/10.1017/S0305000900010679

Friederici, A. D., & Chomsky, N. (2017). *Language in Our Brain*. The MIT Press.

https://doi.org/10.7551/mitpress/9780262036924.001.0001

Garcia-Sierra, A., Ramírez-Esparza, N., & Kuhl, P. K. (2016). Relationships between quantity of language input and brain responses in bilingual and monolingual infants. *International Journal of Psychophysiology*, *110*, 1–17. https://doi.org/10.1016/j.ijpsycho.2016.10.004

Geangu, E., Ichikawa, H., Lao, J., Kanazawa, S., Yamaguchi, M. K., Caldara, R., & Turati, C. (2016). Culture shapes 7-month-olds' perceptual strategies in discriminating facial expressions of emotion. *Current Biology*, *26*(14), R663–R664. https://doi.org/10.1016/j.cub.2016.05.072

Gergely, G. (2001). The obscure object of desire: "Nearly, but clearly not, like me": Contingency preference in normal children versus children with autism. *Bulletin of the Menninger Clinic*, *65*(3), 411–426. https://doi.org/10.1521/bumc.65.3.411.19853

Gergely, G. (2010). Kinds of agents: The origins of understanding instrumental and communicative agency. In U. Goshwami (Ed.), *The Wiley-Blackwell Handbook of Childhood Cognitive Development* (2nd ed., pp. 76–105). Wiley-Blackwell. https://doi.org/10.1002/9781444325485.ch3

Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Oller, D. K., Hansen, J. H. L., & Paul, T. D. (2017). Mapping the early language environment using all-day recordings and automated analysis. *American Journal of Speech-Language Pathology*, *26*(2), 248–265. https://doi.org/10.1044/2016_AJSLP-15-0169

Gliga, T., & Csibra, G. (2007). Seeing the face through the eyes: A developmental perspective on face expertise. *Progress in Brain Research*, *164*, 323–339. https://doi.org/10.1016/S0079-6123(07)64018-7

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. 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

Hall, M. L., Eigsti, I. M., Bortfeld, H., & Lillo-Martin, D. (2018a). Auditory access, language access, and implicit sequence learning in deaf children. *Developmental Science*, *21*(3). https://doi.org/10.1111/desc.12575

Hall, M. L., Eigsti, I. M., Bortfeld, H., & Lillo-Martin, D. (2018b). Executive function in deaf children: Auditory access and language access. *Journal of Speech, Language, and Hearing Research*, *61*(8), 1970–1988. https://doi.org/10.1044/2018_JSLHR-L-17-0281

Han, S., Northoff, G., Vogeley, K., Wexler, B. E., Kitayama, S., & Varnum, M. E. W. (2013). A Cultural Neuroscience Approach to the Biosocial Nature of the Human Brain. *Annual Review of Psychology*, *64*(1), 335–359. https://doi.org/10.1146/annurev-psych-071112-054629

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Paul H Brookes.

Hepper, P. G., & Shahidullah, B. S. (1994). Development of fetal hearing. *Archives of Disease in Childhood*, *71*(2), 81–87. https://doi.org/10.1136/fn.71.2.f81

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., Yust,

P. K. S., & Suma, K. (2015). The contribution of early communication quality to low-income children's language success. *Psychological Science*, *26*(7), 1071–1083. https://doi.org/10.1177/0956797615581493

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*(5), 1368–1378. https://doi.org/10.1111/1467-8624.00612

Huttenlocher, J. (1998). Language input and language growth. *Preventive Medicine*, *27*(2), 195–199. https://doi.org/10.1006/pmed.1998.0301

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236–248. https://doi.org/10.1037/0012-1649.27.2.236

Hutton, J. S., Dudley, J., Horowitz-Kraus, T., Dewitt, T., & Holland, S. K. (2020). Associations between screen-based media use and brain white matter integrity in preschool-aged children. *JAMA Pediatrics*, *174*(1), e193689. https://doi.org/10.1001/jamapediatrics.2019.3869

Jardri, R., Pins, D., Houfflin-Debarge, V., Chaffiotte, C., Rocourt, N., Pruvo, J. P., Steinling, M., Delion, P., & Thomas, P. (2008). Fetal cortical activation to sound at 33 weeks of gestation: A functional MRI study. *NeuroImage*, *42*(1), 10–18. https://doi.org/10.1016/j.neuroimage.2008.04.247

Jarvis, E. D. (2019). Evolution of vocal learning and spoken language. *Science*, *366*(6461), 50–54. https://doi.org/10.1126/science.aax0287

Jones, S. (2017). Can newborn infants imitate? *Wiley Interdisciplinary Reviews: Cognitive*

Science, *8*(1–2), e1410. https://doi.org/10.1002/wcs.1410

Kelly, D. J., Liu, S., Rodger, H., Miellet, S., Ge, L., & Caldara, R. (2011). Developing cultural differences in face processing. *Developmental Science*, *14*(5), 1176–1184. https://doi.org/10.1111/j.1467-7687.2011.01067.x

Kisilevsky, B. S., Hains, S. M. J., Jacquet, A. Y., Granier-Deferre, C., & Lecanuet, J. P. (2004). Maturation of fetal responses to music. *Developmental Science*, *7*(5), 550–559. https://doi.org/10.1111/j.1467-7687.2004.00379.x

Kitamura, C., & Burnham, D. (1998). The infant's response to maternal vocal affect. *Advances in Infancy Research*, *12*, 221–236.

Kubicek, C., De Boisferon, A. H., Dupierrix, E., Pascalis, O., Loevenbruck, H., Gervain, J., & Schwarzer, G. (2014). Cross-modal matching of audio-visual German and French fluent speech in infancy. *PLoS ONE*, *9*(2), e89275. https://doi.org/10.1371/journal.pone.0089275

Kuhl, P. K. (1986). Theoretical contributions of tests on animals to the special-mechanisms debate in speech. *Experimental Biology*, *45*(3), 233–265.

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the "critical period." *Language Learning and Development*, *1*(3–4), 237–264. https://doi.org/10.1080/15475441.2005.9671948

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, *218*(4577), 1138–1141.

Kuhl, P. K., & Miller, J. D. (1975). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science*, *190*(4209), 69–72. https://doi.org/10.1126/science.1166301

Kuhl, P. K., & Padden, D. M. (1983). Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *Journal of the Acoustical Society of America*, *73*(3), 1003–1010. https://doi.org/10.1121/1.389148

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), 13–21. https://doi.org/10.1111/j.1467-7687.2006.00468.x

Kuhl, P. K., Tsao, F., & Liu, H. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(15), 9096–9101. https://doi.org/10.1073/pnas.1532872100

Kuhl, P. K., Tsao, F., Liu, H., & Zhang, Y. (2001). Language/culture /mind /brain progress at the margins between disciplines. *Annals of the New York Academy of Sciences*, *935*(1), 136–174.

Lecanuet, J., Granier-Deferre, C., & Busnel, M. (1995). Human fetal auditory development. In J. Lecanuet, W. P. Fifer, N. A. Krasnegor, & W. P. Smotherman (Eds.), *Fetal development: A psychobiological perspective* (pp. 239–262). Lawrence Erlbaum.

Lewkowicz, D. J., & Ghazanfar, A. A. (2006). The decline of cross-species intersensory perception in human infants. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(17), 6771–6774. https://doi.org/10.1073/pnas.0602027103

Lewkowicz, D. J., & Pons, F. (2013). Recognition of amodal language identity emerges in infancy. *International Journal of Behavioral Development*, *37*(2), 90–94. https://doi.org/10.1177/0165025412467582

Lordier, L., Meskaldji, D. E., Grouiller, F., Pittet, M. P., Vollenweider, A., Vasung, L., Borradori-Tolsa, C., Lazeyras, F., Grandjean, D., De Ville, D. Van, & Hüppi, P. S. (2019). Music in premature infants enhances high-level cognitive brain networks. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(24), 12103–12108. https://doi.org/10.1073/pnas.1817536116

Mastropieri, D., & Turkewitz, G. (1999). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental Psychobiology: The Journal of the International Society for Psychobiology*, *35*(3), 204–214. https://doi.org/10.1002/(SICI)1098-2302(199911)35:3<204::AID-DEV5>3.0.CO;2-V

Matthews, L. G., Walsh, B. H., Knutsen, C., Neil, J. J., Smyser, C. D., Rogers, C. E., & Inder, T. E. (2018). Brain growth in the NICU: Critical periods of tissue-specific expansion. *Pediatric Research*, *83*(5), 976–981.

McCune, L. (1995). A normative study of representational play at the transition to language. *Developmental Psychology*, *31*(2), 198–206. https://doi.org/10.1037/0012-1649.31.2.198

Mehler, J., Jusczyk, P., Lamsertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*(2), 143–178.

Meltzoff, A. N. (2005). Imitation and other minds: The “like me” hypothesis. In *Perspectives on imitation: From neuroscience to social science* (Vol. 2, pp. 55–77). MIT Press. https://doi.org/10.1111/j.1468-0017.2007.00306.x

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*(4312), 75–78. https://doi.org/10.1126/science.198.4312.75

Meltzoff, A. N., & Moore, M. K. (1983). Newborn infants imitate adult facial gestures. *Child*

Development, *54*(3), 702–709.

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, *16*(4), 495–500. https://doi.org/10.1016/0163-6383(93)80007-U

Morales, M., Mundy, P., & Rojas, J. (1998). Following the direction of gaze and language development in 6-month-olds. *Infant Behavior and Development*, *21*(2), 373–377. https://doi.org/10.1016/S0163-6383(98)90014-5

Morokuma, S., Fukushima, K., Kawai, N., Tomonaga, M., Satoh, S., & Nakano, H. (2004). Fetal habituation correlates with functional brain development. *Behavioural Brain Research*, *153*(2), 459–463. https://doi.org/10.1016/j.bbr.2004.01.002

Murray, L., & Trevarthen, C. (1985). Emotion regulation of interactions between 2 month old infants and their mothers. In T. M. Field & N. A. Fox (Eds.), *Social perception in infants* (pp. 137–154). Ablex.

Nadel, J., Carchon, I., Kervella, C., Marcelli, D., & Réserbat-Plantey, D. (1999). Expectancies for social contingency in 2-month-olds. *Developmental Science*, *2*(2), 164–173. https://doi.org/10.1111/1467-7687.00065

Papoušek, M., Bornstein, M. H., Nuzzo, C., Papoušek, H., & Symmes, D. (1990). Infant responses to prototypical melodic contours in parental speech. *Infant Behavior and Development*, *13*(4), 539–545. https://doi.org/10.1016/0163-6383(90)90022-Z

Patterson, M. L., & Werker, J. F. (1999). Matching phonetic information in lips and voice is robust in 4.5-month-old infants. *Infant Behavior and Development*, *22*(2), 237–247. https://doi.org/10.1016/S0163-6383(99)00003-X

Patterson, M. L., & Werker, J. F. (2002). Infants' ability to match dynamic phonetic and gender information in the face and voice. *Journal of Experimental Child Psychology*, *81*(1), 93–115. https://doi.org/10.1006/jecp.2001.2644

Patterson, M. L., & Werker, J. F. (2003). Two-month-old infants match phonetic information in lips and voice. *Developmental Science*, *6*(2), 191–196. https://doi.org/10.1111/1467-7687.00271

Poeppel, D., Emmorey, K., Hickok, G., & Pylkkänen, L. (2012). Towards a new neurobiology of language. *Journal of Neuroscience*, *32*(41), 14125–14131. https://doi.org/10.1523/JNEUROSCI.3244-12.2012

Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d /–/ð / perception: Evidence for a new developmental pattern. *The Journal of the Acoustical Society of America*, *109*(5), 2190–2201. https://doi.org/10.1121/1.1362689

Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, *288*(5464), 349–351. https://doi.org/10.1126/science.288.5464.349

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. E. (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science*, *29*(5), 700–710. https://doi.org/10.1177/0956797617742725

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech vocabulary development. *Child Development*, *83*(5), 1762–1774. https://doi.org/10.1111/j.1467-8624.2012.01805.x

Rowe, M. L., Pan, B. A., & Ayoub, C. (2005). Predictors of variation in maternal talk to children: A longitudinal study of low-income families. *Parenting*, *5*(3), 259–283. https://doi.org/10.1207/s15327922par0503_3

Saffran, J. R., Werker, J. F., & Werner, L. A. (2006). The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language. In R. Siegler & D. Kuhn (Eds.), *Handbook of child psychology* (pp. 58–108). Wiley. https://doi.org/10.1002/9780470147658.chpsy0202

Samples, J. M., & Franklin, B. (1978). Behavioral responses in 7 to 9 month old infants to speech and non-speech stimuli. *Journal of Auditory Research, 18*(2), 115-123.

Senju, A., & Johnson, M. H. (2009). The eye contact effect: Mechanisms and development. *Trends in Cognitive Sciences*, *13*(3), 127–134.

Senju, A., Tucker, L., Pasco, G., Hudry, K., Elsabbagh, M., Charman, T., & Johnson, M. H. (2013). The importance of the eyes: Communication skills in infants of blind parents. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1760), 20130436. https://doi.org/10.1098/rspb.2013.0436

Senju, A., Vernetti, A., Ganea, N., Hudry, K., Tucker, L., Charman, T., & Johnson, M. H. (2015). Early social experience affects the development of eye gaze processing. *Current Biology*, *25*(23), 3086–3091. https://doi.org/10.1016/j.cub.2015.10.019

Sevy, A. B. G., Bortfeld, H., Huppert, T. J., Beauchamp, M. S., Tonini, R. E., & Oghalai, J. S. (2010). Neuroimaging with near-infrared spectroscopy demonstrates speech-evoked activity in the auditory cortex of deaf children following cochlear implantation. *Hearing Research*, *270*(1–2), 39–47. https://doi.org/10.1016/j.heares.2010.09.010

Sheridan, M. A., Sarsour, K., Jutte, D., D'Esposito, M., & Boyce, W. T. (2012). The impact of social disparity on prefrontal function in childhood. *PLoS ONE*, *7*(4), e35744. https://doi.org/10.1371/journal.pone.0035744

Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, *72*(2), B11-B21. https://doi.org/10.1016/S0010-0277(99)00047-5

Shonkoff, J. P., & Phillips, D. (2000). *From neurons to neighborhoods: The science of early child development*. National Academy Press.

Simpson, E. A., Miller, G. M., Ferrari, P. F., Suomi, S. J., & Paukner, A. (2016). Neonatal imitation and early social experience predict gaze following abilities in infant monkeys. *Scientific Reports*, *6*, 20233. https://doi.org/10.1038/srep20233

Simpson, K. L., Weaver, K. J., De Villers-Sidani, E., Lu, J. Y. F., Cai, Z., Pang, Y., Rodriguez-Porcel, F., Paul, I. A., Merzenich, M., & Lin, R. C. S. (2011). Perinatal antidepressant exposure alters cortical network function in rodents. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(45), 18465–18470. https://doi.org/10.1073/pnas.1109353108

Singh, L., Morgan, J. L., & Best, C. T. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy*, *3*(3), 365–394. https://doi.org/10.1207/S15327078IN0303_5

Smith, E. C., & Lewicki, M. S. (2006). Efficient auditory coding. *Nature*, *439*(7079), 978–982. https://doi.org/10.1038/nature04485

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

Spence, M. J., & DeCasper, A. J. (1987). Prenatal experience with low-frequency maternal-voice sounds influence neonatal perception of maternal voice samples. *Infant Behavior and Development*, *10*(2), 133–142. https://doi.org/10.1016/0163-6383(87)90028-2

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, *57*, 1454–1463. http://www.jstor.org/stable/1130423

Vernetti, A., Ganea, N., Tucker, L., Charman, T., Johnson, M. H., & Senju, A. (2018). Infant neural sensitivity to eye gaze depends on early experience of gaze communication. *Developmental Cognitive Neuroscience*, *34*, 1–6. https://doi.org/10.1016/j.dcn.2018.05.007

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. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, *10*(2), 159–171. https://doi.org/10.1111/j.1467-7687.2007.00549.x

Walden, T. A., & Ogan, T. A. (1988). The development of social referencing. *Child Development*, *59*(5), 1230–1240. https://doi.org/10.2307/1130486

Walle, E. A., Reschke, P. J., & Knothe, J. M. (2017). Social Referencing: Defining and Delineating a Basic Process of Emotion. *Emotion Review*, *9*(3), 245–252. https://doi.org/10.1177/1754073916669594

Watson, J. S., & Ramey, C. T. (1972). Reactions to response-contingent stimulation in early infancy. *Merrill-Palmer Quarterly of Behavior and Development*, *18*(3), 219–227. https://www.jstor.org/stable/23084608

Weikum, W. M., Oberlander, T. F., Hensch, T. K., & Werker, J. F. (2012). Prenatal exposure to antidepressants and depressed maternal mood alter trajectory of infant speech perception. *Proceedings of the National Academy of Sciences*, *109*(Supplement_2), 17221–17227. https://doi.org/10.1073/pnas.1121263109

Weikum, Whitney M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., & Werker, J. F. (2007). Visual language discrimination in infancy. *Science*, *316*(5828), 1159. https://doi.org/10.1126/science.1137686

Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology*, *43*(2), 230–246. https://doi.org/10.1037/h0084224

Werker, J.F., & Desjardin, R. . (1995). Listening to speech in the first year of life: Experimental influences on phoneme perception. *Current Directions in Psychological Science*, *4*(3), 76–79.

Werker, J.F., & Hensch, T. K. (2015). Critical periods in speech perception: New directions. *Annual Review of Psychology*, *66*, 173–196. https://doi.org/10.1146/annurev-psych-010814-015104

Werker, J.F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*(1), 49–63. https://doi.org/10.1016/S0163-6383(84)80022-3

Werker, J.F., V Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, *52*(1), 349–355.

Wyttenbach, R. A., May, M. L., & Hoy, R. R. (1996). Categorical perception of sound frequency by crickets. *Science*, *273*(5281), 1542–1544.

Zimmerman, F. J., Gilkerson, J., Richards, J. A., Christakis, D. A., Xu, D., Gray, S., & Yapanel, U. (2009). Teaching by listening: The importance of adult-child conversations to language development. *Pediatrics*, *124*(1), 342–349. https://doi.org/10.1542/peds.2008-2267