Journal of Experimental Child Psychology 183 (2019) 48–64

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

The role of multisensory development in early language learning

Gina M. Mason *, Michael H. Goldstein, Jennifer A. Schwade

Department of Psychology, Cornell University, Ithaca, NY 14853, USA

ARTICLE INFO

Article history:
Available online 8 March 2019

Keywords:
Infant development
Social learning
Multisensory perception
Caregiver–infant interactions
Word learning
Neurodevelopmental disorders

ABSTRACT

In typical development, communicative skills such as language emerge from infants' ability to combine multisensory information into cohesive percepts. For example, the act of associating the visual or tactile experience of an object with its spoken name is commonly used as a measure of early word learning, and social attention and speech perception frequently involve integrating both visual and auditory attributes. Early perspectives once regarded perceptual integration as one of infants' primary challenges, whereas recent work suggests that caregivers' social responses contain structured patterns that may facilitate infants' perception of multisensory social cues. In the current review, we discuss the regularities within caregiver feedback that may allow infants to more easily discriminate and learn from social signals. We focus on the statistical regularities that emerge in the moment-by-moment behaviors observed in studies of naturalistic caregiver–infant play. We propose that the spatial form and contingencies of caregivers' responses to infants' looks and prelinguistic vocalizations facilitate communicative and cognitive development. We also explore how individual differences in infants' sensory and motor abilities may reciprocally influence caregivers' response patterns, in turn regulating and constraining the types of social learning opportunities that infants experience across early development. We end by discussing implications for neurodevelopmental conditions affecting both multisensory integration and communication (i.e., autism) and suggest avenues for further research and intervention.

* Corresponding author. Fax: +1 607 255 8433.
E-mail address: gmm89@cornell.edu (G.M. Mason).

https://doi.org/10.1016/j.jecp.2018.12.011
0022-0965/

Introduction

As multisensory organisms, humans experience the world through many simultaneous pathways that provide information for guiding subsequent behavior. Within the first years of life, our brains and bodies learn not only to organize the many sights, sounds, smells, and other sensory information we encounter into coherent percepts, but also to further associate these percepts with the particular words and communicative signals of our surrounding social environment (Lewkowicz, Minar, Tift, & Brandon, 2015; Yu & Smith, 2012). From an evolutionary perspective, such achievements provide us with a subjective framework both for organizing our own experiences, and for eliciting responses and building cooperation and synchrony with others (Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998; Deák, Triesch, Krasno, de Barbaro, & Robledo, 2013; Locke, 2001; Seyfarth, Cheney, & Marler, 1980). In addition, at the level of proximal development, early multisensory integration and word learning are associated with later social and cognitive outcomes (Marchman & Fernald, 2008), whereas difficulties in multisensory social perception and communicative development are commonly associated with neurodevelopmental conditions such as autism (American Psychiatric Association [APA], 2013; Landa, Holman, & Garrett-Mayer, 2007; Stevenson et al., 2014).

Although it is clear that organizing sensory information into words and social signals is a significant milestone in human development, it remains unknown *how* humans form connections between sensory information and words in the first place. This question is particularly pertinent across infants' first years, as word learning progresses across periods of significant sensory, perceptual, and motoric change (Fenson et al., 1994; Smith, Jayaraman, Clerkin, & Yu, 2018). How, then, do infants learn to associate multisensory information with distinct words and referents, particularly amid the many diverse streams of sensory input that they experience throughout development?

In exploring this question, some research efforts have focused on identifying the specific time course and dynamics of infants' early sensory development, including when and how infants' sensory systems might interact to create percepts prior to the onset of language. One proposal stemming from such research has been that infants' sensory systems are fairly integrated starting early in development, and that infants are born with a basic ability to perceive certain stimulus properties across multiple senses (Bahrick & Lickliter, 2002; Gibson, 1969; see Lewkowicz, 2002, for a review). According to this view, such properties (e.g., rhythm, temporal synchrony) occur as common patterns in infants' everyday social environments, and may help to guide infants' attention adaptively to relevant sources of social information that can facilitate further perceptual refinement and word learning (Bahrick & Lickliter, 2002). However, although infants' sensory systems share information across early development, research has also suggested that infants' sensory systems develop at different relative timescales during certain prenatal and early postnatal sensitive periods (Gottlieb, 1976; Lickliter, 2011; Turkewitz & Kenny, 1982). Moreover, such work has suggested that the specific timing and coordination of infants' sensory development may create unique learning opportunities, both for general perceptual development and for communicative and social learning in particular (Bremner, 2017; Lickliter, 2011; Smith et al., 2018; Turkewitz & Kenny, 1982).

Along with questions surrounding infants' early sensory abilities, developmental researchers have also begun to explore what external and innate information sources must be present for infants to make connections between their multisensory experiences and the presence of words and social signals. Much work has shown that infants' external environments contain reliable patterns of structured social information, and that such patterns are often apparent in the context of caregivers' responses to their infants' behaviors (Bornstein, Putnick, Cote, Haynes, & Suwalsky, 2015; Feldman, 2007; Goldstein, King, & West, 2003; Goldstein & Schwade, 2008, 2010; Goldstein, Waterfall, et al., 2010; Jaffe et al., 2001; Warlaumont, Richards, Gilkerson, & Oller, 2014). However, current perspectives differ on what internal structures or knowledge infants require to detect such social patterns. Specifically, some perspectives maintain that infants require specialized innate knowledge to identify social cues early in development (e.g., Batki, Baron-Cohen, Wheelwright, Connellan, & Ahluwalia, 2000; Csibra & Gergely, 2009), whereas others suggest that general perceptual abilities and domain-general learning mechanisms are sufficient to guide infants toward social understanding through early experience (e.g., Goldstein, Waterfall, et al., 2010; Gredebäck, Astor, & Fawcett, 2018; Heyes, 2012). With some exceptions (e.g., Gredebäck et al., 2018), definitive evidence supporting either perspective has been difficult to acquire, likely in part because the same experimental findings are often interpreted differently by individuals supporting different theories (Heyes, 2014). Furthermore, many previous studies examining mechanisms of infant social learning have done so using looking-time paradigms that differ from infants' everyday learning environments, making it challenging to determine whether infant learning strategies observed in the laboratory are employed in more true-to-life settings.

In response to the question of what structures or experiences are necessary for multisensory social perception and language development, recent years have seen new advancements in research techniques that capture infant behavior and perception in more ecologically relevant contexts, including during everyday social exchanges with caregivers and other adults (e.g., Clerkin, Hart, Rehg, Yu, & Smith, 2017; Deák, Krasno, Triesch, Lewis, & Sepeta, 2014; Goldstein & Schwade, 2008; Karasik, Tamis-Lemonda, & Adolph, 2014; Roy, Frank, DeCamp, Miller, & Roy, 2015; Smith et al., 2018). The data derived from these techniques have begun to inform us of just how structured infants' social world is, and also of how infants' developing sensory, perceptual/cognitive, and motor systems might interact with the structured social world in ways that encourage learning social signals and cues (Clerkin et al., 2017; Deák et al., 2013; Smith et al., 2018). In addition, such work provides insights into whether hypothesized mechanisms of social learning might be plausible within the context of real-world social environments. With these recent findings in mind, we propose that regularities within infants' social environments may interact with infants' developing sensory abilities to facilitate early social learning, including word learning and discrimination of social cues. By doing so, we hope to foster novel ideas regarding mechanisms underlying social and communicative development, and to illuminate new possibilities for further research and intervention.

Infants' early sensory development: Traditional accounts, timing, and the role of experience

Infants' sensory development is not a new topic in developmental science. In William James's reflections on human perception in his seminal *Principles of Psychology,* he famously wondered at the early sensory experiences of infants, assuming that young infants experience the world as an indecipherable jumble of simultaneous sensations, that is, "as one great blooming, buzzing confusion" (James, 1890, p. 488). For James, extensive postnatal sensory experience was essential for overcoming this early confusion, and for eventually learning to discriminate and label different percepts.

Since James's writings, studies in humans and nonhuman animals have elucidated more clearly the possible temporal order and mechanisms of multisensory development, and the more precise roles of experience in forming specific percepts. Regarding the order of sensory development, Gottlieb (1971) found that sensory functions appear to have a fairly conserved sequence of onset across many species of birds and mammals (including humans), with tactile sensations arising first and vestibular, chemical, auditory, and visual sensations following.[1] These findings contrast with James's "blooming, buzzing confusion" given that later senses appear to emerge in a coordinated manner following sufficient organization of earlier-developing sensory systems (Gottlieb, 1971, 1976; Turkewitz & Kenny, 1985). Furthermore, alterations in the timing of this sequence seem to influence certain aspects of sensory integration and perception, depending on the species and developmental period examined (Gottlieb, Tomlinson, & Radell, 1989; Lickliter, 1990; Turkewitz & Kenny, 1985). For example, experimentally opening the eyes of rat pups prior to typical vision onset disrupts the normal progression of homing, in which young rats use sensory cues to find their way back to their family's nest (Turkewitz & Kenny, 1985). Typical rats show an increase in homing behavior followed by a decrease at the end of the second postnatal week, whereas rats whose eyes are opened prematurely do not decrease their homing behavior even as the nest begins to hold less significance as a site of care and nurturing. Turkewitz and Kenny (1985) suggested

[1] Olfactory and gustatory ("chemical") sensations were not discussed in Gottlieb's original writings, although other investigations posited that these sensations likely often arise after vestibular development and prior to audition and vision (Bradley & Mistretta, 1975; Turkewitz & Kenny, 1985).

that these behavioral differences arise because the typical onset of vision helps to reorganize rats' behavior adaptively away from homing, drawing perceptual attention from the previously established olfactory and thermal cues around which homing behaviors were organized. In contrast, prematurely opening rats' eyes inhibits the reorganizing effects of vision given that olfaction and touch-based homing abilities are not yet fully developed. As a result, homing becomes organized around visual and olfactory/tactile input simultaneously, and vision onset no longer shifts perception away from homing. Critically, if young rats' eyes are opened prematurely but visual homing cues are removed, the typical progression and decline of homing remains unchanged. These findings suggest that sensory onset must be coupled with specific sensory experience to influence perceptual organization.

In addition to perceptual differences caused by premature sensory experience, many studies have shown disruptive effects of early sensory deprivation on later sensory and perceptual abilities (Chen, Lewis, Shore, & Maurer, 2017; Greenough, Black, & Wallace, 1987; Maurer, Mondloch, & Lewis, 2007; Putzar, Goerendt, Lange, Rösler, & Röder, 2007). For example, such disruptions have been observed in humans with congenital cataracts who had their vision restored after their first 5 months of infancy. Studies assessing such patients' later multisensory abilities have indicated selective impairments in certain audiovisual capacities, including integration of audiovisual speech cues (Putzar et al., 2007) and audiovisual simultaneity perception (Chen et al., 2017). Intriguingly, however, multisensory perception of visual and tactile simultaneity is not impaired (Chen et al., 2017). In addition, although many unisensory visual abilities are fully restored with treatment and experience, certain more nuanced social perceptual strategies are compromised, including the ability to identify different faces based on spacing of the eyes and internal features (Maurer et al., 2007) and the strategy of processing faces holistically rather than as individual components (Le Grand, Mondloch, Maurer, & Brent, 2004).

Although the presence of such selective impairments may seem unusual, it makes more sense when we consider the specific environmental input and experiences that typical humans encounter during their first 5 months of life. In face processing, studies of infants' everyday visual environments indicate that infants receive the most prominent visual exposure to faces within the first few months of postnatal life (Fausey, Jayaraman, & Smith, 2016; Jayaraman, Fausey, & Smith, 2015). During these months, caregivers' face-to-face interactions with infants are often in very close proximity, unlike interactions during later months that are characterized by infants' increased motor independence. Given that infants' exposure to close-up faces declines rapidly over the first year of life (Jayaraman et al., 2015), it follows that visual deprivation during infants' first months could specifically affect later face perception abilities. However, why are deficits in audiovisual and visual–tactile multisensory integration dissociated among these patients? In response to this question, Bremner (2017) noted that infants do not begin directing visual attention reliably to tactile cues until later in infancy (around 10 months), meaning that early visual deficits might not significantly affect visual–tactile integration. However, audiovisual synchrony detection typically arises very early in development (Lewkowicz, Leo, & Simion, 2010), with typical infants receiving substantial audiovisual input from close social interactions with speaking caregivers (Bremner, 2017; Smith et al., 2018). Taken together, these examples illustrate that both developmental timing and presence of species-typical sensory experiences are important for shaping later perception and for determining which specific sensory and perceptual abilities are impaired or spared.

Adaptive limitations: Considering early sensory trajectories in the context of word learning

Sensory experience and timing are important for overall perceptual development, but what specific sensory experiences are necessary for word learning? And how might variations in early multisensory experience and development create individual differences in later communicative skills?

Among humans, all sensory systems, including vision, are capable of at least some function at birth (Gottlieb, 1976), and responsiveness to auditory information has been reported within the third trimester of fetal development (Birnholz & Benacerraf, 1983). In spite of such early sensory abilities, human infants are otherwise highly altricial and motorically immature, relying on adult caregivers for care and survival. This combination of early sensory ability and long-term dependence on care from others is relatively unique among altricial species, and may support early social and communicative learning (Gottlieb, 1976).

For example, consider potential benefits of early auditory onset for organizing social learning. Studies of the human intrauterine environment have shown that at least some speech-related acoustic information is available for the fetus, especially via mothers' speech (see Lecanuet, 1998, for a review), and that fetuses within the third trimester learn from this acoustic information. For instance, fetuses at 37 weeks react differently to familiar versus unfamiliar speech passages, exhibiting decreases in heart rate when exposed to auditory recordings of nursery rhymes that their mothers had recited to them during the few weeks prior to testing (DeCasper, Lecanuet, Busnel, Granier-Deferre, & Maugeais, 1994). Using a non-nutritive sucking paradigm, DeCasper and Spence (1986) also showed that newborns can distinguish speech passages that their mothers recited to them prenatally, in contrast to novel speech passages recorded by the same person. Furthermore, newborns discriminate and prefer their mother's voice over the voice of another person (DeCasper & Fifer, 1980). These findings suggest that even before birth, infants' auditory systems are prepared to detect the regularities of their social environment, including speech structures of individuals who may continue to be a persistent source of further social stimulation and learning (Webb, Heller, Benson, & Lahav, 2015). Such findings also imply that the fetal auditory system already extracts relevant prosodic information to inform infants' earliest vocal behaviors, potentially setting a foundation for later language development.

Of course, humans' early auditory abilities and prenatal learning do not provide the complete answer to how infants learn words and language. Word learning is a gradual process, in which humans must associate distinct patterns of converging sensory input with specific language symbols (which themselves can be represented through audition, vision, and/or touch) that represent their convergence. Thus, although learning familiar auditory patterns is one small step in language development for typically developing infants, infants must still overcome the challenges of parsing these auditory patterns into individual words and of connecting these words with their corresponding (and often multimodal) external referents. To better understand how infants' trajectory of sensory development might assist with these challenges, we must also then consider the development and integration of other modalities into infants' perception.

Vision, for instance, is the last modality to arise in typical human development, and is one of the most commonly studied in word learning. Although vision is functional for newborn infants, newborns' visual acuity is limited relative to the visual capabilities of older infants and children (Mayer & Dobson, 1982). However, there is evidence that newborns and young infants already have some ability to integrate auditory and visual cues, particularly in detecting temporal synchronies between auditory and visual elements of both social and nonsocial stimuli (Lewkowicz et al., 2010). In addition, young infants' limited acuity may help with focusing infants' attention on only the closest and most salient visual stimuli in their environment, potentially reducing the likelihood of confusion and distraction in complex everyday contexts (Smith et al., 2018; Turkewitz & Kenny, 1985). As mentioned previously, some of the most reliable close-up images that young infants are exposed to are those of their caregivers' talking faces (Fausey et al., 2016; Jayaraman et al., 2015), and deprivation of this early visual input (among others) can lead to differences in certain aspects of face perception and audiovisual integration (Chen et al., 2017; Maurer et al., 2007; Putzar et al., 2007). Thus, it may be that young infants' visual limitations facilitate early speech perception and word learning by inadvertently allowing caregivers' close audiovisual feedback to be prioritized for sensory processing relative to stimuli farther away. The partial overlap of visual limitations with slow motor development may also be beneficial, because motor immaturity may serve to limit infants' exploration of other possible objects while infants initially learn how to integrate caregivers' audiovisual social cues.

Reflecting on the social settings in which infants' auditory and visual abilities emerge allows us to pinpoint some possible ways in which the timing of sensory onset may assist with early communicative learning. Because both audition and vision arise while infants are entirely dependent on caregivers, it is probable that many of infants' early audiovisual experiences will involve some type of social component, whether direct (i.e., caregivers and infants interacting dyadically) or indirect (i.e., infants observing social exchanges between caregivers and others). The number of words spoken during such experiences is estimated to be in the range of hundreds to thousands per hour (Locke, 2001; Shneidman & Goldin-Meadow, 2012), meaning that infants have extensive audiovisual information from which to extract basic patterns from their ambient language. From these patterns, infants can subsequently identify words and discrete language structures both by mapping reoccurring sounds onto related visual events and by building expectations of transitional probabilities (the likelihood that one event will lead or follow another; Saffran, Aslin, & Newport, 1996) between specific syllables. In addition, although early auditory experience and visual limitations already assist infants in narrowing which sensory information to attend to in early social environments, specific characteristics of adults' social feedback toward infants' behaviors may also help to guide infants' attention and augment word learning.

Caregiver–infant interactions and infants' multisensory word learning

In exploring how adult–infant interactions might influence infant word learning, early research placed an emphasis on characterizing how parents change their speech patterns when addressing infants and whether such changes relate to later language (see Snow, 1977, and Spinelli, Fasolo, & Mesman, 2017, for reviews). Although such research provided an alternative to the then-predominant view that language acquisition is innate and essentially independent of experience (e.g., Chomsky, 1967), one critique of early studies was that caregivers' behavior was being measured without regard for infants' own actions (Snow, 1977). That is, caregivers' speech was being studied purely as input for infants, and not as part of a bidirectional conversation in which the efficacy of caregivers' feedback depends on the concurrent attention and actions of infants (and vice versa). In response to this critique, more recent studies have focused on the timing and coordination of caregivers' responses relative to infants' behavior and how infant behaviors may influence the feedback that caregivers provide (Albert, Schwade, & Goldstein, 2018; Goldstein, Schwade, Briesch, & Syal, 2010; Gros-Louis, West, Goldstein, & King, 2006; Hsu & Fogel, 2003; Karasik et al., 2014).

Infant-directed speech

Adult caregivers speak differently to infants than to other adults. Though cultural variation has been observed in caregivers' speech to infants (e.g., Brown & Gaskins, 2014; De Leon, 2000; Farran, Lee, Yoo, & Oller, 2016), differences in aspects of speech such as prosody are common in many languages, with caregivers often employing higher pitch and frequency, wider pitch variability, shorter utterances, and longer pauses in their speech to infants (Broesch & Bryant, 2015; Farran et al., 2016; Fernald & Simon, 1984; Fernald et al., 1989; Grieser & Kuhl, 1988). Caregivers' speech to infants is also highly articulated (Burnham, Kitamura, & Vollmer-Conna, 2002; Newport, 1977), and although caregivers provide infants with a variety of different sentence types, the grammatical and lexical elements of these sentences are often simpler and more repetitive than those of sentences directed to adults (Newport, 1977; Wirén, Björkenstam, Grigonyté, & Cortes, 2016). Together, the speech style resulting from caregivers' characteristic changes in prosody and grammar is known as infant-directed speech (IDS), also historically termed "motherese" (Newport, 1977). IDS has been frequently studied in relation to language learning and other aspects of cognitive development because researchers have hypothesized since its discovery that the features of IDS serve some functional or facilitative role in infant cognition (see Cristia, 2013, Soderstrom, 2007, and Spinelli et al., 2017, for reviews). Given findings connecting IDS with positive language outcomes (Cristia, 2013; Spinelli et al., 2017), how might specific aspects of IDS support multisensory word learning?

Considering the prosodic characteristics of IDS, many perspectives have been put forth regarding how differences in features such as fundamental frequency (perceived acoustically as pitch) and other aspects of prosody might influence word learning. One of the most prominent theories is that such features allow IDS to be more attention-grabbing and engaging for infants, which in turn allows for better encoding of the linguistic patterns present in caregivers' speech (Golinkoff, Can, Soderstrom, & Hirsh-Pasek, 2015; Spinelli et al., 2017). This idea is corroborated by various studies documenting infants' early attentional preferences for IDS, which have been observed in infants as young as 2 days (Cooper & Aslin, 1990; Fernald, 1985; Fernald & Kuhl, 1987). According to such studies, the influence of IDS on infant attention appears to be driven at least in part by fundamental frequency characteristics (i.e., exaggerated frequency modulation and a wider range of pitch variation) rather than by aspects such as amplitude (perceptually "loudness") and duration (Fernald & Kuhl, 1987; see also Kaplan, Goldstein, Huckeby, Owren, & Cooper, 1995, for a discussion of the combined effects of fundamental frequency modulation and harmonic components in influencing infant attention). Although it is not entirely known why the fundamental frequency traits of IDS are more engaging for infants compared with those of adult-directed speech (ADS), some speculate that the dynamic frequency fluctuations and contours of IDS are simply more perceptually interesting and more readily associated with affective cues than fluctuations present in more monotonous speech (Fernald, 1985; Fernald & Kuhl, 1987; Golinkoff et al., 2015). Other researchers posit that the exaggerated prosody present in IDS is perhaps more novel to infants, compared with the attenuated frequencies that infants experience prenatally (Cooper & Aslin, 1990; Lecanuet & Schaal, 1996).

Regardless of why, there is some tentative evidence to suggest that the attention-facilitating aspects of IDS prosody might assist with word segmentation as well as other aspects of word learning (Graf Estes & Hurley, 2013; Ma, Golinkoff, Houston, & Hirsh-Pasek, 2011; Thiessen, Hill, & Saffran, 2005). For instance, Thiessen et al. (2005) found that 6- to 8-month-old infants are better able to segment novel words presented with IDS frequency patterns even when artificially controlling for other IDS cues (e.g., extended pauses at phrase boundaries). Perhaps more relevant to multisensory word learning, other studies have shown that infants more readily learn associations between visual stimuli and spoken labels when such labels are presented using IDS versus ADS (Graf Estes & Hurley, 2013; Ma et al., 2011). However, this effect appears to be dependent on age and prior lexical ability, as 27-month-olds and 21-month-olds with larger vocabularies show evidence of word learning from ADS (Ma et al., 2011). In addition, infants in such studies who showed enhanced word learning from IDS did not necessarily attend longer to IDS presentations during initial label learning. In interpreting this lack of difference in attention duration, Graf Estes and Hurley (2013) suggested that although infants who learn from IDS might not show longer attention to IDS, there may be a difference in the quality of attention elicited by IDS word labels that facilitates more efficient audiovisual encoding. More research is needed to explore the potential relations between the attention-enhancing effects of IDS and word learning, perhaps by relating more physiological indices of attentional arousal (de Barbaro, Clackson, & Wass, 2017; Zangl & Mills, 2007) to IDS exposure and learning outcomes.

Aside from the prosodic features already discussed, there are many other aspects of IDS that may be helpful for early multisensory word learning and language. In natural IDS, caregivers often pronounce new or important words on frequency peaks, and they also tend to place such words at the end of their utterances (Fernald & Mazzie, 1991). There is also some evidence that caregivers exaggerate the syllable lengths of the final words in their utterances (e.g., Albin & Echols, 1996; Church, Bernhardt, Shi, & Pichora-Fuller, 2005), which may assist with overall phrase segmentation in infants. In addition, caregivers' use of partially overlapping sentences ("variation sets") and simplified grammar in IDS may further highlight the general syntactic structure of infants' surrounding language and which words are most meaningful and specific to a given event (Newport, 1977; Waterfall et al., 2010; Wirén et al., 2016). Although it may seem intuitive that such features of IDS must support word learning, more experimental and longitudinal studies are needed to determine their relative contributions to language acquisition (e.g., Hoff-Ginsberg, 1986).

Multimodal IDS

Another observation relevant for the effects of IDS on multisensory word learning is that the auditory aspects of IDS do not occur in isolation; they are often accompanied by synchronous visual (and sometimes tactile) cues, providing *intersensory redundancies* that may aid in language perception and acquisition (Bahrick & Lickliter, 2002; Gogate & Bahrick, 1998; Gogate, Bolzani, & Betancourt, 2006). Researchers have identified two main pathways through which intersensory redundancy is present in IDS; the most obvious is that auditory prosody exaggerations in IDS are temporally synchronized with specific exaggerated visual facial movements and expressions (Chong, Werker, Russell, & Carroll, 2003; Green, Nip, Wilson, Mefferd, & Yunusova, 2010; Shepard, Spence, & Sasson, 2012), whereas the other relates to caregivers' use of gestures and object motion in temporal synchrony with IDS and word labels (Gogate, Bahrick, & Watson, 2000; Gogate et al., 2006).

Concerning multimodal gesturing in IDS, both observational and experimental studies have indicated that synchronies between adults' speech and manual object movements may assist infants in forming word–object associations, specifically during earlier prelinguistic development (Gogate & Bahrick, 1998; Gogate et al., 2000, 2006). Some of these studies have gone beyond paradigms in which infants are simply exposed to IDS in the absence of dyadic interaction to instead explore how multimodal verbal/gesture combinations manifest naturally in semistructured caregiver–infant play settings (Gogate et al., 2000). In one such study, Gogate et al. (2000) examined the trajectory of synchronous verbal/gestural communication among parents of infants at different stages of lexical development. By observing parent–infant interactions in which mothers were tasked with teaching their infants new words, Gogate et al. found that mothers were more likely to use synchronous words and gestures (including object motion) when teaching new words than when saying nontarget nouns or verbs. In addition, mothers' use of synchronous verbalizations and object motion decreased as a function of infants' age, with mothers of "prelexical" infants (5–8 months of age) using multimodal synchrony more often than mothers of 9- to 17-month-olds ("early lexical") (e.g., Werker, Cohen, Lloyd, Casasola, & Stager, 1998) or 21- to 30-month-olds ("advanced lexical"). Given that, in experimental settings, prelexical infants have been shown to receive the most benefit from synchronous words and gestures (Gogate & Bahrick, 1998), these results suggest that caregivers may be using verbal and gestural synchrony as a means of emphasizing word–referent relations for young infants. The decrease in parents' use of synchrony with age also implies that caregivers may be specifically tailoring their feedback to their infants' developmental level. In another study, Gogate et al. (2006) also showed that prelexical infants who switched gaze from their mothers to target objects during synchronous object labeling/motion exhibited more evidence of word learning than those who did not. Combined with prior findings, this study also suggests that infants are not simply passive observers in their own learning, and that infants' own actions help to determine what infants learn in social environments (Begus & Southgate, 2018).

Beyond IDS: Contingent responsiveness and spatial coordination

Given that infants' own behaviors and attention help to determine the effectiveness of IDS in dynamic social settings, it makes sense that caregivers' infant-directed feedback should be most efficacious when caregivers attempt to coordinate such feedback with infants' attention and actions (Bornstein et al., 2015; Goldstein & Schwade, 2008; Goldstein, Schwade, & Bornstein, 2009; Goldstein et al., 2010; Jaffe et al., 2001). To investigate the importance of dyadic coordination on infant language, recent studies have explored how caregivers' *contingent responses* (i.e., responses that are prompt and temporally reliable) to infants' prelinguistic vocalizations and attention might facilitate various aspects of language learning in everyday contexts (Goldstein & Schwade, 2008; Goldstein et al., 2010; Yu & Smith, 2012), particularly when these responses are also *spatially* coordinated with (i.e., refer to) objects and areas with which infants are visually and/or haptically engaged.

Such studies have made a set of key revelations. One is that infants, within their first year of life, seem to extract information from caregivers' responses more effectively when these responses come within a short (2- to 5-s) time window following infants' own prelinguistic vocalizations. When such responses occur reliably on infants' non-cry utterances, infants rapidly learn to incorporate phonological elements of this social feedback into their own vocalizations (Goldstein & Schwade, 2008). Infants also show differences in word and object learning when social partners respond contingently to their vocalizations as well as when infants themselves vocalize toward objects to which they are visually attending. For instance, Goldstein et al. (2010) found that when 11-month-olds looked at objects and vocalized, they were more likely to learn the names of the objects to which they were attending if social partners provided object labels immediately following the infants' vocalization. Infants at this age did not show enhanced learning when objects were labeled contingently on infants' looks without concurrent vocalizations. In addition, when infants explored objects on their own (without social feedback), they learned the visual properties of objects to which they vocalized the most compared with objects that elicited fewer vocalizations. In explaining these findings, Goldstein et al. (2010) proposed that infants' vocalizations serve as a marker of increased attention and arousal, whereby information that proximally follows each vocalization is more easily encoded. It is intriguing to consider whether the proprioceptive and/or auditory feedback that infants receive from their own vocalizations also assists in further priming infants' sensory systems for learning (see Cheng, 1992, for an example of vocal self-stimulation in other species) or whether such feedback is not essential for them to learn more readily within the seconds following their own vocalizations. More work is needed to explore these possibilities.

Another revelation from observations of adult–infant interactions is that as infants become increasingly adept at coordinating their own motor patterns beyond vocalizing, they also appear to learn more readily from social responses that are contingent on their coupled gaze and manual actions. In a study using head cameras and motion tracking, Yu and Smith (2012) showed that 18-month-olds align their manual, visual, and head movements to create transient moments in which the object to which they are attending is stable and visually dominant in their field of view compared with other objects. When parents provided object labels contingently during these visually selective moments, 18-month-olds showed evidence of learning the word–object associations in a subsequent three-alternative forced-choice paradigm. Toddlers did not learn object labels provided by their parents during other (nonselective) moments, suggesting that the temporal alignment of parents' verbal responses with infants' visual attention is an important component of word learning.

Within types of contingent responses, differences in the spatial (and semantic) alignment of social partners' responses to infants' behaviors also appear to be important for language development. For example, in Yu and Smith (2012) study, successful learning events were also characterized by parents moving their heads closer to the object of infants' focus while labeling it contingently on infants' gaze, suggesting that parents' spatial coordination with infants' attention might also facilitate word–referent learning. The fact that parents provided labels for the objects to which infants were attending also indicates that it may be important for the semantic content of parents' responses to align with infants' current focus of attention. This idea is corroborated by other data (Goldstein & Schwade, 2010) suggesting that parental contingent responses that are *contextually* relevant to infants' object of focus are significantly positively correlated with later vocabulary size, in contrast to contingent responses that imitate the phonology of infants' babbles but provide no contextual information.

In some literatures, the concurrent alignment of caregivers' response timing and content with infants' current focus has also been referred to as contingent-"sensitive" responding, and such responding has been broadly associated with better language outcomes and vocabulary development (e.g., Baumwell, Tamis-LeMonda, & Bornstein, 1997). Together, such studies indicate that, in addition to form and timing, the spatial form and semantic content of adults' responses may predict differences in audiovisual word learning.

Effects of infant behavior on caregivers' social feedback

Just as different caregiver responses may encourage different learning outcomes for infants, studies have suggested that differences in infant characteristics, including differences in their current vocal repertoire and motor abilities, may reciprocally influence caregivers' feedback (Albert et al., 2018; Gros-Louis et al., 2006; Karasik et al., 2014; Warlaumont et al., 2014). Regarding infants' vocal abilities, converging evidence shows that caregivers are selective in the types of vocalizations to which they respond contingently and in the types of contingent responses that they provide for different vocalizations (Albert et al., 2018; Gros-Louis et al., 2006; Warlaumont et al., 2014). Specifically, caregivers are more likely to respond to vocalizations that are more mature and speech-like. Infants' motor abilities also predict differences in caregivers' social responding; for example, Karasik et al. (2014) found that mothers are more likely to provide object-relevant action directives in response to their infants' object bids when infants provide "moving bids" (i.e., attentional bids in which infants present objects to their mothers while simultaneously locomoting) than when infants offer objects from a stationary position. The likelihood that infants will provide moving bids also appears to correlate with their overall motor proficiencies given that walking infants are more likely to employ moving bids than crawling infants (Karasik et al., 2014).

Overall, the studies described above indicate that infants' behaviors play a role in shaping the form and types of social responses that caregivers provide. As caregivers selectively respond to infant behaviors that they perceive as more mature or focused, such responses may encourage infants to continue advancing in their vocal and motor abilities, which in turn may provide more opportunities for social learning. This bidirectional relationship between infants' and caregivers' responses has been referred to as a "social feedback loop" for language learning (Goldstein & Schwade, 2010; Warlaumont et al., 2014). In turn, differences in infants' vocal/motor development or in the selectivity of caregivers' feedback may affect the strength of this loop, as may differences in infants' own ability to perceive (or gain reinforcement from) relations between their own behaviors and caregivers' contingent responses.

Implications for neurodevelopmental conditions: Considering bidirectional sensory, motor, and social influences on communicative outcomes in autism

Autism (also known as autism spectrum condition [ASC]) is a common neurodevelopmental condition estimated to have a prevalence of 1% to 2% globally (Centers for Disease Control and Prevention, 2018). Although the range and severity of symptoms vary across individuals, autism is generally characterized by difficulties or delays in social interaction and communication as well as restricted and repetitive behaviors, interests, or activities (APA, 2013). Despite autism's prevalence, there is not currently a common consensus regarding the underlying causes and developmental mechanisms of the autism phenotype; some researchers attribute ASC behaviors to domain-specific atypicalities in "social" brain areas or cognitive modules (e.g., Baron-Cohen, 1989; Leslie, 1992), whereas others argue for more domain-general impairments that may lead to specific deficits over development and experience (e.g., Karmiloff-Smith, 1998; Sinha et al., 2014; Thomas, Davis, Karmiloff-Smith, Knowland, & Charman, 2016).

Until recently, part of the challenge in delineating underlying causes of autism has been that autism cannot be diagnosed until later in development, meaning that many previous studies of autism were conducted in older children (for whom many developmental events potentially contributing to autism have already occurred). In response to this issue, many clinicians have turned their attention during recent years to studying younger siblings of individuals with autism during prediagnostic ages because siblings have a heightened risk of developing autism relative to the general population (Ozonoff et al., 2011; Sandin et al., 2014). Alongside retrospective analyses, these prediagnostic studies of at-risk infants have provided valuable insights into possible precursors of social and communicative impairments, including early differences in sensory processing and motor abilities that may contribute to later language and social learning difficulties (Germani et al., 2014; LeBarton & Iverson, 2013; Northrup, Libertus, & Iverson, 2017; Rogers, 2009). In addition, some studies have specifically explored how at-risk infants' early behaviors might elicit differences in parent feedback, highlighting bidirectional social mechanisms by which broader impairments may lead to specific social difficulties (e.g. Warlaumont et al., 2014).

Early sensory differences in autism

Sensory differences in autism have been well documented in older children and adults (Leekam, Nieto, Libby, Wing, & Gould, 2007; Stevenson et al., 2014); however, fewer studies have assessed sensory processing differences starting in infancy. Those that have studied sensory processing at younger ages have frequently found evidence of atypicalities, including both hypo- and hyperresponsiveness to various sensory stimuli (Baranek et al., 2013; Germani et al., 2014; McCormick, Hepburn, Young, & Rogers, 2016; Nyström, Gredebäck, Bölte, & Falck-Ytter, 2015) and "sensory seeking," defined broadly as unusual fixations or repetitive behaviors that attempt to prolong or enhance sensory input (Damiano-Goodwin et al., 2018). Although findings have historically been mixed regarding whether these sensory atypicalities are specific to autism and/or are predictive of later social impairments (Rogers & Ozonoff, 2005), more recent work has drawn clearer connections between certain early sensory characteristics—particularly increased sensory seeking and decreased sensory responsiveness—and later autism symptoms (Damiano-Goodwin et al., 2018; Germani et al., 2014).

How might the early sensory differences observed in many at-risk infants potentially affect the social and communicative difficulties found in later autism? Regarding sensory seeking, one hypothesis suggests that sensory seeking may indirectly reduce social learning opportunities by drawing attention away from dynamic social cues and more toward enhancing lower-level sensory experiences (Damiano-Goodwin et al., 2018). To test this idea, Damiano-Goodwin et al. (2018) assessed relations between increased sensory seeking and decreased social responding (orienting) in 18-month-olds at risk for autism and further explored whether social orienting differences mediate relations between sensory seeking and later social impairments. They found that sensory seeking corresponds to decreased social responsiveness and that reductions in social responding account for the significant association observed between sensory seeking and later social difficulties. Thus, it appears that sensory seeking may contribute to later social difficulties by drawing infants' attention away from proximal social cues, which in turn diminishes social learning. Independently, hyposensitivities to sensory information (i.e., reduced encoding and reactivity to external sensory stimuli) may also broadly reduce infants' ability to detect, engage and learn from social feedback, whereas hypersensitivities could potentially inhibit infants' ability to focus on the most relevant social information in the midst of distractions (Dunn, 1997). More work is needed to determine whether these more global differences in sensory sensitivity during infancy correspond differentially to social and communicative skills later in development.

In addition to overall differences in sensory responsiveness and sensory seeking, other work in autism has assessed differences in children's ability to detect temporal relations between information presented across specific modalities, including the auditory and visual domains (Foss-Feig et al., 2010; Stevenson et al., 2014). Such studies have found decreased perceptual acuity and an extended *temporal binding window* among older children with autism, meaning that these children are less able to detect differences in onset timing between asynchronous multisensory input. Given that multisensory temporal binding has been assessed only in older children and adults with autism, it is difficult to know when multisensory binding differences first arise in autism and whether they causally affect the social and communicative impairments characteristic of the condition (e.g., Stevenson et al., 2014). However, one could imagine how such temporal binding differences, if they occur during infancy, might affect infants' detection of contingencies between their visual experiences and caregivers' social feedback, which would in turn affect early word learning and communication (Goldstein, Waterfall et al., 2010; Yu & Smith, 2012). This possibility should be further explored.

Another potential area for further research concerns whether those who develop autism exhibit prenatal and perinatal atypicalities in sensory development. Although many clinicians have focused on more obvious obstetric complications as potential risk factors for later autism (e.g., Kolevzon, Gross, & Reichenberg, 2007), the more general prenatal sensory experiences and abilities of individuals later diagnosed with autism have been less well characterized. Because prenatal auditory experience appears to have some initial influence on the speech structures and sounds that typical infants prefer (e.g., DeCasper & Fifer, 1980), it would be intriguing to explore whether at-risk newborns show similar evidence of prenatal learning and whether differences in prenatal auditory development may lead to differences in speech perception and later language. In addition, further work should explore the trajectory of visual and multisensory development within the first few months of life to determine whether at-risk infants experience the same sorts of adaptive limitations that typical infants experience as they interact with an environment containing many sources of sensory stimulation (Smith et al., 2018).

Early motor differences in autism

Although data on sensory development in at-risk infants are currently somewhat limited, motor differences among infants later identified as having autism have been relatively well characterized (Bhat, Galloway, & Landa, 2012; Landa & Garrett-Mayer, 2006; LeBarton & Iverson, 2013, 2016; Nickel, Thatcher, Keller, Wozniak, & Iverson, 2013). Within the first 2 years of life, motor delays have been observed both at the level of gross motor skills (e.g., sitting, postural control, locomotion; Bhat et al., 2012; LeBarton & Iverson, 2016; Nickel et al., 2013) and at the level of fine motor abilities such as oral–motor and manual object manipulations and gestures (LeBarton & Iverson, 2013). In addition, studies have correlated early gross and fine motor delays with differences in later communicative outcomes, including expressive language and gesture articulation (LeBarton & Iverson, 2013, 2016) as well as overall language comprehension and production (Landa & Garrett-Mayer, 2006).

When we consider why motor abilities predict subsequent communication in at-risk infants, data from typical development may provide us with valuable insights. To illustrate this point, consider gross motor skills. As described previously, work in typical infants has shown that changes in gross motor skills, such as locomotor ability, correspond to changes in the types of object-related interactions that infants have with their caregivers (Karasik et al., 2014). Infants' ability to walk and deliver objects to their caregivers (as opposed to offering objects from a sitting position) encourages caregivers to provide qualitatively different types of verbal responses to their infants, including action directives that provide more information about the specific affordances of objects with which infants are engaged. In addition, locomotion combined with fine motor skills such as pointing and object manipulation allows infants to explore a wider array of objects and to play a more active role in guiding their own learning by providing clear signals to caregivers about the locations and items that they are interested in investigating (Iverson, 2010; Olson & Masur, 2011). Based on this information, we can hypothesize that because infants who develop autism are often delayed in their achievement of motor milestones (Landa & Garrett-Mayer, 2006; LeBarton & Iverson, 2013, 2016), they may have fewer early opportunities to interact and receive communicative feedback about objects in ways made possible by more advanced motor abilities. Although this hypothesis requires further study, such reduced opportunities could translate into specific differences in learning of action-specific sentence structures and object/action labels, which could in turn create differences in subsequent language development.

Aside from locomotion and gestures, other motor skills such as sitting and postural stability (also often delayed in infants at risk) may play a more direct role in influencing communicative development than previously assumed. Citing unpublished work by Yingling (1981), Iverson (2010) initially expanded on this notion by describing connections between unsupported sitting and advancements in prelinguistic vocal forms, particularly in the production of phonologically advanced consonant–vowel clusters. According to Yingling (1981), as described by Iverson (2010), the motor accomplishment of unsupported sitting allows for greater breath support (via opening of the rib cage and more consistent maintenance of pressure below the glottis) and also results in a slight repositioning of the speech articulators, such that more phonologically mature consonant–vowel production is facilitated. As we previously noted, phonologically mature infant vocalizations also encourage contingent responses from caregivers (Albert et al., 2018), which in turn facilitate phonological development and word learning (Goldstein & Schwade, 2008; Goldstein et al., 2010). Thus, it is plausible that unsupported sitting aids communicative learning through shaping not only the quality of vocalizations that infants produce but also (indirectly) the frequency and content of caregivers' social feedback. Indeed, subsequent work by LeBarton and Iverson (2016) drew further connections between the delayed onset of sitting and the onset of consonant–vowel clusters in at-risk infants, and retrospective analyses by Patten et al. (2014) showed reduced rates of canonical babbling within the first year among infants who later develop autism. This reduction in canonical babbling may in turn contribute to the reduction in "speech-like" vocalizations found in older children with autism (Warlaumont et al., 2014), and to caregivers' subsequent lack of selectivity in responding to their children's speech-like versus non-speech-like vocalizations (Warlaumont et al., 2014).

Implications of sensory and motor differences for social difficulties in autism

To summarize, early sensory impairments such as those observed in autism may impair communicative learning not only by directly altering infants' perception of the temporal and sensory features of social feedback, but also by diverting infants' attention to other nonsocial sensory stimuli in the broader environment. Similarly, motor impairments may reduce infants' ability to interact with the environment in more directed and socially focused ways, and differences in specific motor abilities may also inhibit the production of vocal forms and gestures that typically elicit language-facilitating social feedback from caregivers (e.g., LeBarton & Iverson, 2016). Future work should focus on exploring sensory differences in individuals with autism at earlier timepoints (i.e., prenatally) as well as on how differences in motor skills observed in at-risk infants proximally affect the form, timing, and content of caregivers' verbal responses in everyday social interactions.

Concluding remarks

When we ponder the problem of language learning, it may seem remarkable at first glance that human infants can somehow take the many streams of information flooding in from all of their senses and parse this information into discrete words and sentence structures within their first years of development. Fortunately, infants are not alone in this endeavor, and their trajectory of sensory development is such that the "blooming, buzzing confusion" of their initial sensory experiences is likely much tamer than initially presumed. The combination of infants' early sensory abilities and limitations, and the extended presence of caregivers early in development adaptively constrain and facilitate the information and statistics that infants extract from their early social environments (Goldstein, Waterfall, et al., 2010; Smith et al., 2018). As infants' sensory and motor abilities continue to develop, infants become progressively more able to expand on the initial social statistics that they have aggregated through their early experiences because they can explore in more sensorimotor detail the objects and social partners surrounding them. In turn, infants' sensorimotor achievements also influence the timing and content of the social responses that caregivers provide, allowing for further communicative learning opportunities as development progresses. Thus, differences in infants' early sensory and motor development may influence communicative learning not only directly, through atypical information encoding and perception, but also indirectly, through influencing caregivers' social feedback. By assessing the influences of infants' sensory and motor development on their own behavior and on caregivers' responding in everyday contexts, we may gain insights into real-world mechanisms of early social and communicative development.

References

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

Albin, D. D., & Echols, C. H. (1996). Stressed and word-final syllables in infant-directed speech. *Infant Behavior and Development, 19*, 401–418.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Bahrick, L. E., & Lickliter, R. (2002). Intersensory redundancy guides early perceptual and cognitive development. *Advances in Child Development and Behavior, 30*, 153–187.

Baranek, G. T., Watson, L. R., Boyd, B. A., Poe, M. D., David, F. J., & McGuire, L. (2013). Hyporesponsiveness to social and nonsocial sensory stimuli in children with autism, children with developmental delays, and typically developing children. *Development and Psychopathology, 25*, 307–320.

Baron-Cohen, S. (1989). The autistic child's theory of mind: A case of specific developmental delay. *Journal of Child Psychology and Psychiatry, 30*, 285–297.

Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., & Ahluwalia, J. (2000). Is there an innate gaze module? Evidence from human neonates. *Infant Behavior and Development, 23*, 223–229.

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

Begus, K., & Southgate, V. (2018). Curious learners: How infants' motivation to learn shapes and is shaped by infants' interactions with the social world. In M. M. Saylor & P. A. Ganea (Eds.), *Active learning from infancy to childhood* (pp. 13–37). Cham, Switzerland: Springer International.

Bhat, A. N., Galloway, J. C., & Landa, R. J. (2012). Relation between early motor delay and later communication delay in infants at risk for autism. *Infant Behavior and Development, 35*, 838–846.

Birnholz, J. C., & Benacerraf, B. R. (1983). The development of human fetal hearing. *Science, 222*, 516–519.

Bornstein, M. H., Putnick, D. L., Cote, L. R., Haynes, O. M., & Suwalsky, J. T. D. (2015). Mother–infant contingent vocalizations in 11 countries. *Psychological Science, 26*, 1272–1284.

Bradley, R. M., & Mistretta, C. M. (1975). Fetal sensory receptors. *Physiological Reviews, 55*, 352–382.

Bremner, A. J. (2017). Multisensory development: Calibrating a coherent sensory milieu in early life. *Current Biology, 27*, R305–R307.

Broesch, T. L., & Bryant, G. A. (2015). Prosody in infant-directed speech is similar across Western and traditional cultures. *Journal of Cognition and Development, 16*, 31–43.

Brown, P., & Gaskins, S. (2014). Language acquisition and language socialization. In N. J. Enfield, P. Kockelman, & J. Sidnell (Eds.), *The Cambridge handbook of linguistic anthropology* (pp. 187–226). Cambridge, UK: Cambridge University Press.

Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's new, pussycat? On talking to babies and animals. *Science, 296*, 1435.

Carpenter, Ma., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (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* (4, Serial No. 176).

Centers for Disease Control and Prevention. (2018, April 26). Data and statistics on autism spectrum disorder. Retrieved June 4, 2018, from <https://www.cdc.gov/ncbddd/autism/data.html>.

Chen, Y. C., Lewis, T. L., Shore, D. I., & Maurer, D. (2017). Early binocular input is critical for development of audiovisual but not visuotactile simultaneity perception. *Current Biology, 27*, 583–589.
Cheng, M.-F. (1992). For whom does the female dove coo? A case for the role of vocal self-stimulation. *Animal Behaviour, 43*, 1035–1044.
Chomsky, N. (1967). Recent contributions to the theory of innate ideas—Summary of oral presentation. *Synthese, 17*(1), 2–11.
Chong, S. C. F., Werker, J. F., Russell, J. A., & Carroll, J. M. (2003). Three facial expressions mothers direct to their infants. *Infant and Child Development, 12*, 211–232.
Church, R., Bernhardt, B., Shi, R., & Pichora-Fuller, K. (2005). Infant-directed speech: Final syllable lengthening and rate of speech. *Journal of the Acoustical Society of America, 117*, 2429–2430.
Clerkin, E. M., Hart, E., Rehg, J. M., Yu, C., & Smith, L. B. (2017). Real-world visual statistics and infants' first-learned object names. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372*, 20160055.
Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61*, 1584–1595.
Cristia, A. (2013). Input to language: The phonetics and perception of infant-directed speech. *Language and Linguistics Compass, 7*, 157–170.
Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*, 148–153.
Damiano-Goodwin, C. R., Woynaroski, T. G., Simon, D. M., Ibañez, L. V., Murias, M., Kirby, A., ... Cascio, C. J. (2018). Developmental sequelae and neurophysiologic substrates of sensory seeking in infant siblings of children with autism spectrum disorder. *Developmental Cognitive Neuroscience, 29*, 41–53.
de Barbaro, K., Clackson, K., & Wass, S. V. (2017). Infant attention is dynamically modulated with changing arousal levels. *Child Development, 88*, 629–639.
De Leon, L. (2000). The emergent participant: Interactive patterns in the socialization of Tzotzil (Mayan) infants. *Journal of Linguistic Anthropology, 8*, 131–161.
Deák, G. O., Krasno, A. M., Triesch, J., Lewis, J., & Sepeta, L. (2014). Watch the hands: Infants can learn to follow gaze by seeing adults manipulate objects. *Developmental Science, 17*, 270–281.
Deák, G. O., Triesch, J., Krasno, A., de Barbaro, K., & Robledo, M. (2013). Learning to share: The emergence of joint attention in human infancy. In B. R. Kar (Ed.), *Cognition and brain development: Converging evidence from various methodologies* (pp. 173–210). Washington, DC: American Psychological Association.
DeCasper, A., & Fifer, W. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science, 208*, 1174–1176.
DeCasper, A. J., Lecanuet, J. P., Busnel, M. C., Granier-Deferre, C., & Maugeais, R. (1994). Fetal reactions to recurrent maternal speech. *Infant Behavior and Development, 17*, 159–164.
DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development, 9*, 133–150.
Dunn, W. (1997). The impact of sensory processing abilities on the daily lives of young children and their families: A conceptual model. *Infants & Young Children, 9*(4), 23–35.
Farran, L. K., Lee, C. C., Yoo, H., & Oller, D. K. (2016). Cross-cultural register differences in infant-directed speech: An initial study. *PLoS One, 11*(3), e151518.
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.
Feldman, R. (2007). Parent–infant synchrony: Biological foundations and developmental outcomes. *Current Directions in Psychological Science, 16*, 340–345.
Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., ... Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59* (5, Serial No. 242).
Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8*, 181–195.
Fernald, A., & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10*, 279–293.
Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology, 27*, 209–221.
Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20*, 104–113.
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*, 477–501.
Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., & Wallace, M. T. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research, 203*, 381–389.
Germani, T., Zwaigenbaum, L., Bryson, S., Brian, J., Smith, I., Roberts, W., ... Vaillancourt, T. (2014). Assessment of early sensory processing in infants at high-risk of autism spectrum disorder. *Journal of Autism and Developmental Disorders, 44*, 3264–3270.
Gibson, E. J. (1969). *Principles of perceptual learning and development*. East Norwalk, CT: Appleton–Century–Crofts.
Gogate, L. J., & Bahrick, L. E. (1998). Intersensory redundancy facilitates learning of arbitrary relations between vowel sounds and objects in seven-month-old infants. *Journal of Experimental Child Psychology, 69*, 133–149.
Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development, 71*, 878–894.
Gogate, L. J., Bolzani, L. H., & Betancourt, E. A. (2006). Attention to maternal multimodal naming by 6- to 8- month-old infants and learning of word–object relations. *Infancy, 9*, 259–288.
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*, 8030–8035.
Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science, 19*, 515–523.
Goldstein, M. H., & Schwade, J. A. (2010). From birds to words: Perception of structure in social interactions guides vocal development and language learning. In M. S. Blumberg, J. H. Freeman, & S. R. Robertson (Eds.). *The Oxford handbook of developmental and comparative neuroscience* (Vol. 29, pp. 737–767). Oxford, UK: Oxford University Press.

Goldstein, M. H., Schwade, J. A., & Bornstein, M. H. (2009). The value of vocalizing: Five-month-old infants associate their own noncry vocalizations with responses from caregivers. *Child Development, 80*, 636–644.

Goldstein, M. H., Schwade, J., Briesch, J., & Syal, S. (2010). Learning while babbling: Prelinguistic object-directed vocalizations indicate a readiness to learn. *Infancy, 15*, 362–391.

Goldstein, M. H., Waterfall, H. R., Lotem, A., Halpern, J. Y., Schwade, J. A., Onnis, L., & Edelman, S. (2010). General cognitive principles for learning structure in time and space. *Trends in Cognitive Sciences, 14*, 249–258.

Golinkoff, R. M., Can, D. D., Soderstrom, M., & Hirsh-Pasek, K. (2015). (Baby)Talk to me: The social context of infant-directed speech and its effects on early language acquisition. *Current Directions in Psychological Science, 24*, 339–344.

Gottlieb, G. (1971). Ontogenesis of sensory function in birds and mammals. *The biopsychology of development*, 67–128.

Gottlieb, G. (1976). Conceptions of prenatal development: Behavioral embryology. *Psychological Review, 83*, 215–234.

Gottlieb, G., Tomlinson, W. T., & Radell, P. L. (1989). Developmental intersensory interference: Premature visual experience suppresses auditory learning in ducklings. *Infant Behavior and Development, 12*, 1–12.

Graf Estes, K., & Hurley, K. (2013). Infant-directed prosody helps infants map sounds to meanings. *Infancy, 18*, 797–824.

Gredebäck, G., Astor, K., & Fawcett, C. (2018). Gaze following is not dependent on ostensive cues: A critical test of natural pedagogy. *Child Development, 89*, 2091–2098.

Green, J. R., Nip, I. S. B., Wilson, E. M., Mefferd, A. S., & Yunusova, Y. (2010). Lip movement exaggerations during infant-directed speech. *Journal of Speech, Language, and Hearing Research, 53*, 1529–1542.

Greenough, W. T., Black, J. E., & Wallace, C. S. (1987). Experience and brain development. *Child Development, 58*, 539–559.

Grieser, D. A. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology, 24*, 14–20.

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development, 30*, 509–516.

Heyes, C. (2012). What's social about social learning? *Journal of Comparative Psychology, 126*, 193–202.

Heyes, C. (2014). False belief in infancy: A fresh look. *Developmental Science, 17*, 647–659.

Hoff-Ginsberg, E. (1986). Function and structure in maternal speech: Their relation to the child's development of syntax. *Developmental Psychology, 22*, 155–163.

Hsu, H.-C., & Fogel, A. (2003). Stability and transitions in mother–infant face-to-face communication during the first 6 months: A microhistorical approach. *Developmental Psychology, 39*, 1061–1082.

Iverson, J. M. (2010). Developing language in a developing body: The relationship between motor development and language development. *Journal of Child Language, 37*, 229–261.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C., Jasnow, M., Rochat, P., & Stern, D. N. (2001). Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the Society for Research in Child Development, 66* (2, Serial No. 265).

James, W. (1890). Discrimination and comparison. In *The principles of psychology* (pp. 483–549). New York: Henry Holt.

Jayaraman, S., Fausey, C. M., & Smith, L. B. (2015). The faces in infant-perspective scenes change over the first year of life. *PLoS One, 10*(5), e123780.

Kaplan, P. S., Goldstein, M. H., Huckeby, E. R., Owren, M. J., & Cooper, R. P. (1995). Dishabituation of visual attention by infant-versus adult-directed speech: Effects of frequency modulation and spectral composition. *Infant Behavior and Development, 18*, 209–223.

Karasik, L. B., Tamis-Lemonda, C. S., & Adolph, K. E. (2014). Crawling and walking infants elicit different verbal responses from mothers. *Developmental Science, 17*, 388–395.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences, 2*, 389–398.

Kolevzon, A., Gross, R., & Reichenberg, A. (2007). Prenatal and perinatal risk factors for autism: A review and integration of findings. *Archives of Pediatrics & Adolescent Medicine, 161*, 326–333.

Landa, R., & Garrett-Mayer, E. (2006). Development in infants with autism spectrum disorders: A prospective study. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 47*, 629–638.

Landa, R. J., Holman, K. C., & Garrett-Mayer, E. (2007). Social and communication development in toddlers with early and later diagnosis of autism spectrum disorders. *Archives of General Psychiatry, 64*, 853–864.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2004). Impairment in holistic face processing following early visual deprivation. *Psychological Science, 15*, 762–768.

LeBarton, E. S., & Iverson, J. M. (2013). Fine motor skill predicts expressive language in infant siblings of children with autism. *Developmental Science, 16*, 815–827.

LeBarton, E. S., & Iverson, J. M. (2016). Associations between gross motor and communicative development in at-risk infants. *Infant Behavior and Development, 44*, 59–67.

Lecanuet, J. P. (1998). Foetal responses to auditory and speech stimuli. In A. Slater (Ed.), *Perceptual development: Visual, auditory, and speech perception in infancy (Studies in Developmental Psychology* (pp. 317–355). London: Psychology Press.

Lecanuet, J. P., & Schaal, B. (1996). Fetal sensory competencies. *European Journal of Obstetrics & Gynecology and Reproductive Biology, 68*, 1–23.

Leekam, S. R., Nieto, C., Libby, S. J., Wing, L., & Gould, J. (2007). Describing the sensory abnormalities of children and adults with autism. *Journal of Autism and Developmental Disorders, 37*, 894–910.

Leslie, A. M. (1992). Pretense, autism, and the theory-of-mind module. *Current Directions in Psychological Science, 1*, 18–21.

Lewkowicz, D. J. (2002). Heterogeneity and heterochrony in intersensory development. *Cognitive Brain Research, 14*, 41–63.

Lewkowicz, D. J., Leo, I., & Simion, F. (2010). Intersensory perception at birth: Newborns match nonhuman primate faces and voices. *Infancy, 15*, 46–60.

Lewkowicz, D. J., Minar, N. J., Tift, A. H., & Brandon, M. (2015). Perception of the multisensory coherence of fluent audiovisual speech in infancy: Its emergence and the role of experience. *Journal of Experimental Child Psychology, 130*, 147–162.

Lickliter, R. (1990). Premature visual-stimulation accelerates intersensory functioning in bobwhite quail neonates. *Developmental Psychobiology, 23*, 15–27.

Lickliter, R. (2011). The integrated development of sensory organization. *Clinics in Perinatology, 38*, 591–603.

Locke, J. L. (2001). First communion: The emergence of vocal relationships. *Social Development, 10*, 294–308.

Ma, W., Golinkoff, R. M., Houston, D., & Hirsh-Pasek, K. (2011). Word learning in infant- and adult-directed speech. *Language Learning and Development, 7*, 185–201.
Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science, 11*, F9–F16.
Maurer, D., Mondloch, C. J., & Lewis, T. L. (2007). Sleeper effects. *Developmental Science, 10*, 40–47.
Mayer, D. L., & Dobson, V. (1982). Visual acuity development in infants and young children, as assessed by operant preferential looking. *Vision Research, 22*, 1141–1151.
McCormick, C., Hepburn, S., Young, G. S., & Rogers, S. J. (2016). Sensory symptoms in children with autism spectrum disorder, other developmental disorders and typical development: A longitudinal study. *Autism, 20*, 572–579.
Newport, E. L. (1977). Motherese: The speech of mothers to young children. In D. B. Pisoni & G. R. Potts (Eds.). *Cognitive theory* (Vol. 2, pp. 177–217). Hillsdale, NJ: Lawrence Erlbaum.
Nickel, L. R., Thatcher, A. R., Keller, F., Wozniak, R. H., & Iverson, J. M. (2013). Posture development in infants at heightened versus low risk for autism spectrum disorders. *Infancy, 18*, 639–661.
Northrup, J. B., Libertus, K., & Iverson, J. M. (2017). Response to changing contingencies in infants at high and low risk for autism spectrum disorder. *Autism Research, 10*, 1239–1248.
Nyström, P., Gredebäck, G., Bölte, S., & Falck-Ytter, T. (2015). Hypersensitive pupillary light reflex in infants at risk for autism. *Molecular Autism, 6*, 10.
Olson, J., & Masur, E. F. (2011). Infants' gestures influence mothers' provision of object, action and internal state labels. *Journal of Child Language, 38*, 1028–1054.
Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., ... Stone, W. L. (2011). Recurrence risk for autism spectrum disorders: A baby siblings research consortium study. *Pediatrics, 128*, e488–e495.
Patten, E., Belardi, K., Baranek, G. T., Watson, L. R., Labban, J. D., & Oller, D. K. (2014). Vocal patterns in infants with autism spectrum disorder: Canonical babbling status and vocalization frequency. *Journal of Autism and Developmental Disorders, 44*, 2413–2428.
Putzar, L., Goerendt, I., Lange, K., Rösler, F., & Röder, B. (2007). Early visual deprivation impairs multisensory interactions in humans. *Nature Neuroscience, 10*, 1243–1245.
Rogers, S. J. (2009). What are infant siblings teaching us about autism in infancy? *Autism Research, 2*, 125–137.
Rogers, S. J., & Ozonoff, S. (2005). What do we know about sensory dysfunction in autism? A critical review of the empirical evidence. *Journal of Child Psychology and Psychiatry, 46*, 1255–1268.
Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences of the United States of America, 112*, 12663–12668.
Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by eight-month-old infants. *Science, 274*, 926–928.
Sandin, S., Lichtenstein, P., Kuja-Halkola, R., Larsson, H., Hultman, C. M., & Reichenberg, A. (2014). The familial risk of autism. *Journal of the American Medical Association, 311*, 1770–1777.
Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour, 28*, 1070–1094.
Shepard, K. G., Spence, M. J., & Sasson, N. J. (2012). Distinct facial characteristics differentiate communicative intent of infant-directed speech. *Infant and Child Development, 21*, 555–578.
Shneidman, L. A., & Goldin-Meadow, S. (2012). Language input and acquisition in a Mayan village: How important is directed speech? *Developmental Science, 15*, 659–673.
Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., ... Held, R. M. (2014). Autism as a disorder of prediction. *Proceedings of the National Academy of Sciences of the United States of America, 111*, 15220–15225.
Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences, 22*, 325–336.
Snow, C. E. (1977). Mothers' speech research: From input to interaction. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children* (pp. 31–49). Cambridge, UK: Cambridge University Press.
Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review, 27*, 501–532.
Spinelli, M., Fasolo, M., & Mesman, J. (2017). Does prosody make the difference? A meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes. *Developmental Review, 44*, 1–18.
Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., & Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders. *Journal of Neuroscience, 34*, 691–697.
Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant directed speech facilitates word segmentation. *Infancy, 7*, 53–71.
Thomas, M. S. C., Davis, R., Karmiloff-Smith, A., Knowland, V. C. P., & Charman, T. (2016). The over-pruning hypothesis of autism. *Developmental Science, 19*, 284–305.
Turkewitz, G., & Kenny, P. A. (1982). Limitations on input as a basis for neural organization and development. *Developmental Psychobiology, 15*, 357–368.
Turkewitz, G., & Kenny, P. A. (1985). The role of developmental limitations of sensory input on sensory/perceptual organization. *Journal of Developmental and Behavioral Pediatrics, 6*, 302–306.
Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science, 25*, 1314–1324.
Waterfall, H. R., Sandbank, B., Onnis, L., & Edelman, S. (2010). An empirical generative framework for computational modeling of language acquisition. *Journal of Child Language, 37*, 671–703.
Webb, A. R., Heller, H. T., Benson, C. B., & Lahav, A. (2015). Mother's voice and heartbeat sounds elicit auditory plasticity in the human brain before full gestation. *Proceedings of the National Academy of Sciences of the United States of America, 112*, 3152–3157.
Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology, 34*, 1289–1309.

Wirén, M., Björkenstam, K. N., Grigonyté, G., & Cortes, E. E. (2016). Longitudinal studies of variation sets in child-directed speech. In A. Korhonen, A. Lenci, B. Murphy, T. Poibeau, & A. Villavicencio (Eds.), *Proceedings of the 7th workshop on cognitive aspects of computational language learning* (pp. 44–52). Stroudsburg, PA: Association for Computational Linguistics.

Yingling, J. (1981). *Temporal features of infant speech: A description of babbling patterns circumscribed by postural achievement. Unpublished doctoral dissertation*. University of Denver.

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition, 125*, 244–262.

Zangl, R., & Mills, D. L. (2007). Increased brain activity to infant-directed speech in 6- and 13-month-old infants. *Infancy, 11*, 31–62.