OXFORD LIBRARY OF NEUROSCIENCE

Editor-in-Chief GORDON M. SHEPHERD

Oxford Handbook of Developmental Behavioral Neuroscience

Edited by

Mark S. Blumberg

John H. Freeman

Scott R. Robinson

OXFORD
UNIVERSITY PRESS

2010

CHAPTER

34

From Birds to Words: Perception of Structure in Social Interactions Guides Vocal Development and Language Learning

Michael H. Goldstein *and* Jennifer A. Schwade

Abstract

Infant songbirds and humans face a similar task: to produce a functional repertoire of sounds that operates within the communication system of conspecifics. The mechanisms by which infants learn to talk and birds learn to sing share parallels at the neural, behavioral, and social levels of organization. By making immature sounds and observing the reactions of others, infants of both taxa learn the acoustic structures and temporal contingencies that define communicative interaction. Social and vocal learning, however, are rarely investigated as part of the same system. This chapter uses avian song learning as a model for socially guided vocal learning in human infants. It then discusses socially guided learning in the development of speech and language.

Keywords: vocal development, language learning, songbirds, conspecifics, immature sounds, reactions, acoustic structures, temporal contingencies

Introduction

In a backyard, near the window of a house, a young bird alights on a branch and sings. His song is immature; its structure bears little resemblance to the rapid sequences of clearly articulated notes that he will produce during the breeding season in a few months. A nearby female, however, listens to his infantile sounds and responds with a quick movement of her wings. The male catches her subtle signal, and repeats the cluster of notes that he just sang.

In the house, a 9-month-old infant watches the birds through a window. She utters a stream of babbling sounds. Her babbling could not be mistaken for mature speech, as she has no words in her productive vocabulary, and her sounds do not follow the phonology of the language that surrounds her. Her mother, however, reacts to her babbling by pointing at the singing bird and saying "that's a birdy." As the infant continues to babble, her mother keeps taking vocal turns with her, and over the next few minutes, the infant's sounds become more speech-like, with more resonant vowels and faster consonant–vowel transitions.

Both of these scenes describe powerful moments of learning, in which the responsiveness of social partners to immature behavior is perceived and used by the young learner to generate more advanced forms of vocalization. Such a socially mediated system of vocal learning relies on specific mechanisms of perception and action in both the infant and the adult. Adults' behavior must be sensitive to acoustic characteristics of vocal precursors, and infants must be capable of using adults' feedback to modify their vocalizations. These mechanisms, described in detail below, play strong roles in the development of song, speech, and language.

The development of birdsong and human speech share parallels at the neural, behavioral, and social levels of organization (Doupe & Kuhl, 1999;

Goldstein, King, & West, 2003; King, West, & Goldstein, 2005; Marler, 1970; Petrinovich, 1972; Wilbrecht & Nottebohm, 2003). Though much of the research described below focuses on humans, our theoretical perspective and even some of the experimental paradigms derive from studies of songbird vocal development. We link research on the two taxa by way of an ecological approach to understanding development. An ecological approach emphasizes connections between opportunities for action that the environment affords and the learner's ability to perceive those affordances (Gibson, 1979; Gibson & Pick, 2000). In this view, the behavioral capacities of animals are best studied in the environmental contexts in which they evolved and developed. Thus an ecological approach to learning and development, in both birds and babies, considers developmental phenomena such as the ability to sing or talk as a product of both the young organism and specific environmental opportunities for learning (Green & Gustafson, 1997; Johnston, 1981; West & King, 1985).

We focus on the role of early vocal precursors in the construction of both birdsong and human speech. In our ecological framework, vocal precursors are not just something that an infant has in its possession, rather they get used as instruments of learning. When the responsiveness of social partners is considered along with infants' ability to perceive changes in others' behavior, babbling becomes part of an active process in which infants' actions contribute to their own development. In this view, the possible functional significance of immature song and speech can best be assessed in a social context. Thus we take a microgenetic approach to social learning. In our infant research, we observe and manipulate parent–offspring interactions at moment-to-moment timescales to understand mechanisms of developmental change. By studying social interaction and learning as they occur from moment to moment, we can connect specific mechanisms of perceptual and cognitive development with social influences on the acquisition of speech and language.

Such an approach stands in contrast to the mainstream of research on communicative and linguistic development. Most work focuses on infant capacities or on the nature of the input, but not both *at the same time*. The studies we describe below integrate the perceptual and cognitive mechanisms by which infants pick up information with the structure available in adult reactions to infant behavior. Data from these studies indicate that the development of intelligent behavior is embedded in social processes. Infants, like many organisms, must rely on the brains and bodies of others as an alternative to evolving specific capacities for surviving in a complex environment. Socially distributed intelligence is evidenced by the foraging and nest-building activities of termites, ants, and bees (e.g., Seeley, 1995), in the movements of rat pups as they huddle to collectively thermoregulate (Alberts, 1978; Schank & Alberts, 1998), and in the vocal development of male cowbirds that rely on the visual reactions of females to shape their immature sounds into functional song (West & King, 1988). In our view, caregivers and infants constitute a system of distributed intelligence, one in which maternal behavior and infant sensory capacities interact to construct more advanced infant behavior. Thus we focus on patterns of interaction between caregivers and infants as a source of developmental change. In this chapter, we will start by describing the implications of social learning for vocal development in songbirds and then show what our research has revealed about the role of social learning for communicative development in human infants.

Social Influences on Vocal Learning in Songbirds

In songbirds, strong evidence indicates that the social environment is crucial for vocal development (Beecher & Brenowitz, 2005; Beecher & Burt, 2004). Vocal learning has long been assumed to depend on early memorization of adult song, followed by a period of vocal development in which the young bird attempts to imitate the memorized song (see Marler, 1997 for review). However, there are multiple models of social influences on vocal learning. For example, in several species, socially guided learning is an important mechanism driving the development of complex articulatory patterns and auditory learning that catalyzes production of well-formed, species-typical song. Social learning also facilitates the selective attrition of functional elements from immature vocalizations (Marler & Nelson, 1993). Investigations of animal communication have shown that responses of conspecifics to the vocal precursors of young songbirds influence developmental outcomes and thus constitute a social mechanism of vocal learning (Doupe & Kuhl, 1999; Marler & Nelson, 1993; West & King, 1988).

An instructive example of the role of immature vocalizations in the socially guided learning of song comes from studies of the brown-headed cowbird (*Molothrus ater*). Cowbirds are brood parasites; the females lay their eggs in the nests of other species, so a young cowbird is not likely to hear the songs of conspecifics early in life (Friedman, 1929). Despite their lack of early species-specific experience, males sing to females during the breeding season (May–June) and a mature song is necessary for successful courtship and copulatory success (Chapter 33). Thus cowbirds were long assumed to have an innate, unmodifiable program of song development (Mayr, 1974). The reasons for such a closed developmental program are easy to understand: Since song is necessary for reproductive success, there is too much at stake for species survival to leave song development to the vagaries of individual experience, especially in a brood parasite. A cowbird being reared by another species might learn the wrong song or court the wrong species.

Evidence that learning played a crucial role in song production came from studies that focused on the early stages of song development. The ontogeny of song in cowbirds is gradual, lasting from 2 to 12 months of age, and can be broken down into several stages (Smith, King, & West, 2000; see also Chapter 33). Subsong, the initial stage, is characterized by low amplitude and high variability in structure and timing. Plastic song, beginning at approximately 60–75 days of age, is composed of poorly articulated notes that are sung in varying, unstable sequences. Formatted song, beginning at approximately 200 days, has elements with stable ordering and timing, but with variable content. Finally, stereotyped song typically develops at 250–300 days. These are mature, stable songs, and a male's repertoire typically consists of 2–6 song types (Rothstein, Yokel, & Fleischer, 1988).

To assess the relative influences of nature and nurture on song production, West and King (1988) housed individual males over the winter and spring in 1.3 m^3 sound attenuating chambers. Each male was housed with a female, so the males would have a social companion but would still live in acoustic isolation, as females do not sing. Results showed that being housed with a female was a vocally enriching experience. Males housed in acoustically impoverished conditions developed song that had high potency, as measured by females' responses to playbacks of the song (West & King, 1988). What were the females doing to create males with high song potency? Video analyses showed that the females were using small, infrequent wing movements (called wingstrokes) to selectively reinforce elements of the male's plastic song. Females produced wingstrokes either during or immediately after songs, which often were 1 s in duration. Wingstrokes were typically under 200 ms in duration and were infrequent, occurring once for every 100 songs. Males that received a wingstroke were, on average, three times more likely to repeat the element that they just sang (West & King, 1988).

Were the females responding to more advanced vocalizations, thus shaping the males in a developmentally advanced direction? To test the functional value of wingstrokes, West and King took immature songs from the early spring that preceded, elicited, and followed wing strokes, and played them back to a new set of females. They used a female bioassay procedure (West, King, Eastzer, & Staddon, 1979) in which females during the breeding season respond with copulatory postures to immature songs, as long as those songs contain some functional elements. They found that females were most responsive to plastic songs that had earlier elicited wing strokes. In addition, songs that followed wingstrokes received higher responding than songs that preceded wingstrokes (West & King, 1988).

Thus the immature vocalizations of young male cowbirds elicit orientation and wing movements from adult females, and this feedback facilitates the development of more advanced forms of song (King et al., 2005; Smith et al., 2000; West & King, 1988). The functional significance of plastic song only became apparent in social contexts; the value of early vocalizing would not have been discovered if the males had been studied in isolation. Later studies showed that females from different regions, with differing levels of responsiveness, create differences in the acoustics of plastic song, the rate of song development, and the potency of song (King et al., 2005; Smith et al., 2000; West & King 1988). A focus on the functional significance of vocal precursors has been successfully employed in studies of vocal learning in multiple species (Marler & Nelson, 1993; Snowdon & Hausberger, 1997; West, King, & Duff, 1990).

Socially Guided Vocal Learning in Human Infants

Those who study the vocal precursors of songbirds and human infants face similar challenges. Like the subsong and plastic song of songbirds, the early noncry vocalizations of human infants are highly variable and immature in acoustic form. In both taxa, mechanisms of social learning have often taken a backseat to a focus on innate templates guiding vocal development. When social learning has been found, processes of imitation are usually invoked to explain the results. Our work on vocal learning and development in human infants was a direct result of the cowbird work described above. If vocal precursors in songbirds could have functional significance in social settings, the babbling of human babies might also be a strong force in socially guided vocal learning.

Infant sounds have been extensively studied from a taxonomic perspective, in which vocal development has been described in terms of universal and invariant stages (Oller, 1980, 1985, 2000; Oller & Lynch, 1992; Stark, 1979, 1980). The aim of these studies was not to consider the effects of social context or other environmental influences on vocal production, but rather to map out the species-typical course of vocal development. In our vocal learning studies, we quantify babbling using Oller's categories of prelinguistic production, as they describe early sounds in terms of their "infraphonology" or the acoustic prerequisites of well-formed speech. The infraphonological acoustic classification system incorporates both acoustic parameters (e.g. fundamental frequency [F_0], formant transitions) and qualitative descriptors (e.g. phonetic categories). From 1 to 4 months of age (primitive articulation stage), vocal production incorporates articulator movements at the back of the vocal tract using tongue and epiglottal contact with the soft palate to produce some consonant stops. Combinations of stops with quasi-resonant nuclei produce the first "goo"- and "coo"-type syllables. At 3–8 months (Expansion/Exploratory stage), the vocal tract becomes more open and fully resonant sounds are produced. Infants produce marginal syllables, which are slow sequences of consonant–vowel articulation, with long transitions between consonant and vowel. From 5 to 10 months (Canonical Syllable stage), infants begin to produce fully resonant sounds and faster formant transitions, resulting in canonical syllables (e.g., [ba], [da]), a language-general unit of mature vocal production (Figure 34.1).

What are the underlying mechanisms mediating transitions to new vocal forms? At the level of the articulators, there are constraints imposed by the developing vocal tract. The nasalized quasi-resonant vocalizations of the primitive articulation stage are

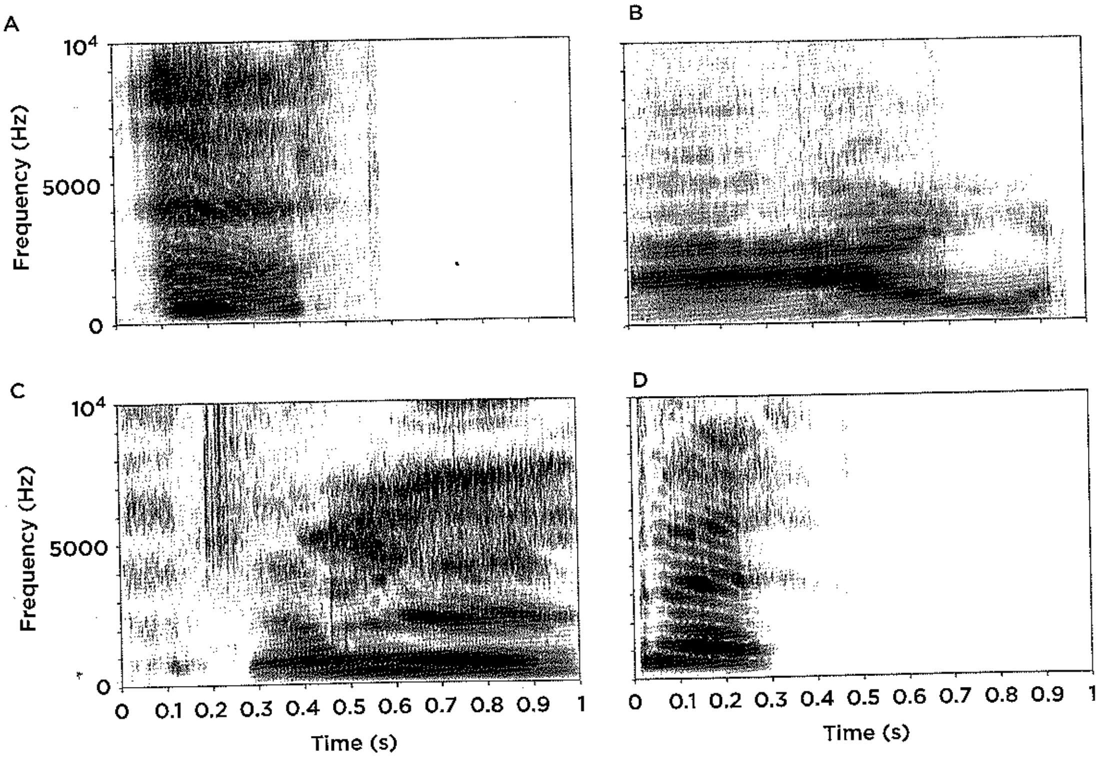

Figure 34.1 Spectrograms of the four vocal types. (A) Quasi-resonant vowel. (B) Fully resonant vowel. (C) Marginal syllable. (D) Canonical syllable.

due to the overlapped larynx and nasopharynx at the back of the vocal tract, which forces infants to breathe nasally. Also, the small size of the oral cavity restricts tongue movement (Kent, 1981a, 1981b). As the larynx descends into the throat, the fully resonant vowels of the expansion/exploratory stage can be produced. These physiological constraints are most important in the first months of life. By approximately 7 months of age, a new type of articulatory pattern, rhythmic jaw oscillations, emerges and permits reduplicated canonical syllables (Iverson, Hall, Nickel, & Wozniak, 2007; Iverson & Thelen, 1999; MacNeilage, 1998; MacNeilage & Davis, 2000; Meier, 1997). MacNeilage attributes the oscillatory pattern to evolutionary forces that shaped the physiology and neuroanatomy of jaw control, while Iverson shows that rhythmic behavior is a global, whole-organism event, occurring in limbs at the same time as the jaw (Iverson et al., 2007).

While the presence of rhythmic, reduplicated vocalizations is a ubiquitous feature of vocal development after 7 months of age, there are also large individual differences in early vocal development. For example, there are large age overlaps across the stages of vocal development discussed above. Data from deaf and hearing-impaired infants reveal that their babbling is acoustically different from that of hearing infants from 6 months onward, with abnormalities in marginal syllables and a marked delay in onset of canonical syllables (Locke, 1993; Locke & Pearson, 1992; Oller, 1985; Oller & Eilers, 1988; Oller, Eilers, Bull, & Carney, 1985; Stoel-Gammon & Otomo, 1986). These studies suggest that typical vocal development requires environmental input.

The phonology of prelinguistic vocalizations also shows effects of the ambient language on early vocalizing. The "babbling drift" hypothesis emphasizes the influence of the ambient language on babbling. By approximately 8 months of age, several suprasegmental features of vocalizing have come to resemble that of the ambient language. For example, the intonation patterns (rising and falling pitch) and rhythm (final syllable lengthening) of 7–11-month-old French and English infants resemble those of their ambient languages (Levitt & Wang, 1991; Whalen, Levitt, & Wang, 1991). Segmental features, such as the patterns of vowels and consonants that comprise syllables, are also influenced by experience. Nigerian infants are exposed to Yoruba, a language in which many words begin with a vowel. The babbling of 10- to 12-month-old Nigerian infants has a typical disyllable structure of vowel–consonant–vowel (VCV), in contrast to the typical CVCV disyllable pattern found in the babbling of French infants (Boysson-Bardies, 1993). In comparisons of French, English, Japanese, and Swedish infants, the modes of consonant production are similar in distribution to that of the ambient language (Boysson-Bardies & Vihman, 1991), and comparisons of French, English, Chinese, and Algerian 10-month-old infants has revealed that the distribution of vowels in their babbling is similar to that of the ambient language (Boysson-Bardies, Halle, Sagart, & Durand, 1989). Nine-month-old infants can change their babbling to reflect sound patterns in their mothers' speech within minutes (Vihman & Miller, 1988). Though these studies reveal a gradual shift toward the phonology of the ambient language, they do not shed light on the nature of the interactions that lead to phonological learning.

The ambient language also exerts a strong influence on speech perception. Studies of environmental influence on speech perception yield clues as to the mechanisms that might be involved in vocal learning. By 8–10 months of age, infants form categories corresponding to the phonemic contrasts of their language environment (e.g., Best, 1994, 2002; Jusczyk, 1992; Werker & Tees, 1984). Social interaction facilitates the development of categorical speech perception (Kuhl, 2007a, 2007b; Kuhl, Tsao, & Liu, 2003). Social feedback may focus attention on relevant features of the speech signal and/or provide additional sources of information that specify perceptual categories (Kuhl, 2007a, 2007b). Categorical speech perception is negatively impacted when infants attend to speech to learn word–object associations (Fennell, Byers-Heinlein, & Werker, 2007; Stager & Werker, 1997; Werker & Curtin, 2005). Social interaction may be beneficial for speech perception because its attention-organizing effects buffer the increased cognitive load of word learning. Social interaction may benefit vocal learning in similar ways.

Given the important role of language experience in the development of speech perception and production, what mechanisms incorporate the phonology of the ambient language into babbling? Infants develop in a responsive, structured social environment, and can perceive regularities in others' behavior (Baldwin, Baird, Saylor, & Clark, 2001). Structured and consistent social responses to babbling, if present, can create opportunities for infants to learn from the consequences of their vocalizing and acquire an understanding of the contingencies defining communicative interaction. Socially facilitated patterns of vocalizing can entrain the underlying motor movements, creating stable articulatory patterns. Practice can entrain complex motor systems. For example, patterns of interlimb coordination that comprise walking are facilitated by practice at stepping on a treadmill (Thelen, 1986; also see Chapter 9). Consistent social feedback may be an important control parameter in the organization of vocal behavior.

Social Responses to Early Vocalizations

Guided by the cowbird work, we first investigated the amount and type of social reactions that occur in response to a babbling infant. For socially guided learning to occur, adults' behavior must be sensitive to acoustic characteristics of babbling, and infants have to be capable of using feedback from caregivers to modify their vocalizations. Throughout the first year, infants typically vocalize during parent–child interactions, and these sounds elicit responses from caregivers. During play sessions with their 2- to 5-month-old infants, mothers often matched their infants' vocalizations and did so accurately (Papoušek, 1991; Papoušek & Papoušek, 1989). When engaged in face-to-face interactions, mothers and infants take turns vocalizing (Anderson, Vietze, & Dokecki, 1977; Papoušek, Papoušek, & Bornstein, 1985). The synchrony of vocal turn-taking is regulated by both mother and infant, with each partner's behavior predictive of the other's (Jaffe, Beebe, Feldstein, Crown, & Jasnow, 2001). Further evidence for the role of social interaction in vocal development is provided by social environments that do not produce organized patterns of responses to babbling. For example, when hearing infants of deaf parents had little systematic exposure to spoken language, they produced vocalizations that were acoustically disorganized and developmentally delayed (Petitto, Holowka, Sergio, Levy, & Ostry, 2004).

Caregivers are sensitive to infant sounds from early in development, as shown by multiple playback studies in which adults view audio and video recordings of infant behavior. Crying causes rapid responses from both parents and nonparents (Green & Gustafson, 1997; Gustafson & Harris, 1990; Wood & Gustafson, 2001). Noncry prelinguistic vocalizations also influence caregivers' behavior, causing reliable changes in adults' perceptions of infant behavior (Beaumont & Bloom, 1993; Bloom, D'Odorico, & Beaumont, 1993; Bloom & Lo, 1990; Papoušek, 1989). In naturalistic play sessions, caregivers are more likely to respond to vocalizations that contain more speech-like elements, such as full resonance and CV-structured syllables (Gros-Louis, West, Goldstein, & King, 2006; Hsu & Fogel, 2003; Keller & Scholmerich, 1987). When canonical syllables first emerge, parents recognize that their infants' babbling has become more speech-like (Oller, Eilers, & Basinger, 2001).

All previous playback studies of adults' responses to babbling, however, had participants code or rate infants' behavior, rather than recording their own behavioral responses to the infants. Because the adults' functional responses were not considered, the ratings employed by these studies do not directly reveal how caregivers' responses could be part of a system of vocal learning. There has been only a single published study of social responsiveness to the prelinguistic vocalizations of infants between 6 and 12 months of age (Goldstein & West, 1999). In this study, we assessed the sensitivity of mothers' responses to the behavior of *unfamiliar* infants. We used a playback paradigm to measure the relative salience of infants' vocal and visual cues. Stimuli were taken from prerecorded episodes of prelinguistic behavior from three infants with vocal repertoires of different sizes. The audio and video components of each infant's reference episodes were digitally separated and permuted to create new combinations of cues. Mothers viewed both the original and the recombined episodes. Participants responded to each episode by choosing one of four possible action categories (e.g., "I think the baby wants something").

Mothers' responses were reliably similar. The examples of prelinguistic infant behavior thus contained sufficient information to guide mothers' reactions toward consensus as to communicative content. The similarity of responses across participants suggests a social mechanism of reliable feedback for infant behavior. The communicative potency of prelinguistic vocalizations was demonstrated by playback mothers' responses to the recombined stimuli. As the vocal repertoire size of the stimulus infants increased, changes in vocalizations resulted in more shifts in playback mothers' responses. When exposed to the two infants with larger and more complex vocal repertoires, mothers shifted their responses when the audio component was changed. Thus, when presented with behavior of sufficient acoustic diversity, the playback mothers' responses were sensitive to changes in prelinguistic sounds (Goldstein & West, 1999). Feedback from caregivers is consistent enough to provide reliable cues about the consequences of vocalizing.

We are currently examining the role of caregiving experience in responding to prelinguistic vocalizations. As infants become more skilled, caregivers increase the complexity of their interactions with their infants (e.g., Bruner, 1983). This process is called "scaffolding" because parents provide age-appropriate support for their infants' emerging cognitive skills. We hypothesize that caregivers will show evidence of "acoustic scaffolding," in which caregivers raise their thresholds of responsiveness as a result of experience with more developmentally advanced forms of infant vocalizations. Caregiver responsiveness should thus be matched to infants' current level of vocal development. Infants aged 3–6 months primarily produce quasi-resonant and fully resonant vowels and have begun producing marginal syllables. Based on our acoustic scaffolding hypothesis, we predict that mothers of 3- and 6-month-olds will be more responsive to quasi-resonant and fully resonant vowels and marginal syllables. Although infants aged 9–12 months produce some quasi-resonant vowels and marginal syllables in addition to canonical syllables, mothers of 9- and 12-month-olds should be most responsive to canonical syllables, as they represent the infants' most advanced vocal form. Thus we anticipate interactions between caregiving experience and responsiveness to changes in vocalizations.

Infants can use the information available in the social environment to structure their vocalizations. When exposed to face–voice matches and mismatches, 3- to 4-month-old infants were more likely to imitate vowel sounds when the face and voice were matched (Kuhl & Meltzoff, 1996; Legerstee, 1990). Through the experience of producing prelinguistic vocalizations, the infant learns to map its acoustic output with articulatory movements. Infants may then use the visual information provided by the articulatory movements of others to assemble a visual/auditory mapping of speech sounds (Masataka, 2003). In adults, comprehension of spoken speech is facilitated when the speaker's face is visible (e.g., Schwartz, Berthommier, & Savariaux, 2004); the use of visual information in speech processing may have its roots in infant vocal production.

Sensitivity of Prelinguistic Vocalizing to Social Behavior

Social Shaping of Babbling

The playback studies above were designed to document the capacity of the social environment to provide feedback to infants' sounds. The next step was to ascertain the role of social feedback in vocal development. Are the reactions of caregivers to babbling perceived and used by infants to refine their vocal repertoires? Social contingencies influence vocalizations and generate possibilities for trial and error learning of phonology during the prelinguistic period (Locke, 2001; Masataka, 2003). Studies of contingent versus noncontingent social reinforcement have demonstrated robust effects of receiver responses on vocal production. Contingent vocal turn-taking causes 3-month-olds to produce vocalizations that are less nasalized and more fully resonant (Bloom, Russell, & Wassenberg, 1987; Masataka, 2003). Infants increase their rate of babbling when their sounds are reinforced by touches from adults and shakes of a rattle (Poulson, 1983; Rheingold, Gewirtz, & Ross, 1959; Routh, 1969; Weisberg, 1963). These conditioning studies indicate that social reinforcement may facilitate vocal development. However, these studies shared a set of limitations: the infants were very young, before the age at which many different types of sounds are produced, and the infants were in chairs or cribs, with few degrees of freedom for movement. In essence, they were locked into a social interaction.

What would be the course of vocal learning if older infants, with a diversity of sounds in their repertoire, were tested with all of the degrees of freedom that they would naturally have? We tested 30 mothers and their 8- to 10-month-old infants in a new type of vocal learning paradigm (Goldstein et al., 2003). Mothers and infants played in a large (12 × 18 ft) room that allowed infants to become distal from their mothers if they so desired. Infants were able to explore the room and play with toys independently of their mothers. Mothers wore wireless headphones, allowing them to receive instructions from an experimenter. Infants wore denim overalls that held a wireless microphone and transmitter so we could obtain high-quality recordings of their sounds (Figure 34.2). Dyads came to the laboratory for two 30-min play sessions, spaced 24 h apart (Figure 34.3).

The first "familiarization" session allowed mothers and infants to get used to the playroom and its toys. The second play session, an ABA design, was divided into 10-min periods of baseline recording, social response, and a second baseline period. During both baseline periods, all mothers were instructed to play as they would at home. During the social response period, half of the mothers (the "Contingent" condition) were instructed to immediately respond to their infants' sounds by smiling, moving closer, and touching them. The other half (the "Yoked Control" condition) were instructed to react based on response schedules that had been generated by the Contingent mothers. Yoked Control mothers' responses were thus linked to those of the Contingent mothers so that the Control infants received the same amount of social stimulation as the Contingent infants, but the stimulation was not synchronized with their vocalizations. In both conditions, mothers were asked not to speak to their infants when responding. Thus the infants were not presented with vocal material with which to imitate.

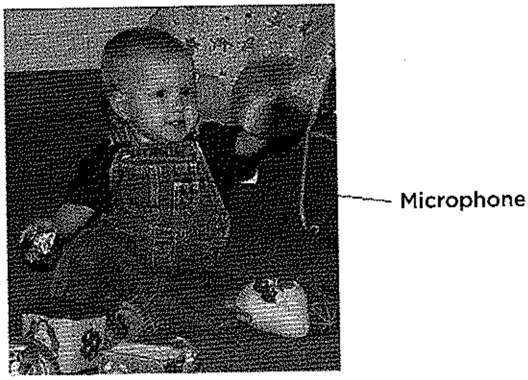

Figure 34.2 An infant dressed in the overalls containing the wireless microphone and transmitter.

The babbling of Contingent infants changed dramatically across the experimental periods (Goldstein et al., 2003). Relative to Baseline-1, the vocalizations produced during the social response period were characterized by increased voicing and faster consonant–vowel transitions, which are developmentally more advanced forms of production (Table 34.1). Infants continued to produce these new vocal forms during Baseline-2. In contrast, vocalizations of the Yoked Control infants did not change in form across the three periods. Only contingent social feedback created changes in babbling. Why did infants need contingency? The effect occurred because contingency is a rewarding event (Watson, 1972). Increased maternal contingency caused increased arousal in these infants; we observed louder and more resonant sounds, and more repeated and rhythmic articulatory movements that created sounds with CV form. As the social response period progressed, all new sounds were reinforced. Infants must therefore recognize that their sounds produce a change in the environment in order for their sounds to mature. These results demonstrate real-time infant vocal learning from social feedback.

Infants are capable of learning from many sources of contingent feedback (e.g., Watson, 1972). A recent study from our laboratory tested the effects of contingent social feedback for

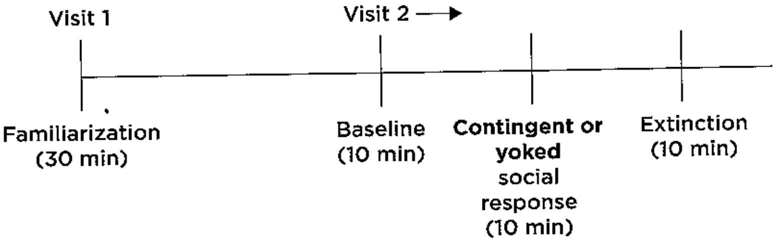

Figure 34.3 Timeline of vocal learning paradigm.

Table 34.1 Vocal Learning as a Function of Contingency and Caregiver Speech

	Contingent-Response Conditions			Yoked Controls
	Silent	Resonance	CV	All Groups
Proportion of vocalizations with full resonance	*	*	ns	ns
Proportion of vocalizations with CV structure	*	ns	*	ns
Proportion of vocalizations with canonical form	*	ns	ns	ns

Note: Silent group from Goldstein, King, and West (2003); Resonance and CV groups from Goldstein and Schwade (2008).

vocalizing separately from the effects of animacy (Narayan, Goldstein, & Schwade, 2008). To do so, we replicated Goldstein et al. (2003), substituting a nonanimate interaction partner—a remote-controlled car—for the infants' caregivers. During the familiarization session, infants played in a playroom with their caregiver as they became familiarized with the car's movements. To simulate maternal responsiveness during unstructured play, the car responded to approximately 20% of infant vocalizations during familiarization and the two baseline periods. After one in five babbles, the car responded to the infant by approaching the infant, then retreating. During the social response period, the car responded to all babbles for infants in the Contingent condition. For infants in the Yoked Control condition, the car responded during social response according to a yoked schedule from a randomly selected Contingent infant.

Vocal quality (vowel resonance and proportion of canonical syllables) significantly increased across the test session, but only for the Contingent infants. The form of vocal learning was similar overall to the learning shown when interacting with contingent caregivers (Goldstein et al., 2003). In addition, the number of vocalizations produced by Contingent infants during the social response period was positively correlated with improved vocal quality. This is because the number of vocalizations that they produced determined the number of opportunities they had to learn the connection between their own vocalizing and the movement of the car. In contrast, the number of vocalizations and amount of vocal learning were not correlated for Yoked Control infants. Thus, regardless of animacy, infants only learn when the social partner provides contingent feedback for babbling. Vocal learning is more sensitive to the temporal structure of social interaction than to the physical attributes of the interaction partner.

Socially Guided Learning of Phonological Rules

The study above showed that contingent silent feedback facilitated vocal learning. Most of the time, however, caregivers also respond vocally to their infants babbling. In this study, we manipulated the phonological structure present in caregivers' speech as they reacted to their infant's babbling (Goldstein & Schwade, 2008). We used an ABA design similar to the study above. Again we asked half the mothers (Contingent condition) to immediately respond to their infants' sounds by smiling, moving closer, and touching them. These Contingent mothers were divided into two groups. Mothers in the Resonance group produced fully resonant versions of their infants' babbling as they responded. Mothers in the Consonant–Vowel (CV) group were asked to interpret each babble as if it were a word, and to say that word back to their infants as they responded. Thus infants were given acoustic models of specific parameters of well-formed speech. The other half of the mothers were assigned to the Yoked Control (YC) conditions. There were two YC groups, matching the Resonance and CV groups. Mothers in the Resonance and Consonant-Vowel YC groups responded using the same utterances as the Contingent mothers with whom they were paired, and the timing of YC mothers' responses was governed by the Contingent mothers. To do this, we created a digital file of the contingent utterances, removed extraneous sounds while preserving the timing of speech, and created CD tracks to play back to YC mothers over wireless headphones. The YC mothers were asked to repeat the vocalizations that they heard over their headphones while getting closer to, touching, and smiling at their infants.

The form of mothers' responses influenced how their infants' vocalizations changed. Relative to the initial baseline, the vocalizations of infants in the Resonance and CV groups changed in specific ways (Table 34.1). In the Resonance group, infants' sounds became more fully resonant. In the CV group, infants increased the proportion of syllables with CV structure. These changes were maintained throughout the second baseline, even after the rate of vocalizing had decreased to initial baseline levels. In contrast, vocalizations of the YC infants did not change in form across the three periods, though they increased in amount. The increase in vocalizing was likely due to the increase in noncontingent reinforcement caused by mothers' unpredictable responses. Noncontingent responding did not lead to greater attrition or infant fussiness.

Since the phonological patterns of infants in the contingent groups were similar to those of their mothers, we next analyzed the amount of infant imitation of specific phonemes. We did so by comparing phonetic transcriptions of each maternal utterance with the phonemes in the following infant vocalization (see Goldstein & Schwade, 2008, for details). To compute a chance level of vocal matching, we shuffled the sequences of vocalizations to create a random distribution of mother–infant vocal pairings. Amount of matching in the original transcripts was compared against amount of matching in the scrambled transcripts. We found that the amount of infant vocal imitation was not different from chance levels. The kinds of sounds that infants produced shared phonological rules but not phonetic characteristics with their mothers' speech. This means that vocal learning is a nonimitative process.

Thus the form and timing of maternal behavior influenced the learning of specific acoustic parameters of babbling. Changes in the sounds mirrored acoustic changes that occur during development. Infant imitation is not a likely explanation for these data. Though these infants produced sounds with more resonance or CV syllables, they did not produce the same phonemes their mothers modeled. Because mothers produced a diversity of phonemes, the underlying infraphonological patterns became more salient. Previous research on statistical learning, using an auditory discrimination task, has shown that 9-month-old infants can rapidly learn to recognize phonological patterns (Gerken, 2006; Saffran & Thiessen, 2003). In an artificial grammar learning task, greater variability in the input results in infants learning higher-order patterns across nonadjacent elements in the input (Gómez, 2002). We believe that variability in the input plays an important role in determining the speech patterns that infants incorporate into their vocal repertoires, as infants who receive repetitions of the same utterance produce sounds that imitate the input (Kuhl & Meltzoff, 1996). If infants' prelinguistic vocalizations are guided by statistical learning, input consisting of repetitions of a single syllable would lead to infants' learning of the phonetic surface features rather than learning the underlying CV structure. As a result, infants would not have a phonological framework with which to generalize to new combinations of sounds. Thus vocal production resulting from repeated exposure to a single syllable would not show evidence of generalization beyond the phonemes the infants heard, yielding sounds that seemed imitative but were really the output of a statistical learning mechanism that is operating on input with no variability. Current studies in our laboratory are systematically manipulating the amount of phonological and phonetic variability in the input to understand the role of variability in vocal learning.

Infants in the YC Resonance and CV groups, who were exposed to the same maternal vocalizations and behaviors as Contingent infants, did not show changes in vocal quality. These data suggest that the timing of feedback is important, as mere exposure to their mothers' utterances was not enough to get infants to change their sounds. Why is contingency crucial for vocal learning? In studies of statistical learning in speech perception, infant learning does not require contingent presentations of stimuli. There are several reasons why contingency is important for learning new forms of vocal production. The learning task is more difficult because infants are creating, rather then recognizing, utterances that obey phonological rules. Thus focused attention and increased arousal, which may be facilitated by social contingency (Kuhl, 2007a, 2007b), should have stronger effects on tasks in which the infant is required to produce as well as to recognize new phonological patterns. Contingent vocal responses also tend to be prompt (within 5 s), which means that the utterances of infants and adults are contiguous in time. Such temporal proximity may facilitate infants' detection of differences between their own and adults' utterances. Computer models of unsupervised language learning rely on the perception of utterances that are partially redundant to learn grammatical structure (e.g., Solan, Horn, Ruppin, & Edelman, 2005). For these models, temporal contiguity between utterances would enhance the distillation of structure in the input. We are currently testing the role of contingency by manipulating the promptness of social feedback in new studies of vocal learning.

In summary, infants learned new patterns of vocalizing by recognizing phonological patterns in caregivers' contingent speech and using those patterns to restructure their own babbling. Infants can learn from social feedback as they vocalize; vocal learning is guided in real time by phonological patterns in caregivers' speech. Thus the activity of babbling, in combination with caregivers' responsiveness to vocal precursors, creates a process of socially guided vocal learning. Taken together, our studies support socially guided learning as an important mechanism in early vocal development. Prelinguistic vocal learning may lay the foundation for later advances in language acquisition, as discussed below.

Role of Social Feedback in the Transition from Babbling to Words

How do infants learn that their sounds map onto objects in the world around them? Observations of infants in our laboratory have revealed new patterns of infant actions and social reactions that accompany babbling. These patterns of interaction provide a source of nonimitative learning of correspondences between sounds and objects. Infants often babble while looking at or manipulating an object, and we are currently studying the effect of these "object-directed" sounds on the behavior of caregivers. By investigating the effects of directed versus undirected babbling on caregiver behavior, we can elucidate the earliest social interactions that support the association of sounds and referents.

Maternal Responses to Babbling are Related to Later Lexical Development

Patterns of parental responsiveness to infant behavior have been linked to long-term language development (e.g. Hart & Risley, 1995; Tamis-LeMonda & Bornstein, 2002; Tamis-LeMonda, Bornstein, & Baumwell, 2001). Social interaction clearly plays an important role in language development, but what role does responsiveness to babbling play in later language learning? There is strong acoustic continuity between babbling and early words, so that socially guided learning during the babbling phase may carry over into early lexical development. Observations of language development indicate that babbling blends into word production (Stoel-Gammon, 1992; Vihman, Macken, Miller, Simmons, & Miller, 1985). The onset of first words is correlated with the onset and number of syllable types of canonical babbling (Stoel-Gammon, 1992).

Given prior demonstrations of the sensitivity of babbling to caregiver feedback, what are the effects on language of different styles of caregiver responses to babbling? Mothers and their 9-month-old infants were recorded as they played together in our large playroom. The room contained toys and pictures. Mothers were asked to play as they would at home. We coded and categorized several styles of caregiver behavior based upon their contingent verbal responses to the babbling of their 9-month-old infants (Table 34.2). We then assessed the long-term effects of those response styles on early vocabulary acquisition at 15 months.

We focused on mother's responses to their infants' object-directed vocalizations (ODVs) during naturalistic play. An ODV is one that is emitted when the infant is looking at an object that is within reach or being held. We found several stable categories of contingent verbal responses to ODVs. We then used the MacArthur Communicative Development Inventory (CDI; Fenson et al., 1994) to examine vocabulary when infants were approximately 15 months old (mean age 15;11, range 14;11–16;25). The CDI is a parental report measure of productive and receptive vocabulary that has been validated with laboratory measures of vocabulary development (Bornstein & Haynes, 1998). In addition, the CDI shows long-term stability in predictive value for later language and cognitive development (Fenson et al., 2000).

We obtained correlations between two categories of maternal response to babbling and later vocabulary. We found a significant positive linear association between the proportion of *proximal object labeling* and later productive vocabulary, $r(11) = .81$, $p < .01$. A proximal object label occurs when the mother says the name of an object that infant was touching or looking at when he/she vocalized. We also found a negative curvilinear relationship between proportion of *phonological resemblance* responses and later receptive vocabulary, $r(11) = -.68$, $p < .01$. A phonological resemblance response occurs when the mother utters a word that sounded like the infant's vocalization, for example, infant utters "ba," for which the mother says "ball."

Table 34.2 Categories of Mothers' Verbal Responses to Infant Vocalizations

Maternal Response	Description
Proximal object label	Name of object that infant was touching or looking at when he or she vocalized, e.g., "cup," "ball"
Phonological resemblance	Word that sounded like the infant's vocalization, e.g., infant utters "da," mother says "dog", but no dog is present
Conversational placeholder	An utterance that does not convey information, e.g., "uh-huh," "oh really"
Internal state	Description of infant's emotional state or needs, e.g., "happy," "up," "thirsty"
Imitation	Mother imitated her infant's vocalization
Object/property descriptor	Word describing an aspect of an object that the infant was touching or looking at when s/he vocalized, e.g., "red," "big," "round"
Object action	Word describing an action could be done on an object that the infant was touching or looking at when s/he vocalized, e.g., "roll," "throw," "eat"
Other	Other vocal maternal responses, e.g., laughing

Overall maternal responsiveness at 9 months was not significantly correlated with either vocabulary measure at 15 months [comprehension r (11) = .06, p = .86, production, r (11) = .39, p = .19].

Thus mothers' spontaneous verbal responses to their infants' ODVs were related to their children's later vocabulary. Mothers who used more proximal object labels in response to their infants' ODVs at 9 months had children with larger vocabularies at 15 months. Mothers who responded with more phonologically similar (but contextually irrelevant) words at 9 months had children with smaller vocabularies at 15 months. Other response types (e.g., conversational placeholders) and overall amount of maternal responsiveness were not related to later vocabulary development.

What factors explain the relationships between social responses to ODVs and vocabulary? Mothers who respond to their infants' ODVs by providing object labels may facilitate word learning by helping their infants recognize the connection between sounds and objects present in the environment. When mothers label the objects their infants are attending to in response to their babbling, they provide structure that may promote associative learning of word–object pairings. In contrast, mothers who react with words that are similar to their infants' sounds, but unrelated to the context of infants' ODVs, may be inhibiting later word learning because they are providing inconsistent word–object pairings.

Mothers who provide labels for nearby objects may also be more verbal, or interact with their infants in other ways that facilitate infants' language development. Mothers' verbal skills are a significant predictor of their children's language development (e.g., Bornstein, Tamis-LeMonda, & Haynes, 1999; Hart & Risley, 1995; Hoff-Ginsberg, 1986). Mothers who respond to their infants' ODVs with an object label may also be more likely to later report that their infant understands more words and produces more words. However, the CDI correlates with laboratory and observational measures of vocabulary for children in this age range (Fenson et al., 1994). Given the CDI's validity and the strong negative correlation between phonological resemblance responses and later vocabulary, we believe the positive correlation between proximal object labeling and child vocabulary arises from the opportunities for associative learning created when mothers label objects in response to ODVs.

Though the manner in which caregivers responded to babbling predicted later language development, the overall quantity of mothers' responsiveness does not. This finding stands in contrast to the results of previous research (e.g., Tamis-LeMonda & Bornstein, 2002). Perhaps the reason that higher maternal responsiveness has been correlated with positive language outcomes in previous studies is that increased responding in general is related to more responses to ODVs. The present research focused on mothers' responses to ODVs rather than to all infant vocalizations. Narrowing the type of vocalizations considered has uncovered more specific underlying relations between maternal responsiveness and later language. Future longitudinal studies relating caregiver responsiveness to language learning could be informed by including caregiver behaviors that relate to mechanisms of real-time infant learning.

This study represents a first step in establishing links between prelinguistic vocal development and early word learning. The link is illustrated only by examining caregivers and infants together, as a system. Specific forms of feedback from caregivers may provide reliable cues about the relationship between sounds and objects, thus serving as a source of learning for infants, facilitating the transition between babbling and words.

Perceptual Learning While Vocalizing

What infant learning mechanisms explain the correlations between maternal responses to ODVs and lexical development that were found in the study above? In combination with our vocal learning studies, the findings suggest that prelinguistic vocalizing occurs when an infant is in a state conducive to learning. We hypothesize that ODVs signal a state of focused attention, so that maternal responses to those sounds occur when an infant is in an optimal state for learning the properties of an object. In addition to learning about the speech acoustics of mothers' contingent vocal responses to their vocalizations, infants may also be learning about the objects at which they vocalize. To test this idea, we investigated whether infants learned more about the visual features of the objects to which they were attending when they produced ODVs (Briesch, Schwade, & Goldstein, 2008; Goldstein, Schwade, Briesch, & Syal, 2009). Twelve-month-old infants explored a set of 12 novel objects, each for 40 s. For each infant, we determined the object that elicited the highest (HV) and lowest number (LV) of vocalizations. In a subsequent preferential looking task, each infant was presented with pictures pairing the original HV and LV objects with shape-distorted versions of each (6 s trials, 2 trials each, counterbalanced for side and order). We predicted that infants would learn more about the objects that elicited vocalizations than about objects that did not.

Infants looked significantly longer at the distorted version of the HV object than to the original. In contrast, infants did not differ in looking times to distorted versus original LV objects. We interpret the looking times to mean that infants learned the visual features of the HV object, thus their attention was drawn to the distorted, novel version during the test trials. Infants did not learn the features of the LV object, so the familiar and distorted versions were equally interesting during the test trials. The amount of looking and handling of objects during object exploration was not different for HV and LV objects (Briesch et al., 2008; Goldstein et al., 2009). In a second experiment, we extended the findings to word learning by investigating the effect of labeling novel objects contingently on infant vocalizing or on infant looking (Goldstein et al., 2009). Twelve-month-old infants learned associations between novel words and objects more readily when the words were presented contingently on an ODV. Data from these studies have thus linked prelinguistic vocal learning and language learning, phenomena that are usually studied as separate entities.

Thus prelinguistic infants might be especially sensitive to perceptual information just after vocalizing. Mothers who respond contingently to babbling do so at exactly the time when their infants are ready to learn and remember new information. Specific forms of feedback from caregivers may provide reliable cues about the relationship between sounds and objects, thereby serving as a source of information for infants, facilitating the transition between babbling and words.

Socially Guided Learning of Words and Syntax

As infants' vocal, social, and cognitive skills develop, so do opportunities for socially guided learning of more advanced forms of communication. In word learning outside of laboratory experiments, children usually associate labels and objects after multiple experiences with the words and objects. Often, an adult pairs the words and objects during several social interactions with the child. Acquiring a novel word outside of an experiment has advantages for children. Learning new nouns and verbs is accomplished with fewer trials when an adult provides the new words over several play sessions than when the adult provides the word many times within a single play session (Childers & Tomasello, 2002). Word learning is possible even when the objects and labels are not presented in isolation. Children typically see many objects when hearing a new object name, and the problem of associating the new label with the correct object is thus logically complex (Quine, 1960). When given a variety of objects and novel labels for those objects, adults are able to correctly associate the labels with the correct objects given the cross-situational regularities present across *several* exposures, even when the words and objects are never paired in isolation (Yu & Smith, 2007). Experiments on word learning show that children are capable of similar cross-situational learning (Smith & Yu, 2008). This type of learning is possible given only the statistics present in the input, without the benefit of additional social information. Two recently developed computational models are capable of extracting words from speech and correctly correlating those words with raw visual information; learning was facilitated with the additional of social information such as pointing or the adult's eye gaze (Roy & Pentland, 2002; Yu, Ballard, & Aslin, 2005).

Children are also capable of using social information when learning words. Many explanations of word learning focus on children's interpretation of an adult's behavior while learning the association between an object and a word. In these accounts, the argument is made that children rely on a parent's intention in naming new objects in order to discern the referent of a parent's label (Akhtar & Tomasello, 2000; Baldwin, 1993, Baldwin et al., 1996). Children's task while learning words would thus be to determine the parent's intention, then to associate the object label with the thing to which the parent intended to refer. These approaches often test the ability of children to learn words when the object and word are not present at the same time so as to evaluate children's reliance on adult intention (e.g., Akhtar, Carpenter, & Tomasello, 1996; Baldwin & Moses, 1996).

However, children's performance on theory-of-mind tasks is poor at the onset of word learning, then improves over the first few years (Flavell, 2004). Rather than incorporating children's inferences about adult intentions into an explanation of word learning, other approaches have sought to identify structure in social interactions, especially in the predictable behaviors of caregivers, that children could detect and use for word learning.

Children's sensitivity to such adult behaviors could serve as a foundation for the development of theory of mind (Moore & Povinelli, 2007). In their interactions with infants and young children, adults create a highly structured environment of co-occurrences between speech, movement, and objects that facilitates associative learning (Gogate & Bahrick, 2001; Graf Estes, Evans, Alibali, & Saffran, 2007).

One such study on the social context of word learning observed mothers as they labeled objects and actions for their toddlers (Yu, Smith, Christensen, & Pereira, 2007). During their interaction, mothers and children each wore a head-mounted camera to capture an approximation of what they saw during the interaction. The study found that word-object pairings were disambiguated, in part, through mothers' and children's actions on the objects. Parents labeled objects while holding them, or as children closely examined them. Thus, parents' timing in labeling objects and actions allows children to coordinate visual and auditory perception to learn labels, rather than forcing them to rely on reading parents' intentions while labeling. Other studies of naturalistic parent–child interaction have discovered additional sources of information about word–object pairings. Parents spontaneously move objects in synchrony with their voices as they label (Gogate, Bahrick, & Watson, 2000). Experiments have shown that such intersensory redundancy of label and object improves object–label associative learning in infants (Gogate & Bahrick, 2001). Thus, in the course of social interactions, parents provide multiple sources of information to infants about the associations between words and objects. As parents label objects, they do so in ways that match children's perceptual abilities, which facilitates word learning. The ways in which adults organize the attention of infants and children is an effective mechanism of socially guided word learning.

The size of children's productive vocabularies increases dramatically, beginning around 18 months of age (Bloom, 1973; McMurray, 2007). Around the same time, children begin combining words (around 18–24 months of age), and can produce complex sentences by about 3–5 years of age (Bloom, 2000; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002). The acquisition of syntax, in which children learn to combine words according to the grammatical rules of their native language, is more complex than word learning and has been subject to more debate. The debate has centered on whether syntax is too complex to be learnable, or whether there is instead sufficient reliable information about syntax that children can detect and use to learn the grammatical rules of their language.

Language acquisition is a species-typical human behavior; all neurologically typical humans with a normal social upbringing learn language. Given the universal acquisition and the complexity of syntax, Chomsky (1957) famously argued that that syntax is not learnable and therefore must be innately specified (see also Lenneberg, 1967; Pinker, 1991; Pinker & Jackendoff, 2005). In recent years, an increasing number of linguists, psychologists, and computational modelers have argued against the necessity of innate specifications for syntactic acquisition (e.g., Bates & Goodman, 1999; Edelman, 2008; Elman et al., 1996; Goldberg, 2006; MacWhinney, 2004; Tomasello, 2003). They argue that the impossibility of learning of syntax has been overstated (e.g., Edelman, 2008; MacWhinney, 1999), that not all native speakers acquire syntax with the same degree of competence, reflecting in part differences in input, especially differences in parents' child-directed speech (e.g., Dabrowska & Street, 2006; Hoff, 2003), and that the trajectory of syntactic development is predictable from a combination of the information available to children in the input and a statistical learning mechanism (e.g., Bates & Goodman, 1999; Chater & Manning, 2006; MacWhinney, 2004; Regier & Gahl, 2004).

As discussed above, caregivers' speech to infants is structured in ways that facilitate word learning. Child learners also have many sources of information about syntax from caregivers' child-directed speech. Adults model correct forms of children's ungrammatical utterances, often in the next conversational turn (e.g., Chouinard & Clark, 2003). In infant-directed speech, mothers change their pitch and shorten their sentences, but they also include repetitions and partial repetitions of their utterances (Masters & McRoberts, 2001; Newport, Gleitman, & Gleitman, 1977). Although exact repetitions have a negative effect language development (Newport et al., 1977), partial repetitions have beneficial effects on language learning.

Variation Sets

Child-directed speech often contains partial repetitions, either within parents' own speech (e.g., "Go walky! Want to go walky? Go walky! Walky") or as expansions of children's utterances (e.g., *Child*: "Disappear," *Mother* "It disappeared," *Child:* "Yes, it did disappear," Waterfall, 2006). Such clusters of related utterances are called variation sets. Variation sets are common in parents' speech directed to children between 5 and 30 months of age; approximately 17%–21% of parents' child-directed speech occurs as part of variation sets (Küntay & Slobin, 1996; Waterfall, 2006, 2009a).

Variation sets contain related utterances that are contiguous in time (Waterfall, 2009a). Utterances are related in having the same nonlinguistic content and similar lexical items. Although the sentences contain similar lexical items, they also contain differences within a set. Parents can form variation sets by inserting an item in a previous sentence (e.g., "A hat. A *funny little* hat"), by deleting an item (e.g., "*You* dance. Dance."), by substituting an item in the sentence (e.g., "Let *me* see. Let *Mommy* see"), or by restructuring an item (e.g., "Is this it? This is it," Waterfall, 2009b). Because of their temporal contiguity and their structural similarity, variation sets are structured as to assist children's comparison of sentences. The similarities in topic and structure allow children to discover classes of words or phrases that have equivalent roles in a sentence. For example, in the variation set "Put it in the basket. Throw it in the basket," a child can compare the two sentences and discover that "put" and "throw" can play the same role in a sentence. Harris (1946, 1954) argued that one method for finding lexical classes is to compare similar sentences to determine what words or strings of words can be substituted for each other. This process of comparing similar utterances has been used successfully in computational approaches to extract syntactic patterns and generate new utterances (Solan et al., 2005).

Variation sets promote child language learning, thus it is likely that children can detect the structure available in variation sets. Children are disproportionately likely to learn nouns and verbs that appeared in parents' variation sets (Waterfall, 2009a). In addition, parents' use of variation sets is correlated with children's acquisition of particular syntactic structures, such as use of direct objects and grammatical subjects in commands (Waterfall, 2009b).

Construction-Based Grammar

Parents and children often engage in routines during play, such as peekaboo, book reading, tickling games, etc. (Ninio & Bruner, 1978; Ratner & Bruner, 1978). Bruner (1983) argued that routines provide a context that supports language learning. Because the social content of the interaction is predictable and familiar, infants learn to predict what mothers say and associate her speech with objects and actions. In a longitudinal observational study of picture-book reading in a single parent–child dyad, Ninio and Bruner (1978) observed that the mother used many predictable sentence frames when looking at picture books with her child, such as "Where's the X?" and "Look at the X." These frames could enhance word learning by providing familiar sentence frames with only one unfamiliar element. Recent theories of syntactic development have analyzed language acquisition in terms of these common frames or items (e.g., Cameron-Faulkner, Lieven, & Tomasello, 2003; Tomasello, 2000). These item-based analyses successfully predict children's language production and generalization of new forms.

Frames are very common in parents' child-directed speech; 51% of parents' utterances directed to their 12-month-old children began with one of 52 frames (e.g., *Look at*..., *It's a*....; Cameron-Faulkner et al., 2003). In addition, when children began combining words, they produced utterances that began with these sentence frames. Their rate of use of particular sentence frames was correlated with the rate of parents' use of those frames. Children's production of more complex syntactic structure can also be predicted from the patterns present in parents' speech. For example, the extent to which children produce generalizations of verbs and phrases depends on the extent to which they heard these patterns generalized in parents' speech (e.g., Goldberg, 2006; Tomasello, 2003). Taken together, these findings suggest that children's word learning and syntax are closely tied to the language they hear in social interactions.

Neural Mechanisms of Socially Guided Learning

Scientists who have argued for an innate predisposition for language, especially for syntax acquisition, have sought to identify brain regions responsible for the human language capacity (see Pinker, 1994 for a review). Brain regions implicated in adult language processing were usually identified through studies of brain-injured adults, who had already acquired language and, later, imaging studies of adult language processing. However, developmental studies have identified additional neural mechanisms recruited for language *learning* (e.g., Bates & Roe, 2001, Thomas & Karmiloff-Smith, 2005).

Though the experiments presented in this chapter focus on the behavioral level, there are important parallels between avian and human vocal learning at the neural level. Data on avian socially guided vocal learning have influenced neural analyses by showing that regions of the brain involved in social perception and learning are fundamental to birdsong learning, and thus possibly to human language learning (Freeberg, West, King, Duncan, & Sengelaub, 2002; Hamilton, King, Sengelaub, & West, 1997, 1998; Jarvis, Scharff, Grossman, Ramos, & Nottebohm, 1998; also see Chapter 35). Vocal learning in songbirds is mediated by the anterior frontal pathway (AFP), a neural loop connecting cortical, thalamic, and basal ganglia areas (e.g., Brenowitz & Beecher, 2005). The AFP is homologous to circuitry connecting cortical and basal ganglia areas in primates, including humans (Doupe & Kuhl, 1999; Doupe, Perkel, Reiner, & Stern, 2005).

The basal ganglia may play a particularly important role in socially guided vocal learning, as these structures are involved in reward-based learning of acoustic patterns. When adults receive feedback for performance on a phonological learning task, the caudate nucleus is activated (Tricomi, Delgado, McCandliss, McClelland, & Fiez, 2006). The caudate is involved in the reinforcement of actions that lead to a reward (Tricomi, Delgado, & Fiez, 2004). In addition, the striatum of the basal ganglia, which forms part of the cortical–basal ganglia–cortical loop in primates, is involved in the recognition and control of patterned behavioral sequences and may be important in syntax learning (see review in Osterhout, Kim, & Kuperberg, in press).

The basal ganglia are necessary for attaching value to sound patterns and can influence the organization of auditory cortex (Gao & Suga, 2000). Such structures allow for close associations to be made between forms of vocalizing and emotional state (Cheng, 2003; Cheng & Durand, 2004). Basal ganglia mediation of vocal learning would thus allow an infant bird or baby to have differential reactions to similar auditory stimuli based on emotional value, such as the interaction history between an infant and adult (Martinez-Garcia, Novejarque, & Lanuza, 2007). Thus the basal ganglia are likely to be important substrates of socially guided vocal learning. The robust effects of socially guided learning in multiple aspects of language development indicate that the basal ganglia should be considered in future studies of the neurobiology of language.

Conclusions: From Birds to Words to Language

In summary, our research, grounded in an ecological framework, has shown that social learning is a crucial part of vocal development. Caregivers use the babbling of infants to inform their reactions to the infants (Goldstein & West, 1999). At the same time, vocal learning is sensitive to social feedback (Goldstein et al., 2003; Goldstein & Schwade, 2008). When a new type of sound is produced in the context of contingent social interaction, there are thus multiple processes—both in the parent and in the infant—that support the ability of the infant to retain the new sound form in his or her repertoire. Taken together, these findings suggest that social responses to prelinguistic vocalizations have functional significance for vocal development. Infants learn how to make well-formed syllables, and learn how to order those syllables, as a result of social feedback. Such a system of mutually engaged infants and receivers constitutes an active feedback mechanism that plays a strong role in vocal learning (Figure 34.4). Recent research on cowbird song development has continued the ecological tradition and has moved beyond the dyad to focus on patterns of interaction and learning at the level of the flock (Miller, 2007; Miller, King, & West, 2008). Studying cowbirds and humans in seminaturalistic settings allows them to express species-typical

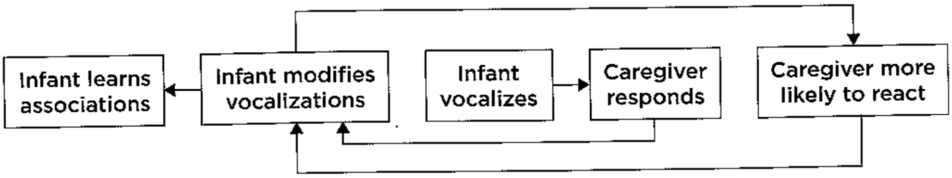

Figure 34.4 A schematic of a socially distributed mechanism of vocal development. When an infant vocalizes, observational studies and our playback studies show that caregivers provide consistent responses to the infant's babbling. Our vocal learning studies show that caregiver responses cause infants to modify their vocalizations; the modifications make the vocalizations more speech-like, which (as shown by our playback studies) elicit stronger responses from caregivers. In addition, infants who receive responses after object-directed vocalizing are more likely to learn from those responses, thus facilitating the learning of word–object associations.

degrees of behavioral freedom and demonstrate multiple forms of socially guided learning.

In both songbirds and human infants, vocal precursors take on functional significance for the development of more advanced forms of vocalizing, but only when the immature sounds are studied in a social context. In both taxa, adults provide rapid, real-time feedback for infants' vocalizations. Infants learn to construct new, more developmentally advanced vocal forms from the reactions of adults to their immature sounds. The processes that create communicative competence from social feedback to vocal precursors are grounded in general learning mechanisms. Learning from contingent responses (social shaping) and learning from recognizing patterns in adults' speech (statistical learning) are both effects of mechanisms that are also used in domains outside of vocal communication. In addition, vocal learning in infants is guided by contingency rather than the perception of intentionality in social partners. We have shown that babbling can be shaped by contingent feedback from an inanimate vehicle.

In our view, learning while babbling lays the foundation for later advances in language. Our findings blur the boundary between prelinguistic and linguistic learning, as phonological and lexical development are necessary components of language. The connections between social learning of phonology, words, and syntax need to be better explored so we can determine how perceptual and statistical mechanisms contribute to the development of language. For example, the role of statistical learning in speech segmentation and grammar learning is well-understood, but there is little research on the statistical learning of patterns in the behavior of social partners. Understanding mechanisms by which infants build expectations of social feedback would help connect socially guided vocal learning with later, more traditionally linguistic advances such as syntax. Caregivers' speech to children as they learn language continues to provide important sources of information about individual lexical items and larger syntactic structures; we need to better characterize the mechanisms by which infants perceive and use that information.

We conclude that socially guided learning is a fundamental characteristic of vocal development. In humans, this includes phonological, lexical, and syntactic development. Taken together, our studies show that the infant's social world is a crucial component of the learning process. When studied as a form of social interaction, babbling is an important stage in the development of communication. The social environment is structured by consistent behavior that captures infants' attention and increases arousal. Infants are active participants in these interactions, creating new patterns of vocalizing that catalyze developmental changes in speech and language.

Acknowledgments

We thank Shimon Edelman and Heidi Waterfall for stimulating discussions on speech and language development. We thank Jennifer Miller and Ambrose Gosling for their helpful comments on the manuscript. The preparation of this manuscript was supported by grants from the Cornell Population Program and the Institute for the Social Sciences at Cornell University.

References

Akhtar, N., Carpenter, M., & Tomasello, M. (1996). The role of discourse novelty in early word learning. *Child Development, 67,* 635–645.

Akhtar, N., & Tomasello, M. (2000). The social nature of words and word learning. In R. M. Golinkoff, K. Hirsh-Pasek, L. Bloom, L. B. Smith, A. L. Woodward, N. Akhtar, M. Tomasello, & G. Hollich (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 114–135). New York: Oxford University Press.

Alberts, J. R. (1978). Huddling by rat pups: Group behavioral mechanisms of temperature regulation and energy conservation. *Journal of Comparative and Physiological Psychology, 92*(2), 231–246.

Anderson, B. J., Vietze, P., & Dokecki, P. R. (1977). Reciprocity in vocal interactions of mothers and infants. *Child Development, 48,* 1676–1681.

Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language, 20*(2), 395–418.

Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic action. *Child Development, 72,* 708–717.

Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word–object relations. *Child Development, 67,* 3135–3153.

Baldwin, D. A., & Moses, L. J. (1996). The ontogeny of social information gathering. *Child Development, 67,* 1915–1939.

Bates, E., & Roe, K. (2001). Language development in children with unilateral brain injury. In C. A. Nelson & M. Luciana (Eds.) *Handbook of developmental cognitive neuroscience* (pp. 281– 308). Cambridge, MA: MIT Press.

Bates, E., & Goodman, J. C. (1999). On the emergence of grammar from the lexicon. In B. MacWhinney (Ed.), *The emergence of language* (pp. 29–80). Hillsdale, NJ: Lawrence Erlbaum Associates.

Beaumont, S. L., & Bloom, K. (1993). Adults' attributions of intentionality to vocalizing infants. *First Language, 13,* 235–247.

Beecher, M. D., & Brenowitz, E. A. (2005). Functional aspects of song learning in songbirds. *Trends in Ecology and Evolution, 20,* 143–149.

Beecher, M. D., & Burt, J. M. (2004). The role of social interaction in bird song learning. *Current Directions in Psychological Science, 13,* 224–228.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167–224). Cambridge, MA: MIT Press.

Best, C. T. (2002). Revealing the mother tongue's nurturing effects on the infant ear. *Infant Behavior and Development, 25*, 134–139.

Bloom, K., D'Odorico, L., & Beaumont, S. (1993). Adult preferences for syllabic vocalizations: Generalizations to parity and native language. *Infant Behavior and Development, 16*, 109–120.

Bloom, K., & Lo, E. (1990). Adult perceptions of vocalizing infants. *Infant Behavior and Development, 13*, 209–219.

Bloom, K., Russell, A., & Wassenberg, K. (1987). Turn taking affects the quality of infant vocalizations. *Journal of Child Language, 14*, 211–227.

Bloom, L. (1973). *One word at a time: The use of single word utterances before syntax.* The Hague: Mouton.

Bloom, P. (2000). *How children learn the meanings of words.* Cambridge, MA: MIT Press.

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

Bornstein, M. H., Tamis-LeMonda, C. S., & Haynes, O. M. (1999). First words in the second year: Continuity, stability, and models of concurrent and predictive correspondence in vocabulary and verbal responsiveness across age and context. *Infant Behavior and Development, 22*, 65–85.

Boysson-Bardies, B. (1993). Ontogeny of language-specific syllabic productions. In B. Boysson-Bardies, S. Schonen, P. Jusczyk, P. MacNeilage & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 353–363). Dordrecht: Kluwer.

Boysson-Bardies, B., Halle, P., Sagart, L., & Durand, C. (1989). A cross-linguistic investigation of vowel formants in babbling. *Journal of Child Language, 16*, 1–17.

Boysson-Bardies, B., & Vihman, M. M. (1991). Adaptation to language: Evidence from babbling and first words in four languages. *Language, 67*, 297–320.

Brenowitz, E. A., & Beecher, M. D. (2005). Song learning in birds: Diversity and plasticity, opportunities and challenges. *Trends in Neuroscience, 28*, 127–132.

Briesch, J., Schwade, J. A., & Goldstein, M. H. (2008). *Responses to prelinguistic object-directed vocalizations facilitate word learning in 11-month-olds.* Poster presented at the International Conference on Infant Studies, Vancouver, BC.

Bruner, J. S. (1983). *Child's talk: Learning to use language.* New York: W.W. Norton.

Cameron-Faulkner, T., Lieven, E., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science, 27*, 843–873.

Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences, 10*, 335–344.

Cheng, M. F. (2003). Vocal self-stimulation: From the ring dove story to emotion-based vocal communication. *Advances in the Study of Behavior, 33*, 309–353.

Cheng, M. F., & Durand, S. E. (2004). Song and the limbic brain: A new function for the bird's own song. *Annals of the New York Academy of Sciences, 1016*, 611–627.

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology, 38*, 967–978.

Chomsky, N. (1957). *Syntactic structures.* The Hague: Mouton.

Chouinard, M. M., & Clark, E. V. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language, 30*, 637–669.

Dabrowska, E., & Street, J. (2006). Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. *Language Sciences, 28*, 604–615.

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

Doupe, A. J., Perkel, D. J., Reiner, A., & Stern, E. A. (2005). Birdbrains could teach basal ganglia research a new song. *Trends in Neurosciences, 28*(7), 353–363.

Edelman, S. (2008). *Computing the mind: How the mind really works.* New York: Oxford University Press.

Elman, J., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press.

Fennell, C. T., Byers-Heinlein, K., & Werker, J. F. (2007). Using speech sounds to guide word learning: The case of bilingual infants. *Child Development, 78*(5), 1510–1525.

Fenson, L., Bates, E., Dale, P., Goodman, J., Reznick, J. S., & Thal, D. (2000). Measuring variability in early child language: Don't shoot the messenger. *Child Development, 71*(2), 323–328.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5, Serial No. 242), 1–173.

Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merril-Palmer Quarterly, 50*(3), 274–290.

Freeberg, T. M., West, M. J., King, A. P., Duncan, S. D., & Sengelaub, D. R. (2002). Cultures, genes, and neurons in the development of song and singing in brown-headed cowbirds (*Molothrus ater*). *Journal of Comparative Physiology, A, 188*, 993–1002.

Friedmann, H. (1929). *The cowbirds: A study in the biology of social parasitism.* Springfield, Ill: C. C. Thomas.

Gao, E., & Suga, N. (2000). Experience-dependent plasticity in the auditory cortex and the inferior colliculus of bats: Role of the corticofugal system. *Proceedings of the National Academy of Science USA, 97*, 8081–8086.

Gerken, L. A. (2006). Decisions, decisions: Infant language learning when multiple generalizations are possible. *Cognition, 98*, B67–B74.

Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development.* Oxford: Oxford University Press.

Gibson, J. J. (1979). *The ecological approach to perception.* Boston: Houghton Mifflin.

Gogate, L. J., & Bahrick, L. E. (2001). Intersensory redundancy and 7-month-old infants' memory for arbitary syllable-object relations. *Infancy, 2*(2), 219–231.

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*(4), 878–894.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language.* Oxford: Oxford University Press.

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, 100*(13), 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., Briesch, J., & Syal, S. (2009). *Learning while babbling: Prelinguistic object-directed vocalizations facilitate learning about objects and words.* Manuscript submitted for publication.

Goldstein, M. H., & West, M. J. (1999). Consistent responses of human mothers to prelinguistic infants: The effect of prelinguistic repertoire size. *Journal of Comparative Psychology, 113*(1), 52–58.

Gómez, R. (2002). Variability and detection of invariant structure. *Psychological Science, 13,* 431–436.

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science, 18*(3), 254–260.

Green, J. A., & Gustafson, G. E. (1983). Individual recognition of human infants on the basis of cries alone. *Developmental Psychobiology, 16*(6), 485–593.

Green, J. A., & Gustafson, G. E. (1997). Perspectives on an ecological approach to social communicative development in infancy. In P. Zukow-Goldring & C. Dent-Read (Eds.), *Evolving explanations of development: Ecological approaches to organism-environment systems* (pp. 515–546). Washington, DC: American Psychological Association.

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development, 30*(6), 509–516.

Gustafson, G. E., & Harris, K. L. (1990). Women's responses to young infants' cries. *Developmental Psychology, 26*(1), 144–152.

Hamilton, K. S., King, A. P., Sengelaub, D. R., & West, M. J. (1997). A brain of her own: A neural correlate of song assessment in a female songbird. *Neurobiology of Learning and Memory, 68*(3), 325–332.

Hamilton, K. S., King, A. P., Sengelaub, D. R., & West, M. J. (1998). Visual and song nuclei correlate with courtship skills in brown-headed cowbirds. *Animal Behaviour, 56,* 973–982.

Harris, Z. S. (1946). From morpheme to utterance. *Language, 22,* 161–183.

Harris, Z. S. (1954). Distributional structure. *Word, 10,* 140–162.

Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD: Brookes.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74,* 1368–1378.

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., & Fogel, A. (2003). Social regulatory effects of infant nondistress vocalization on maternal behavior. *Developmental Psychology, 39*(6), 976–991.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology, 45,* 337–374.

Iverson, J. M., Hall, A. J., Nickel, L., & Wozniak, R. H. (2007). The relationship between reduplicated babble onset and laterality biases in infant rhythmic arm movements. *Brain and Language, 101,* 198–207.

Iverson, J. M., & Thelen, E. (1999). Hand, mouth, and brain: The dynamic emergence of speech and gesture. *Journal of Consciousness Studies, 6,* 19–40.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). Rhythms of dialogue in infancy. *Monographs of the Society for Research in Child Development, 66*(2, Serial No. 265), 1–132.

Jarvis, E. D., Scharff, C., Grossman, M. R., Ramos, J. A., & Nottebohm, F. (1998). For whom the bird sings: Context-dependent gene expression. *Neuron, 21*(4), 775–788.

Johnston, T. D. (1981). Contrasting approaches to a theory of learning. *Behavioral and Brain Sciences, 4,* 125–173.

Jusczyk, P. W. (1992). Developing phonological categories from the speech signal. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 17–64). Timonium, MD: York Press.

Keller, H., & Scholmerich, A. (1987). Infant vocalizations and parental reactions during the first 4 months of life. *Developmental Psychology, 23*(1), 62–67.

Kent, R. D. (1981a). Articulatory-acoustic perspectives on speech development. In R. E. Stark (Ed.), *Language behavior and early childhood* (pp. 105–126). New York: Elsevier.

Kent, R. D. (1981b). Sensorimotor aspects of speech development. In R. N. Aslin, J. R. Alberts & M. R. Petersen (Eds.), *Development of perception* (Vol. 1, pp. 161–189). New York: Academic Press.

King, A. P., West, M. J., & Goldstein, M. H. (2005). Non-vocal shaping of avian song development: Parallels to human speech development. *Ethology, 111,* 101–117.

Kuhl, P. K. (2007a). Cracking the speech code: How infants learn language. *Acoustical Science and Technology, 28*(2), 71–83.

Kuhl, P. K. (2007b). Is speech learning "gated" by the social brain? *Developmental Science, 10,* 110–120.

Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America, 100,* 2425–2438.

Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences USA, 100,* 9096–9101.

Küntay, A., & Slobin, D. (1996). Listening to a Turkish mother: Some puzzles for acquisition. In D. Slobin, J. Gerhardt, A. Kyratzis, & J. Guo (Eds.), *Social interaction, social context, and language: Essays in honor of Susan Ervin-Tripp* (pp. 265–286). Hillsdale, NJ: Lawrence Erlbaum Associates.

Legerstee, M. (1990). Infants use multimodal information to imitate speech sounds. *Infant Behavior and Development, 13,* 343–354.

Lenneberg, E. (1967). *Biological foundations of language.* New York: Wiley.

Levitt, A., & Wang, Q. (1991). Evidence for language-specific rhythmic influences in the reduplicative babbling of French- and English-learning infants. *Language and Speech, 34,* 235–249.

Locke, J. L. (1993). *The child's path to spoken language.* Cambridge, MA: Harvard University Press.

Locke, J. L. (2001). First communion: The emergence of vocal relationships. *Social Development, 10*(3), 294–308.

Locke, J. L., & Pearson, D. M. (1992). Vocal learning and the emergence of phonological capacity: A neurobiological approach. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 91–129). Timonium, MD: York Press.

MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences, 21*, 499–546.

MacNeilage, P. F., & Davis, B. L. (2000). On the origin of the internal structure of word forms. *Science, 288*, 527–531.

MacWhinney, B. (1999). *The emergence of language.* Mahwah, NJ: Lawrence Erlbaum.

MacWhinney, B. (2004). A multiple process solution to the logical problem of language acquisition. *Journal of Child Language, 31*, 883–914.

Marler, P. (1970). Birdsong and speech development: Could there be parallels? *American Scientist, 58*, 669–673.

Marler, P. (1997). Three models of song learning: Evidence from behavior. *Journal of Neurobiology, 33*, 501–516.

Marler, P., & Nelson, D. A. (1993). Action-based learning: A new form of developmental plasticity in bird song. *Netherlands Journal of Zoology, 43*(1–2), 91–103.

Martinez-Garcia, F., Novejarque, A., & Lanuza, E. (2007). Evolution of the amygdala in vertebrates. In J. H. Kaas (Ed.), *Evolution of nervous systems: A comprehensive reference* (Vol. 2, pp. 255–334). Oxford: Academic Press.

Masataka, N. (2003). *The onset of language.* Cambridge: Cambridge University Press.

Masters, R., & McRoberts, G. W. (2001, April). *Verbal repetition in speech to infants.* Poster presented at the Biennial Meeting of the Society for Research in Child Development, Minneapolis, MN.

Mayr, E. (1974). Behavior programs and evolutionary strategies. *American Scientist, 62*, 650–659.

McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science, 317* (3), 631.

Meier, R. P. (1997). Silent mandibular oscillations in vocal babbling. *Phonetica, 54*, 153–171.

Miller, J. L. (2007, November). *Vocal strategies in brown-headed cowbirds* (Molothrus ater) *are sensitive to social ecology.* Paper presented at the International Society for Developmental Psychobiology, San Diego, CA.

Miller, J. L., King, A. P., & West, M. J. (2008). Female social networks influence male vocal development in brown-headed cowbirds (*Molothrus ater*). *Animal Behaviour, 76*, 931–941.

Moore, C., & Povinelli, D. J. (2007). Differences in how 12- and 24-month-olds interpret the gaze of adults. *Infancy, 11*, 215–231.

Narayan, A. J., Goldstein, M. H., & Schwade, J. A. (2008). *What characteristics of social interaction facilitate early learning? Effects of animacy and contingency on vocal learning in prelinguistic infants.* Poster presented at the International Conference on Infant Studies, Vancouver, BC.

Newport, E. L., Gleitman, H., & Gleitman, L. (1977). Mother, I'd rather do it myself: Some effects and noneffects of maternal speech style. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 109–150). Cambridge: Cambridge University Press.

Ninio, A., & Bruner, J. (1978). The achievement and antecedents of labeling. *Journal of Child Language, 5*(1), 1–15.

Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In G. Yeni-Komshian, J. Kavanagh & C. Ferguson (Eds.), *Child phonology.* New York: Academic Press.

Oller, D. K. (1985). Infant vocalizations: Traditional beliefs and current evidence. In S. Harel & N. J. Anastasiow (Eds.), *The at-risk infant: Psycho/socio/medical aspects* (pp. 333–339). Baltimore, MD: Paul H. Brookes.

Oller, D. K. (2000). *The emergence of the speech capacity.* Mahwah, NJ: Lawrence Erlbaum Associates.

Oller, D. K., & Eilers, R. E. (1988). The role of audition in infant babbling. *Child Development, 59*, 441–449.

Oller, D. K., Eilers, R. E., & Basinger, D. (2001). Intuitive identification of infant vocal sounds by parents. *Developmental Science, 4*(1), 49–60.

Oller, D. K., Eilers, R. E., Bull, D. H., & Carney, A. E. (1985). Prespeech vocalizations of a deaf infant: A comparison with normal metaphonological development. *Journal of Speech and Hearing Research, 28*, 47–63.

Oller, D. K., & Lynch, M. P. (1992). Infant vocalizations and innovations in infraphonology: Toward a broader theory of vocal development and disorders. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (Eds.), *Phonological development* (pp. 509–536). Timonium, MD: York Press.

Osterhout, L., Kim, A., & Kuperberg, G. (in press). The neurobiology of sentence comprehension. In M. Spivey, M. Joannisse & K. McRae (Eds.), *The Cambridge handbook of psycholinguistics.* Cambridge: Cambridge University Press.

Papoušek, M. (1989). Determinants of responsiveness to infant vocal expression of emotional state. *Infant Behavior and Development, 12*, 507–524.

Papoušek, M. (1991). Early ontogeny of vocal communication in parent–infant interactions. In H. Papoušek, U. Jergens & M. Papoušek (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches* (pp. 230–261). Cambridge: Cambridge University Press.

Papoušek, M., & Papoušek, H. (1989). Forms and functions of vocal matching in interactions between mothers and their precanonical infants. *First Language, 9*, 137–158.

Papoušek, M., Papoušek, H., & Bornstein, M. H. (1985). The naturalistic vocal environment of young infants: On the significance of homogeneity and variability in parental speech. In T. Field & N. Fox (Eds.), *Social perception in infants* (pp 269–297). Norwood, NJ: Ablex.

Petitto, L. A., Holowka, S., Sergio, L. E., Levy, B., & Ostry, D. J. (2004). Baby hands that move to the rhythm of language: Hearing babies acquiring sign languages babble silently on the hands. *Cognition, 93*, 43–73.

Petrinovich, L. P. (1972). Psychobiological mechanisms in language development. In G. Newton & A. H. Riesen (Eds.), *Advances in psychobiology* (Vol. 1, pp. 259–285). New York: Wiley-Interscience.

Pinker, S. (1991). Rules of language. *Science, 253*, 530–535.

Pinker, S, (1994). *The language instinct.* New York: Harper Collins.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it. *Cognition, 95*, 201–236.

Poulson, C. (1983). Differential reinforcement of other-than-vocalization as a control procedure in the conditioning of infant vocalization rate. *Journal of Experimental Child Psychology, 36*, 471–489.

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

Ratner, N., & Bruner, J. (1978). Games, social exchange and the acquisition of language. *Journal of Child Language, 5*(5), 391–401.

Regier, T., & Gahl, S. (2004). Learning the unlearnable: The role of missing evidence. *Cognition, 93*, 147–155.

Rheingold, H. L., Gewirtz, J. L., & Ross, H. W. (1959). Social conditioning of vocalizations in the infant. *Journal of Comparative and Physiological Psychology, 52*, 68–73.

Rothstein, S. I., Yokel, D. A., & Fleischer, R. C. (1988). The agonistic and sexual functions of vocalizations of male brown-headed cowbirds (*Molothrus ater*). *Animal Behaviour, 36*, 73–86.

Routh, D. K. (1969). Conditioning of vocal response differentiation in infants. *Developmental Psychology, 1*(3), 219–226.

Roy, D. K., & Pentland, A. P. (2002). Learning words from sights and sounds: A computational model. *Cognitive Science, 26*, 113–146.

Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Developmental Psychology, 39*, 484–494.

Schank, J. C., & Alberts, J. (1997). Self-organized huddles of rat pups modeled by simple rules of individual behavior. *Journal of Theoretical Biology, 189*, 11–25.

Schwartz, J.-L., Berthommier, F., & Savariaux, C. (2004). Seeing to hear better: Evidence for early audio-visual interactions in speech identification. *Cognition, 93*, B69–B78.

Seeley, T.D. (1995). *The wisdom of the hive.* Boston: Harvard University Press.

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

Smith, V. A., King, A. P., & West, M. J. (2000). A role of her own: Female cowbirds, *Molothrus ater*, influence the development and outcome of song learning. *Animal Behaviour, 60*, 599–609.

Snowdon, C. T., & Hausberger, M. (Eds.). (1997). *Social influences on vocal development.* Cambridge: Cambridge University Press.

Solan, Z., Horn, D., Ruppin, E., & Edelman, S. (2005). Unsupervised learning of natural languages. *Proceedings of the National Academy of Sciences USA, 102*, 11629–11634.

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388*(6640), 381–382.

Stark, R. E. (1979). Prespeech segmental feature development. In P. Fletcher & M. Garman (Eds.), *Language acquisition.* Cambridge: Cambridge University Press.

Stark, R. E. (1980). Stages of speech development in the first year of life. In G. Yeni-Komshian, J. Kavanagh & C. Ferguson (Eds.), *Child phonology.* New York: Academic Press.

Stoel-Gammon, C. (1992). Prelinguistic vocal development: Measurement and predictions. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 439–456). Timonium, MD: York Press.

Stoel-Gammon, C., & Otomo, K. (1986). Babbling development of hearing-impaired and normally hearing subjects. *Journal of Speech and Hearing Disorders, 51*, 33–41.

Tamis-LeMonda, C. S., & Bornstein, M. H. (2002). Maternal responsiveness and early language acquisition. *Advances in Child Development and Behavior, 29*, 89–127.

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child Development, 72*(3), 748–767.

Thelen, E. (1986). Treadmill-elicited stepping in seven-month-old infants. *Child Development, 57*, 1498–1506.

Thomas, M. S. C., & Karmiloff-Smith, A. (2005). Can developmental disorders reveal the component parts of the human language faculty? *Language Learning and Development, 1*, 65–92.

Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences, 4*, 156–163.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Tricomi, E., Delgado, M. R., McCandliss, B. D., McClelland, J. L., & Fiez, J. A. (2006). Performance feedback drives caudate activation in a phonological learning task. *Journal of Cognitive Neuroscience, 18*(6), 1029–1043.

Tricomi, E. M., Delgado, M. R., & Fiez, J. A. (2004). Modulation of caudate activity by action contingency. *Neuron, 41*(2), 281–292.

Vihman, M. M., Macken, M. A., Miller, R., Simmons, H., & Miller, J. (1985). From babbling to speech: A reassessment of the continuity issue. *Language, 61*, 397–446.

Vihman, M. M., & Miller, R. (1988). Words and babble at the threshold of language acquisition. In M. D. Smith & J. L. Locke (Eds.), *The emergent lexicon: The child's development of a linguistic vocabulary* (pp. 151–183). San Diego: Academic Press.

Waterfall, H. R. (2006). *A little change is a good thing: Feature theory, language acquisition and variation sets.* Unpublished doctoral dissertation, University of Chicago.

Waterfall, H. R. (2009a). *The relation of variation sets to children's noun and verb use.* Manuscript submitted for publication.

Waterfall, H. R. (2009b). *Things change for the better: The relation of variation sets to syntactic development.* Manuscript submitted for publication.

Watson, J. S. (1972). Smiling, cooing, and "the game." *Merrill-Palmer Quarterly, 18*, 323–329.

Weisberg, P. (1963). Social and nonsocial conditioning of infant vocalizations. *Child Development, 34*, 377–388.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development, 1*, 197–234.

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*, 49–63.

West, M. J., & King, A. P. (1985). Learning by performing: An ecological theme for the study of vocal learning. In T. D. Johnston & A. T. Pietriwicz (Eds.), *Issues in the ecological study of learning* (pp. 245–272). Hillsdale, NJ: LEA Press.

West, M. J., & King, A. P. (1988). Female visual displays affect the development of male song in the cowbird. *Nature, 334*(6179), 244–246.

West, M. J., King, A. P., & Duff, M. A. (1990). Communicating about communicating: When innate is not enough. *Developmental Psychobiology, 23*(7), 585–598.

West, M. J., King, A. P., Eastzer, D. H., & Staddon, J. E. R. (1979). A bioassay of isolate cowbird song. *Journal of Comparative and Physiological Psychology, 93*, 124–133.

Whalen, D. H., Levitt, A., & Wang, Q. (1991). Intonational differences between the reduplicative babbling of French- and English-learning infants. *Journal of Child Language, 18*, 501–516.

Wilbrecht, L., & Nottebohm, F. (2003). Vocal learning in birds and humans. *Mental Retardation and Developmental Disabilities Research Reviews, 9*, 135–148.

Wood, R. M., & Gustafson, G. E. (2001). Infant crying and adults' anticipated caregiving responses: Acoustic and contextual influences. *Child Development, 72*(5), 1287–1300.

Yu, C., Ballard, G. A., & Aslin, R. N. (2005). The role of embodied intention in early lexical acquisition. *Cognitive Science, 29*, 961–1005.

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science, 18*, 414–420.

Yu, C., Smith, L. B., Christensen, M., & Pereira, A. (2007). *Two views of the world: Active vision in real-world interaction.* Paper presented at the Proceedings of the 29th annual meeting of the Cognitive Science Society, Nashville, TN.