

Social Mechanisms in Early Language Acquisition: Understanding Integrated Brain Systems Supporting Language

Patricia K. Kuhl

Abstract

This chapter advances a new hypothesis—that in the earliest phases of language acquisition, infants combine a powerful set of domain-general computational skills with their equally extraordinary social skills to “crack the speech code.” Further, it is hypothesized that the social brain—in ways we have yet to understand—“gates” the computational mechanisms underlying learning in the domain of language. Four levels at which the gating mechanism might work are developed. It is argued that the assertion that social factors gate language explains not only how typically developing children acquire language, but also why children with autism exhibit twin deficits in social cognition and language, and why nonhuman animals with impressive computational abilities do not acquire language.

Keywords: language acquisition, infants, domain-general computational skills, social skills, social brain

The earliest phases of language acquisition invoke a special fascination because they allow us to examine humans’ extraordinary abilities to acquire what is arguably our most unique skill—the ability to speak. Humans’ capacity for speech and language provoked classic debates on nature versus nurture by equally strong proponents of nativism (Chomsky, 1959) and learning (Skinner, 1957) as the explanation for children’s remarkable abilities. While we are far beyond those debates and now informed by a great deal of data about infants, their innate predispositions, and their incredible abilities to learn once exposed to natural language (Kuhl, 2009; Saffran, Werker, & Werner, 2006), we are still at ground zero with regard to the mechanisms of language acquisition in the brain and mind of a child.

In this chapter I will advance a new hypothesis—that in the earliest phases of language acquisition, infants combine a powerful set of domain-general computational skills with their equally extraordinary social skills to “crack the speech code.” Further, I hypothesize that the social brain—in ways we have yet to understand—“gates” the computational mechanisms underlying learning in the domain of language. I will develop four levels at which the gating mechanism might work.

The assertion that social factors gate language learning, I argue, explains not only how typically developing children acquire language, but also why children with autism exhibit twin deficits in social cognition and language, and why nonhuman animals with impressive computational abilities

do not acquire language. Moreover, this gating hypothesis may explain why social factors play a far more significant role than previously realized in human learning across domains throughout our lifetimes (Meltzoff, Kuhl, Movellan, & Sejnowski, 2009).

In the next decade, the methods of modern neuroscience will be used to explore how the integration of information across specialized brain systems such as language and social cognition take place, and this will advance our understanding of the potent role social interaction plays in language learning. These approaches, as well as others described here, will lead us towards a transformational view of language acquisition in the human child.

The Learning Problem

Speech learning is a deep puzzle that our theories and machines struggle to solve. How do infants discover the sounds and words used in their particular language(s) when the most sophisticated computers cannot? What is it about the human mind that allows the young child, merely a year old, to understand the words that induce meaning in our collective minds, and to begin to use those words to convey their innermost thoughts and desires? Children's ability to express a thought through words is a breathtaking feat of the human mind.

In explaining a new approach to this problem, I will focus on perception of the elementary units of

language, the consonants and vowels that make up words, and children's early word recognition, to show how our computational and social skills combine to form a very powerful learning mechanism. Interestingly, the new solution does not resemble Skinner's operant conditioning and reinforcement model of learning, nor Chomsky's parameter setting, but rather a view of the process that takes into account new data on children's learning via computational and social means.

Language Exhibits a "Critical Period" for Learning

A stage-setting concept for human language learning is the graph shown in Figure 43.1, redrawn from a study by Johnson and Newport on English grammar in speakers of Korean learning English (1989). The graph as rendered shows a simplified schematic of second language learning as a function of age.

The graph is surprising from the standpoint of human learning more generally. In the domain of language, infants and young children are superior learners when compared to adults, in spite of adults' cognitive superiority. Language is one of the classic examples of a "critical" or "sensitive" period in neurobiology (Bruer, 2008; Johnson & Newport 1989; Knudsen, 2004; Kuhl, 2004; Newport, Bavelier, & Neville, 2001).

Scientists are in agreement that this curve represents data across a wide variety of language-learning studies (Bialystok & Hakuta, 1994;

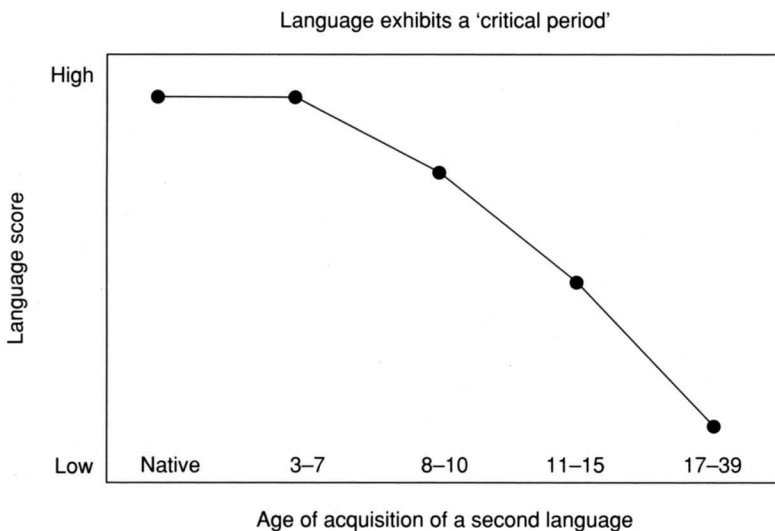


Fig. 43.1 The relationship between age of acquisition of a second language and language skill. From Johnson and Newport (1989); reprinted with permission of Elsevier.

Birdsong & Molis 2001; Flege, Yeni-Komshian, & Liu, 1999; Johnson & Newport, 1989; Kuhl, Conboy, Padden, Nelson, & Pruitt, 2005; Kuhl et al., 2008; Maryberry & Locke, 2003; Neville et al., 1997; Newport & Supalla, 1987; Weber-Fox & Neville 1999; Yeni-Komshian, Flege, & Liu, 2000). The learning function describes our current understanding of second-language learning, though it is composed of many individual sensitive periods for phonological, semantic, and grammatical learning, over time. Given widespread agreement on the fact that we do not learn equally well over the lifespan, theory is currently focused on attempts to explain the phenomenon. What accounts for adults' inability to learn a new language with the facility of an infant?

One of the candidate explanations was Lenneberg's hypothesis that development of the corpus callosum changed the brain in a way that affected learning (Lenneberg, 1967; Newport et al., 2001). More recent hypotheses take a different perspective. Newport has raised the "less is more" hypothesis, which suggests that infants' limited cognitive capacities actually allow superior learning of the simplified language spoken to infants (Newport, 1990). Work in my laboratory has led me to advance the concept of *neural commitment*, the idea that the neural architecture is established early in infancy to detect the phonetic and prosodic patterns of speech (Kuhl, 2004; Zhang et al., 2005; Zhang et al., 2009). This architecture is designed to maximize processing for the language experienced by the infant. Once established, the neural architecture for French or Tagalog, for example, impedes learning of new patterns that do not conform. We will return to the concept of the critical period for language learning, and the role that computational and social skills may play in accounting for the relatively poor performance of adults attempting to learn a second language, later in this chapter.

Focal Example: Phoneme and Word Learning

The world's languages contain approximately 600 consonants and 200 vowels (Ladefoged, 2001). Each language uses a unique set of about 40 distinct elements, *phonemes*, which change the meaning of a word (e.g., from *bat* to *pat*). But phonemes are actually groups of nonidentical sounds, *phonetic units*, which are functionally equivalent in the language. The baby's task is to make some progress in figuring out the composition of the 40-odd phonemic categories before trying to acquire words

which depend on these elementary units. Japanese-learning infants have to group the phonetic units *r* and *l* into a single phonemic category (Japanese *r*), whereas English-learning infants must uphold the distinction to separate *rake* from *lake*. Similarly, Spanish-learning infants must distinguish phonetic units critical to Spanish words (*mano* and *piano*), whereas English-learning infants must combine them into a single category (English *b*). If infants were exposed only to the subset of phonetic units that will eventually be used phonemically to differentiate words in their language, the problem would be trivial. But infants are exposed to many more phonetic variants than will be used phonemically. Simple exposure to a category of sounds does not explain phonetic learning.

Learning to produce the sounds that will characterize infants as speakers of their "mother tongue" is equally challenging, and is not completely mastered until the age of 8 years (Ferguson, Menn, & Stoel-Gammon, 1992). Yet, by 10 months of age, differences can be discerned in the babbling of infants raised in different countries (de Boysson-Bardies, 1993), and in the laboratory, vocal imitation can be elicited by 20 weeks (Kuhl & Meltzoff, 1982). The speaking patterns we adopt early in life last a lifetime (Flege, 1991). My colleagues and I have suggested that this kind of indelible learning stems from a linkage between sensory and motor experience; sensory experience with a specific language establishes auditory patterns stored in memory that are unique to that language and these representations guide infants' successive motor approximations until a match is achieved (Kuhl & Meltzoff, 1996). The ability to imitate vocally may depend on the brain's social understanding mechanisms, those that form a human mirroring system for social understanding (Hari & Kujala, 2009), and we will revisit the brain's social understanding systems later in this chapter.

What enables the kind of learning we see in infants for speech? No machine in the world can derive the phonemic inventory of a language from natural language input (Rabiner & Huang, 1993), though models improve when exposed to "motherese," the linguistically simplified and acoustically exaggerated speech that adults universally use when speaking to infants (de Boer & Kuhl, 2003). The variability in speech input is simply too enormous; Japanese adults produce both English *r*- and *l*-like sounds, exposing Japanese infants to both sounds (Lotto, Sato, & Diehl, 2004; Werker, Pons, Dietrich, Kajikawa, Fais, & Amano, 2007). How do Japanese infants learn

that these two sounds do not distinguish words, and that their differences should be ignored? Similarly, English speakers produce the Spanish *b* and *p*, exposing American infants to both categories of sound (Abramson & Lisker, 1970). How do infants learn that these sounds are not important in distinguishing words in English? An important discovery in the 1970s was that infants initially hear all these phonetic differences (Eimas, 1975; Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Lasky, Syrdal-Lasky, & Klein, 1975; Werker & Lalonde, 1988). What we have to explain is how infants learn which phonetic categories make a difference in their language.

The Timing of Infant Learning

An important discovery in the 1980s identified the timing of the change in infant perception. The transition from an early universal perceptual ability for all languages to language-specific perception occurred very early in development—between 6 and 12 months of age (Werker & Tees, 1984). At this age, infants' perception of nonnative distinctions declines (Best & McRoberts, 2003; Rivera-Gaxiola, Silva-Pereyra, & Kuhl, 2005; Tsao, Liu, & Kuhl, 2006; Werker & Tees, 1984). Work in my laboratory also established a new fact: At the same time that nonnative perception declines, native language speech perception shows a significant increase. We showed that Japanese infants' discrimination of English *r-l* declined between 8 and 10 months of age, while at the same time in development, American infants' discrimination of the same sounds showed an increase (Kuhl et al., 2006) (Figure 43.2).

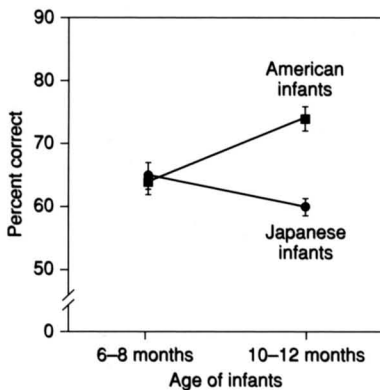


Fig. 43.2 Effects of age on discrimination of the American English /ra-la/ phonetic contrast by American and Japanese infants at 6–8 and 10–12 months of age. Mean percent correct scores are shown with standard errors indicated. From Kuhl et al. (2006); reprinted with permission of John Wiley and Sons.

We argued that the increase observed in native-language phonetic perception represented a critical step in initial language learning (Kuhl et al., 2006; Kuhl et al., 2008). Many studies in our laboratory now show that native-language discrimination between 6 and 7 months predicts the rate of language growth between 11 and 30 months (Conboy, Rivera-Gaxiola, Klarman, Aksoylu, & Kuhl, 2005; Kuhl, Conboy, et al., 2005; Kuhl et al., 2008; Rivera-Gaxiola, Klarman, Garcia-Sierra, & Kuhl, 2005; Tsao, Liu, & Kuhl, 2004). Intriguingly, our data show that while better performance on *native* contrast discrimination predicts rapid growth in later language abilities, better performance on *nonnative* contrasts predicts slower language growth (Kuhl, Conboy, et al., 2005; Kuhl et al., 2008) (Figure 43.3). In other words, phonetic learning does not depend on auditory acuity, but something else. Based on these findings we argued that exposure to language commits the brain's neural circuitry to the properties of native-language speech, and that neural commitment has bi-directional effects—it increases learning for patterns (such as words) that are compatible with the learned phonetic structure, while decreasing perception of nonnative patterns that do not match the learned scheme (Kuhl, 2004).

A Computational Solution to Phonetic and Word Learning

Studies in the decade of the 1990s demonstrated that infants are capable of a surprising new form of learning, referred to as “statistical learning” (Saffran, Aslin, & Newport, 1996). Statistical learning is computational in nature, and reflects implicit rather than explicit learning. It relies on the ability to automatically pick up and learn from the statistical regularities that exist in the stream of sensory information we process, and both phonetic learning and early word learning have been shown to be strongly influenced by this form of learning.

Consider phonetic learning: What causes the developmental change in phonetic perception between the ages of 6 and 12 months? Recent studies show that infants analyze the statistical distributions of sounds that they hear in ambient language, and this affects perception. Although adult listeners hear /r/ and /l/ as either distinct (English speakers) or identical (Japanese), speakers of both languages produce highly variable sounds. Japanese adults produce both English *r*- and *l*-like sounds, so Japanese infants are exposed to both. Similarly, in Swedish there are 16 vowels, whereas English uses 10 and Japanese uses only 5, but speakers of these languages

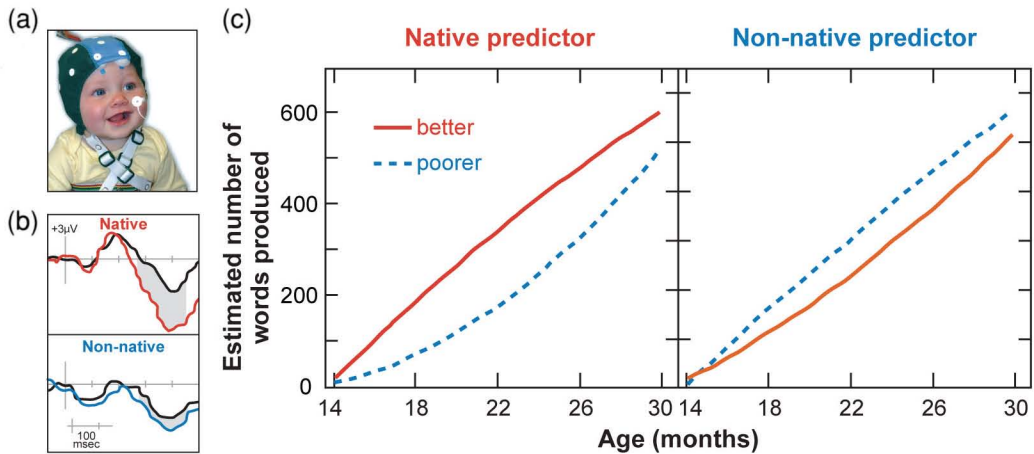


Fig. 43.3 (a) A 7.5-month-old infant wearing an ERP electrocap. (b) Infant ERP waveforms at one sensor location (CZ) for one infant are shown in response to a native (English) and nonnative (Mandarin) phonetic contrast at 7.5 months. The mismatch negativity (MMN) is obtained by subtracting the standard waveform (black) from the deviant waveform (English = red; Mandarin = blue). This infant's response suggests that native-language learning has begun because the MMN negativity in response to the native English contrast is considerably stronger than that to the nonnative contrast. (c) Hierarchical linear growth modeling of vocabulary growth between 14 and 30 months for MMN values of +1SD and -1SD on the native contrast at 7.5 months (c, left) and vocabulary growth for MMN values of +1SD and -1SD on the nonnative contrast at 7.5 months (c, right). Analyses show that both contrasts predict vocabulary growth but that the effects of better discrimination are reversed for the native and nonnative contrasts. From Kuhl et al. (2008); reprinted with permission of Royal Society.

produce a wide range of sounds. The mere presence of a particular sound, therefore, does not account for infant learning, but the distributional patterns of such sounds differ across languages. In each language, distributional frequency is high for phonetic units at the center of phonemic categories, and low at the borders between categories. Distributional patterns of sounds thus provide clues about the phonemic structure of a language. If infants are sensitive to the relative distributional frequencies of phonetic segments in the language that they hear, this could explain native language phonetic learning.

Cross-cultural studies indicate that infants are sensitive to these statistical properties. Infants tested in Sweden and the United States at 6 months of age show a unique response to the “prototypical” vowels in their language—ones that are not only the distributional mean in productions of adults but the ones judged as perceptually preferred by adults; this response is not shown to foreign-language vowel prototypes (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992). Moreover, when tested with very simple stimuli in the laboratory, infants can also learn from distributional patterns in language input after short-term exposure to phonetic stimuli (Maye, Werker, & Gerken, 2002). Six- and 8-month-old

infants were exposed for 2 minutes to 8 sounds that formed a series. They heard all the stimuli on the entire continuum, but experienced different distributional frequencies (Figure 43.4). A “bimodal” group heard more frequent presentations of stimuli at the ends of the continuum; a “unimodal” group heard more frequent presentations of stimuli from the middle of the continuum. After familiarization, infants in the bimodal group discriminated the two sounds, whereas those in the unimodal group did not.

Statistical learning also supports word learning. A previously unsolved problem in both human and machine language learning is how words are detected in the ongoing stream of speech. Unlike written language, spoken language has no reliable markers to indicate word boundaries in typical phrases. How do infants find words? New experiments show that before 8-month-old infants know the meaning of even a single word, they detect likely word candidates through sensitivity to the transitional probabilities between adjacent syllables. In typical words, like in the phrase, “pretty baby,” the transitional probabilities between the two syllables within a word, such as those between “pre” and “ty,” and those between “ba” and “by,” are higher than those between syllables that cross word boundaries, such as “ty” and “ba.”

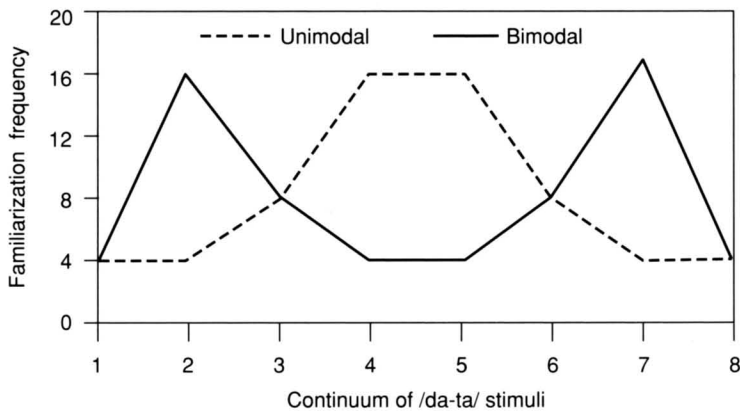


Fig. 43.4 Infants are familiarized for 2 min with a series of /da-ta/ stimuli, with higher frequencies of either stimuli 2 and 7 (bimodal group) or stimuli 4 and 5 (unimodal group). Only infants in the bimodal condition discriminated the /da-ta/ end-point stimuli.

From Maye et al. (2002); reprinted with permission of Elsevier.

Infants are sensitive to these probabilities. When exposed to a 2-min string of nonsense syllables, with no acoustic breaks or other cues to word boundaries, they treat syllables that have high transitional probabilities as “words” (Saffran et al., 1996). Statistical learning is not limited to humans (Hauser, Newport, & Aslin, 2001), nor to speech; it operates for musical and visual patterns in the same way (Fiser & Aslin, 2002; Kirkham, Slemmer, & Johnson, 2002; Saffran, Johnson, Aslin, & Newport, 1999).

Effects of Social Interaction on Computational Learning

Human infants have an intense social interest in people and their behavior, and the newly discovered statistical learning mechanisms, which provide powerful sources of leverage for early learning, appear to be strongly modulated by social interaction.

Demonstrations of statistical learning in the laboratory when infants are exposed to the speech material for only a few minutes, as well as demonstrations that the phenomenon was domain-general, suggested that statistical learning was an automatic process. However, recent studies done in this laboratory suggest that in complex natural language-learning situations social interaction may be necessary for learning. In other words, in natural language learning situations, there are constraints on infants’ computational abilities.

The new experiments showing that social interaction is critical began with a simple question—can infants learn phonetically from first-time natural exposure to a foreign language at 9 months?

We wondered whether statistical learning at 9 months required a long-term history of listening to that language—we reasoned that infant learning at this age might depend on the build-up of statistical distributions over the initial 9 months of life. Alternatively, the transition might occur at 9 months because a learning process initially became available at that age.

We designed a foreign-language intervention to test whether learning the statistics of a new language would occur. Nine-month-old American infants listened to four different native speakers of Mandarin during 12 sessions scheduled over 4–5 weeks time. The foreign language “tutors” read books and played with toys in sessions that were unscripted. A control group was also exposed for 12 sessions but heard only English from native speakers. After infants in the experimental Mandarin exposure group and the English control group completed their sessions, all were tested with a Mandarin phonetic contrast that does not occur in English. Both behavioral and ERP methods were used. The results indicated that infants had a remarkable ability to learn from the live sessions—they performed significantly better on the Mandarin contrast when compared to the control group that heard only English. In fact, they performed equivalently to infants of the same age tested in Taiwan who had listened to the language for 10 months (Kuhl, Tsao, & Liu, 2003).

Learning was durable. Infants returned to the laboratory for their behavioral discrimination tests between 2 and 12 days after the final exposure session, with a median of 6 days, and for their ERP

measurements between 8 and 33 days following the last exposure session, with a median of 15 days. These delays allowed us to examine whether longer periods between exposure and test resulted in poorer discrimination. A median-split approach was used to subdivide infants based on the median delay in days between exposure and test. The results indicated no significant differences between discrimination performance for infants above and below the median delay for either the behavioral or brain tests. No “forgetting” of the Mandarin contrast occurred during the 2 to 33 day delay.

Infants’ remarkable learning led us to test two additional conditions. We were struck by the fact that infants exposed to Mandarin were socially very

engaged in the language sessions. We began to wonder about the role of social interaction in learning, and what would happen if infants were exposed to the same information in the absence of a human being, say, via television or via an audiotape. The results of these tests were surprising. Infants exposed to the same foreign-language material, at the same time in development, and at the same rate, but via standard television or via audio-tape only, showed no learning—their performance equaled that of infants in the control group who had not been exposed to Mandarin at all (Figure 43.5).

Thus, the presence of a human being interacting with the infant during language exposure, while not required for simpler statistical-learning tasks

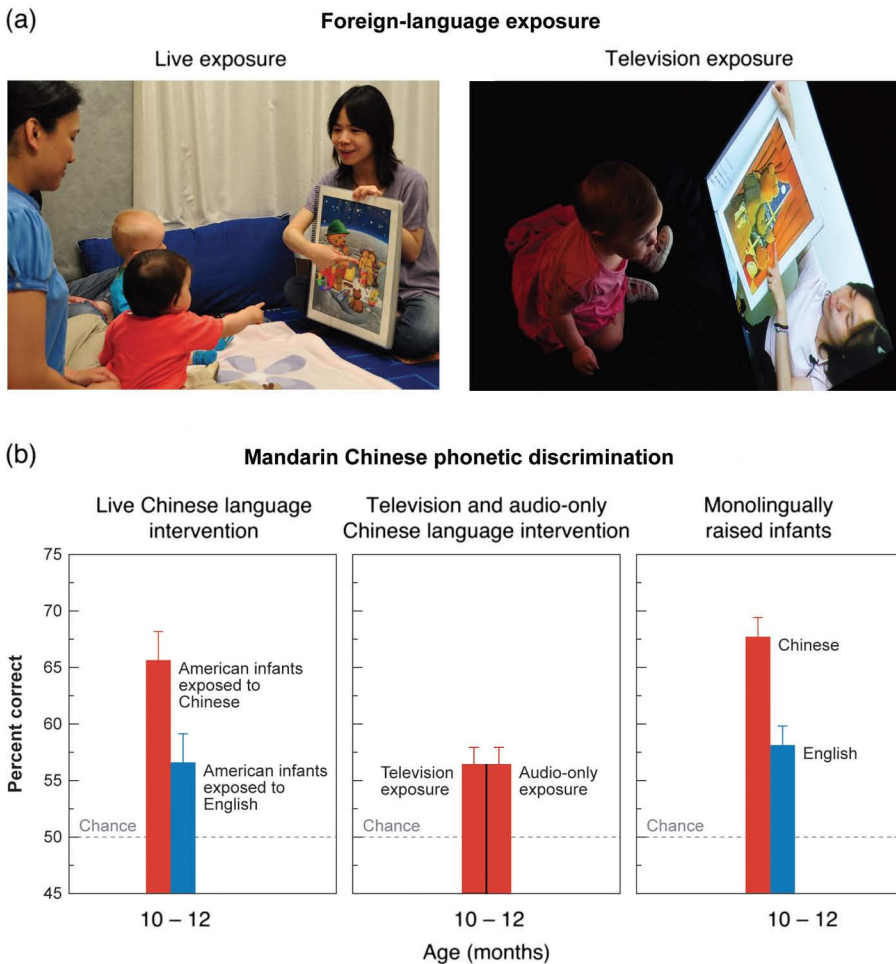


Fig. 43.5 The need for social interaction in language acquisition is shown by foreign-language learning experiments. Nine-month-old infants experienced 12 sessions of Mandarin Chinese through (a) natural interaction with a Chinese speaker (left) or the identical linguistic information delivered via television (right) or audiotape (not shown). (b) Natural interaction resulted in significant learning of Mandarin phonemes when compared with a control group who participated in interaction using English (left). No learning occurred from television or audiotaped presentations (middle). Data for age-matched Chinese and American infants learning their native languages are shown for comparison (right).

From Kuhl et al. (2003); adapted with permission of National Academy of Sciences, USA.

(Maye et al., 2002; Saffran et al., 1996), is critical for learning in complex natural language-learning situations in which infants heard, on average, 33,000 Mandarin syllables, from a total of four different talkers, over a 4–5-week period (Kuhl et al., 2003).

Explaining the Effect of Social Interaction on Language Learning

The 2003 (Kuhl et al., 2003) findings led us to develop a *social gating hypothesis* (Kuhl, 2007) that was in turn modified and expanded by interaction with an interdisciplinary group of scientists in the LIFE Center, an NSF Science of Learning Center. As expanded, the social gating hypothesis examines four concepts that may explain how social settings and interactions fundamentally alter language learning. The concepts revolve around mechanisms that support learning via: (1) attention, (2) information, (3) relationship, and/or (4) brain mechanisms linking perception and action.

The first, the *attention gate*, captures the idea that social contexts and interactions often increase arousal, and increased arousal can lead to increased attention, motivation, and encoding of content. Infant learning in the context of our “live” social exposure could be explained by increased attention and arousal produced by the social interaction. The second is the *intersubjective coordination gate*, and captures the idea that social learning is potent because of the close interaction between people which provides information that enhances learning. An example is joint visual attention—cued by gesture, bodily orientation, and speech—which signals adults’ goals, intentions, desires, and emotions. These coordinated actions provide information that is not available in nonsocial learning settings. Infants’ tracking of adults’ eye movements could help parse speech. An adult’s gaze toward a newly introduced toy, while simultaneously naming that toy, would indicate the correspondence between the object and its name. Thus, the “secret ingredient” provided by social interaction could go beyond arousal and reward and involve intersubjective coordination between people, and the information it provides in social settings.

The third concept is codified by the *sense of relationship gate*, which highlights the idea that a person’s perception of self in relation to others can affect whether they learn. Infants in our social language learning experiment may require a social “tutor” that they perceive to be “like me” (Meltzoff, 2007). In this case learning may occur only from

other humans, or possibly from machines, such as robots, when they exhibit human-like social properties. Finally, a fourth option identifies something more fundamental as the explanation for the potent effects of social learning on language. The *socially adapted brain gate* describes the ways that human beings—who developed evolutionarily to learn from and adapt to others in their social group—are predisposed to place special value on human features (faces, voices), patterns of action (biological movement), and interactions (reciprocal exchanges and interactivity). The emerging field of social cognitive neuroscience is beginning to uncover the brain systems that underlie these preferences and more importantly the human brain systems linking social perception and action (Hari & Kujala, 2009). These brain systems, sometimes called shared representations or “mirroring systems” link what we see and hear others do in response to our own actions. These systems may underpin the parity between self and other that supports seamless interpersonal communication and reciprocity.

The four “gates” are not mutually exclusive. More than one could play a role in explaining the effects of a social other on early language learning. Empirical studies can identify whether a simpler more global mechanism, such as attention, is sufficient to explain these effects, or whether a more complex and fundamentally social system, such as the brain’s evolved mechanisms for social interaction, are necessary. We have made some progress on elucidating these mechanisms in further experiments, as described below.

Attention and Arousal as a Mechanism

Attention and arousal affect learning in a wide variety of domains (Posner, 2004). Could they impact infant learning during exposure to a new language? Infant attention, measured in our studies, was significantly higher in response to the live person than to either inanimate source (Kuhl et al., 2003). Attention has been shown to play a role in the statistical learning studies as well. “High-attender” 10-month-olds learned from bimodal stimulus distributions when “low-attenders” did not (Yoshida, Pons, Cady, & Werker, 2006). Infant arousal, while not measured in our first tests, appeared to be enhanced in some conditions. Infants in the live exposure sessions were visibly aroused before the sessions—they watched the door expectantly, and were excited by the tutor’s arrival, whereas infants in the nonsocial conditions did not. Heightened attention and arousal could produce an overall increase



Fig. 43.6 Infant interactions with a touch-screen TV result in contingent presentation of the Mandarin video clips.

in the quantity or quality of the speech information that infants encode and remember.

In recent tests we examined whether increasing infants' attention would result in learning from the same television presentations we used in the original study (Kuhl et al., 2003). The critical difference was that in these new tests, children's interactions with the TV screen resulted in contingent presentation of the Mandarin video clips (Figure 43.6). We used a touch-screen TV—in other words, the infants had control over the presentation of the language—turning it on with their own touch (Roseberry, Nash, Garcia-Sierra, & Kuhl, in preparation).

Preliminary results, measured using ERPs and the same Mandarin stimuli used in the original study, suggest that some learning does occur. This learning is not as robust as that produced during “live” exposure (Kuhl et al., 2003), but there is some evidence of learning. The group-level behavioral data do not show significant learning of the Mandarin contrast. However, the ERP data show evidence of discrimination at a left hemisphere electrode site (F7), indicating Mandarin sound discrimination after the interactive TV exposure (Roseberry, et al., in preparation). Interestingly, children's ERP responses are related to their vocalizations. Children who vocalize at the TV screen during the exposure sessions show increased negativity to Mandarin sounds at left hemisphere electrode sites. This link to vocalizations is of interest because in other work on infant word learning, infant vocalizations were linked to learning object-word correspondences, and vocalizations were interpreted as indicators of increased attention (Goldstein, Schwade, Briesch, & Syal, 2010). Our interactive TV experiment increases infants' attention—while everything else remains as it was in the original

Kuhl et al. (2003) study. The intermediate level of learning we observed may reflect the effects of attention. Interactive touch-screen TV increases infants' attention to the person on the screen and her toys, and we are currently hypothesizing that this prompts the infant to “test,” through vocalizing at the screen, whether the person on the screen will react to them. Vocalizing may provide a way for infants to attempt to interact with the inanimate device.

Contingency plays a role in human vocalization learning (Bloom, 1975; Bloom & Esposito, 1975; Goldstein, King, & West, 2003), as well as in infant cognition more broadly (Watson, 1979, 2005). Interactivity, the reciprocity that is integral in social exchange, could therefore be a key component of early language learning in the experiments we have designed. Infants have a great deal of experience with people whose vocalizations are contingent on their own: Reciprocity in adult-infant language is common as infants alternate their vocalizations with those of an adult (Bloom, Russell, & Wassenberg, 1987), and the pervasive use of motherese by adults tends to encourage infant reciprocity (Kuhl & Meltzoff, 1982, 1996). Infants may have been “testing” the televised speaker by vocalizing, expecting reciprocity.

Intersubjective Coordination as a Mechanism

We raised a second hypothesis to explain the effectiveness of social interaction—our live learning situation allowed the infants and tutors to interact and this produced added information that could foster learning. During live exposure, tutors focus their visual gaze on pictures in the books or on the toys they talk about, and infants' gaze tends to follow the speaker's gaze (Baldwin, 1995; Brooks, & Meltzoff, 2002) beginning at this point in

development. Referential information is present in both the live and televised conditions, but it is more difficult to pick up via television, and is totally absent during audio-only presentations. Gaze following is a significant predictor of receptive vocabulary (Baldwin, 1995; Brooks & Meltzoff, 2005; Mundy & Gomes, 1998), and may help infants segment foreign speech into word-like units that assist phonetic learning. When 9-month-old infants follow a tutor's line of regard in our foreign-language learning situation, the tutor's specific meaningful social cues, such as eye gaze and pointing to an object of reference, might help infants segment words from ongoing speech, thus facilitating phonetic learning of the sounds contained in those words.

Several key developments coincide with the ability to understand reference. By 9 months, infants begin to engage in triadic “person–person–object games”—they systematically combine attention to objects with looks that promote interest from another human, reflecting a “secondary intersubjectivity” (Trevarthen & Hubley, 1978). Shared perception of communicative intentions, which emerges at around 9 months of age, has been argued to be crucial for the acquisition of language (Akhtar & Tomasello, 1998; Tomasello, 2003a, 2003b). Attending to objects of another person's reference is linked to the infant's growing ability to understand others as intentional agents (Meltzoff, 1995; Tomasello, 2003a). The timing of these social abilities coincides with the beginnings of word comprehension. The suggestion here is that attunement to

the communicative intentions of other humans enhances attention to linguistic units at several levels. Attunement to the meaning of a communicative act would be expected to enhance the uptake of units of language present in that act.

If this hypothesis is correct, then the degree to which infants in our social language-learning situation interact and engage socially with the tutor should correlate with learning. In studies testing this hypothesis, we exposed 9-month-old infants to Spanish. We wanted to extend our findings to a new language, and to examine both Spanish phonetic learning and Spanish lexical learning. To address the social question, we also designed the study to measure specific interactions between the tutor and the infant to examine whether specific kinds of interactive episodes could be related to learning of either phonemes or words.

The results confirmed Spanish learning, both of the phonetic units of the language (Conboy & Kuhl, 2010a) and the lexical units of the language (Conboy & Kuhl, 2010b). In addition, these studies answered a key question—does the degree of infants' social engagement during the Spanish exposure sessions predict the degree of language learning as shown by brain measures? Our results (Figure 43.7) show that they do (Conboy, Brooks, Meltzoff, & Kuhl, submitted; Stevens, Mehus, Mertl, Conboy, & Kuhl, submitted). Infants' eye-gaze data show that when the tutor introduced new toys, infants who shifted attention between the tutor and the toy learned more than infants who simply gazed at the tutor, or the toy.

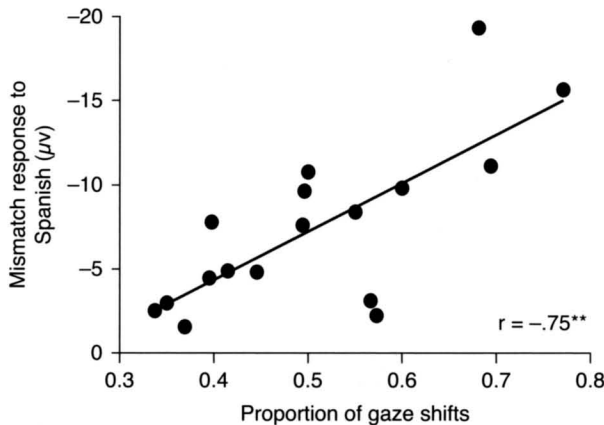


Fig. 43.7 ERP N250–450 effect (mismatch response) to Spanish phonetic contrast /d-t/, by proportion of gaze shifts during Spanish sessions.
From Conboy et al. (submitted).

A parallel set of ERP measures show that lexical learning also occurred from these Spanish exposure sessions. The ERP results show infants learned Spanish words that were presented during sessions, as opposed to those never presented, indicating that infants learned these new word patterns. The degree of infants' social engagement during sessions predicted word learning and the results mirror the phonetic findings—the more socially engaged the infant, the greater the word learning an infant showed (Conboy & Kuhl, 2010b).

There is evidence to suggest that cognitive (inhibitory) control is improved in bilingual speakers (Bialystok, 1999, 2001; Bialystok & Hakuta, 1994), and tests done both before and after the Spanish exposure sessions show an association between the degree of phonetic and/or word learning and infants' scores on tests of inhibitory control (Conboy, Sommerville, & Kuhl, in preparation). Infants who show greater learning from Spanish exposure show higher skill on tests of inhibitory control after, but not before, Spanish exposure. These results suggest that even short-term second-language exposure *may* play some role in enhancing inhibitory control skills (Conboy et al., in preparation), a finding consistent with those showing higher inhibitory control skills in bilingual adults (Bialystok, 1999, 2001; Bialystok & Hakuta, 1994).

Finally, the findings on Spanish exposure were extended to speech production and show that infants' exposure to Spanish affects their speech production as well their speech perception skills (Ward, Sundara, Conboy, & Kuhl, 2009). Infants exposed to 12 sessions of Spanish showed changes in their patterns of vocalization. After Spanish exposure, infants' vocalizations change to reflect the prosodic patterns of Spanish that are not characteristic of English when listening to a Spanish speaker but not when listening to an English speaker. This does not occur before exposure to Spanish in the laboratory or in infants who did not experience the 12-session Spanish exposure. Infants' learning from perceptual experience extends to the motor system.

The Spanish exposure experiments provide strong support for the idea that social interactivity and engagement plays an important role in initial language learning. The fact that social engagement promotes word learning has been shown in previous studies (Baldwin, 1995; Brooks & Meltzoff, 2005), but this effect has not previously been shown for phonetic learning. The findings suggest a more fundamental role for social interaction in language learning, one more akin to that seen in social

learning in birds, in which conspecific song learning occurs only when engaged socially with tutor birds (Woolley & Doupe, 2008).

A Sense of Relationship as a Mechanism

The finding that social interaction is critical to learning raise a more fundamental question: What defines a “social agent” for infants? Must a social agent involve a human being (with sight, smell, and all other indicators of humanness), or would an inanimate entity, imbued with certain interactive features, induce infant perception of a social being? And if so, could infants learn language from such a socially augmented entity? Social interaction might be effective *because* it involves other humans, or because features inherent in social settings, such as interactivity and contingency, are critical for learning.

In another set of studies, we took a different approach to address the question. The interactive television experiment made only one change from the original television condition tested in Kuhl et al. (2003). Infants had control over the TV presentation, using a touchscreen to provide a contingent presentation of Mandarin. Using a different approach, we tested a more “human-like” machine, one that delivered contingencies automatically, and had certain key characteristics of a human actor.

We examined 18–24-month-old children's learning from a foreign language (Finnish) via a “social robot.” The robot interacts with the children in a daycare setting by (1) orienting its head to them when they hand the robot a toy, (2) naming the toy in Finnish in a sentence context, and (3) taking the toy from the child using its pincer (Tanaka, Cicourel, & Movellan, 2007). Learning was assessed using a pre- and post-test examining children's knowledge of 10 Finnish words, with 10 English words used as controls. At the group level, English word learning did not change over the 2-week period of the intervention, while Finnish word learning increased significantly (Figure 43.8). Of most interest were the individual differences. Excellent learning was shown in some of the children, but not in all of them. Some children learned all of the 10 Finnish words and repeated these new words throughout the day in the daycare setting, as reported by the teachers. Supporting our hypothesis, learning was greatest in children who had the most sustained interaction with the robot during the learning episodes. This sustained interaction appears to be more important than overall time spent in interacting with the robot. During sustained periods of interaction, young children engage at different levels



Fig. 43.8 A social robot can operate autonomously with children in a preschool setting. In this photo, toddlers play a game with the robot. One long-term goal is to engineer systems that test whether young children can learn a foreign language through interactions with a talking robot.

with the robot, and analyses of these data are continuing.

Socially Adapted Brain as a Mechanism

While attention, intersubjective coordination, and a sense of relationship may help explain our social language learning effects, it is also possible that social contexts are connected to language through more fundamental mechanisms. Social interaction may activate brain mechanisms of social understanding that link perception and action (Hari & Kujala, 2009). Neuroscience research focused on shared neural systems for perception and action have a long tradition in speech, and interest in “mirror systems” for social cognition (Kuhl, 2007; Kuhl & Meltzoff, 1996; Meltzoff & Decety, 2003; Pulvermuller, 2005; Rizzolatti, 2005; Rizzolatti & Craighero, 2004) have re-invigorated this tradition. Might the brain systems that link perception and production for speech be engaged when infants experience social interaction during language learning?

The theoretical linkage between perception and action in speech began with the motor theory (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967) and direct realism (Fowler, 1986), both of which posited that the mechanisms of speech production are involved in speech perception. The discovery of “mirror neurons” in monkeys that react both to the sight of others’ actions and the same actions produced by themselves (Gallese 2003; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996; Rizzolatti, Fogassi, & Gallese, 2002) rekindled interest in a potential mirror system for speech, as has work on the origins of infant

imitation (Meltzoff & Moore, 1997). The perception-action link for speech has been viewed as potentially innate (Liberman & Mattingly, 1985) and also viewed as forged early in development through experience (Kuhl & Meltzoff, 1982, 1996).

Neuroscience studies using speech and imaging techniques have the capacity to examine whether the brain systems involved in speech production are activated when infants listen to speech. Two new infant studies take a first step towards an answer to this developmental issue. Imada et al. (2006) used magnetoencephalography (MEG) to study newborns, 6-month-old infants, and 12-month-old infants while they listened to nonspeech, harmonics and syllables (Figure 43.9). Dehaene-Lambertz, Hertz-Pannier, Dubois, Meriaux, and Roche (2006) used fMRI to scan 3-month-old infants while they listened to sentences. Both studies show activation in brain areas responsible for speech production (the inferior frontal, Broca’s area) in response to auditorially presented speech. Imada et al. reported synchronized activation in response to speech in auditory and motor areas at 6 and 12 months, and Dehaene et al. reported activation in motor speech areas in response to sentences in 3-month olds. Is activation of Broca’s area to the pure perception of speech present at birth? Newborns tested by Imada et al. showed no activation in motor speech areas for any signals, whereas auditory areas responded robustly to all signals, suggesting the possibility that perception-action linkages for speech develop by 3 months of age as infants produce vowel-like sounds. Further work must be done to answer the

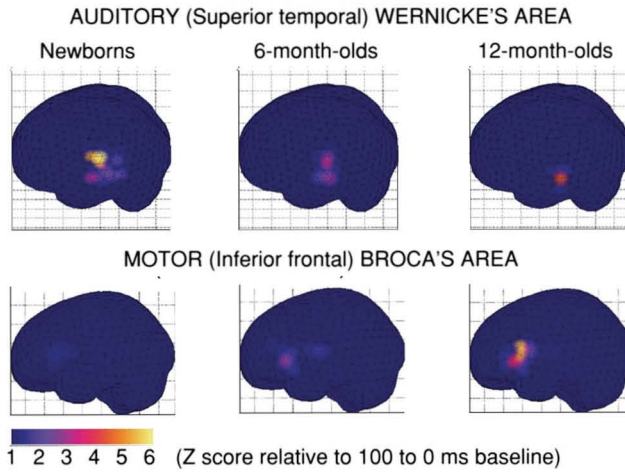


Fig. 43.9 (top) Neuromagnetic signals were recorded in newborns, 6-month-old (shown) and 12-month-old infants in the MEG machine while listening to speech and nonspeech auditory signals. (bottom) Brain activation in response to speech recorded in auditory (top row) and motor (bottom row) brain regions showed no activation in the motor speech areas in the newborn in response to auditory speech, but increasing activity that was temporally synchronized between the auditory and motor brain regions was observed in 6- and 12-month-old infants when listening to speech.

From Imada et al. (2006); reprinted with permission of Wolters Kluwer Health.

question—whether binding of perception and action requires experience with the language, or is activated by speech with or without experience. This is one of the exciting questions that can now be addressed. Using the tools of modern neuroscience, we can now ask how the brain systems responsible for speech perception and speech production forge links in early development, and if these same brain areas are involved when language is presented socially, but not when language is presented through a disembodied source such as a television set.

Implications for Children with Autism Spectrum Disorder (ASD)

Scientific discoveries on the progression toward language by typically developing children are providing new insights into the language deficit shown by children with autism spectrum disorder (ASD). Neural measures of language processing

in children with autism, involving both phonemes and words, when coupled with measures of social interest in speech in children with ASD, are revealing a tight coupling between social interaction skills and language acquisition. These measures hold promise as potential diagnostic markers of risk for autism in very young children, and therefore there is a great deal of excitement surrounding the application of these basic measures of speech processing as potential biomarkers in very young children with autism.

In typically developing children, behavioral and/or brain responses to a change in speech predicts the growth of language to the age of 30 months (Kuhl et al., 2008; Rivera-Gaxiola, Klarman, et al., 2005). We are therefore interested in the degree to which ERP measures of phonetic processing are sensitive to the degree of severity of autism, and also the degree to which the brain's responses to syllables can

be predicted by other factors, such as a social interest in speech.

We examined phonetic perception in preschool-age children with ASD (Kuhl, Coffey-Corina, Padden, & Dawson, 2005), using ERPs to measure responses to a simple change in two speech syllables. A measure of the children's social interest in speech was also taken. The social measure was a listening choice that allowed children with autism to select between listening to "motherese" or an acoustically matched nonspeech signal. Motherese is a social signal that typical infants prefer when given a choice between infant-directed and adult-directed speech (Fernald & Kuhl, 1987; Grieser & Kuhl, 1988), and the phonetic units in motherese are acoustically stretched (Kuhl et al., 1997), a feature of maternal speech that can be linked to infants' performance on tests of phonetic perception (Liu, Kuhl, & Tsao, 2003).

In the tests on children with autism and typical controls, the nonspeech signal was created by replacing the formant frequencies of speech with pure tones at the same frequencies. The resulting signal was a computer warble that followed exactly the frequencies and amplitudes of the 5-second speech samples over time. Slight head turns to one direction versus the other allowed the children to choose their preferred signal on each trial. Our goal was to compare performance at the group level between typically developing children and children with ASD, as well as to examine the relationship between brain measures of speech perception and measures of social processing of speech in children with ASD.

The ERP measures of phonetic perception showed that, as a group, children with ASD did not exhibit the mismatched negativity (MMN) that marks discrimination of the simple change in syllables. However, when children with ASD were subgrouped on the basis of their preference for motherese, very different results were obtained. Children with ASD who preferred motherese produced MMN responses that resembled those of typically developing children, whereas those who preferred the non-speech analog did not show an MMN response to the change in a speech syllable.

The results of auditory preference testing showed that while typically developing children listen to both motherese and non-speech analog signals, children with autism strongly preferred the non-speech analog signals. Moreover, the degree to which they did so was significantly correlated with both the severity of autism symptoms (see Figure 43.10) and MMN responses to speech syllables.

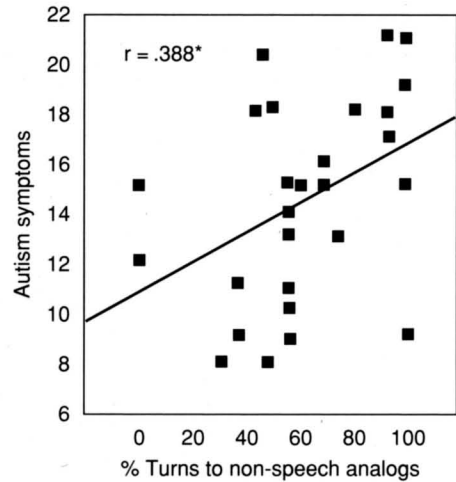


Fig. 43.10 Children with ASD: scatter plots relating autism symptoms (ADOS social-communication total) and percent head-turns to non-speech analogs ($r_s = .388$, $p = .037$, $n = 29$). From Kuhl, Coffey-Corina, et al. (2005); reprinted with permission of John Wiley and Sons.

These results underscore the importance of a social interest in speech early in development, especially an interest in motherese. The acoustic stretching in motherese makes phonetic units more distinct from one another (Burnham, Kitamura, & Vollmer-Conna, 2002; Englund, 2005; Kuhl et al., 1997; Liu et al, 2003; Liu, Tsao, & Kuhl, 2007). Mothers who use the exaggerated phonetic patterns to a greater extent when talking to their typically developing 2-month-old infants have infants who, months later in the laboratory, show significantly better performance in phonetic discrimination tasks (Liu et al., 2003). New data show that the potential benefits of early motherese extend to the age of 5 years (Liu, Tsao, & Kuhl, submitted). In the absence of a listening preference for motherese, children with autism would miss the benefit these exaggerated phonetic cues provide.

Infant-directed speech also produces unique brain responses in typically developing infants. Brain measures of typical infants' response to infant-directed speech, used by Pena et al. (2003) in the first study using near infrared spectroscopy (NIRS), showed more activation in left temporal areas when infants were presented with infant-directed speech as opposed to backward speech or silence. Bortfeld, Wruck, and Boas (2007) obtained analogous results using NIRS in a sample of 6–9-month-old infants presented with infant-directed speech and visual stimulation. It would be of interest to examine brain activation while children with autism listen

to motherese as opposed to acoustically matched nonspeech signals. In children with ASD, brain activation to carefully controlled speech versus nonspeech signals may provide clues to these children's aversion to highly intonated speech signals typical of motherese.

Neurobiological Foundations of Social Learning

Humans are not the only species in which communicative learning is affected by social interaction. Communicative learning in songbirds provides an example. Young zebra finches need visual interaction with a tutor bird to learn song in the laboratory (Eales, 1989). A zebra finch will override its innate preference for conspecific song if a Bengalese finch foster father feeds it, even when adult zebra finch males can be heard nearby (Immelmann, 1969). White-crowned sparrows, which reject the audiotaped songs of alien species, learn the same alien songs when a live tutor sings them (Baptista & Petrinovich, 1986). In barn owls (Brainard & Knudsen, 1998) and white-crowned sparrows (Baptista & Petrinovich, 1986), a richer social environment extends the duration of the sensitive period for learning. Social contexts also advance song production in birds; male cowbirds respond to the social gestures and displays of females, which affect the rate, quality, and retention of song elements in their repertoires (West & King, 1988), and white-crowned sparrow tutors provide acoustic feedback that affects the repertoires of young birds (Nelson & Marler, 1994).

In birds, social interaction can take various forms. Blindfolded zebra finches that cannot see the tutor, but can interact through pecking and grooming, learn their songs. And young birds operantly conditioned to present conspecific song to themselves by pressing a key learn the songs they hear (Adret, 1993; Tchernichovski, Mitra, Lints, & Nottebohm, 2001). In other words, in birds, interactivity and contingency play critical roles even in the absence of another bird.

Neural Underpinnings of Social Influences on Language Learning

Language evolved to address a need for social communication and evolution may have forged a link between language and the social brain in humans (Adolphs, 2003; Dunbar, 1998; Pulvermuller, 2005). Work on mirror neurons in nonhuman primates indicates a neural link between the self and other; seeing an action and producing it oneself are

neurally equivalent in adult monkeys, and shared neural systems for action and perception assist imitation and social understanding (Meltzoff & Decety, 2003; Rizzolatti, 2005). Research on the development of the neural networks that constitute the "social brain" is beginning to appear (Hari & Kujala, 2009).

The fact that MEG has now been demonstrated to be feasible for developmental studies of speech perception in infants during the first year of life (Imada et al., 2006) opens up new possibilities for examining shared brain systems for action and perception involving speech in infancy. Examining brain activation using MEG with infants during social versus nonsocial language experience will allow us to test whether shared brain networks are activated differentially in the two conditions.

If social factors "gate" computational learning, infants would be protected from meaningless calculations—learning would be restricted to signals that derive from humans (or close facsimiles) rather than other sources (Doupe & Kuhl, 1999; Evans & Marler, 1995; Marler, 1991). Constraints of this kind appear to exist for infant imitation: When infants hear sounds that resemble vowels, but ones that could not be produced by human vocal tracts, they fail to imitate (Kuhl, Williams, & Meltzoff, 1991), and in action imitation experiments, infants infer and reproduce intentions displayed by humans but not by machines (Meltzoff, 1995). Our current experiments with social robots will explore the extent to which children will treat a social robot as a social agent. The data establish an interesting set of boundary conditions for language: Exposure to a new language in a live social interaction situation induces remarkable learning in 9-month-old infants, but *no* learning when the exact same language material is presented to infants by a disembodied source. We now are exploring various mechanisms that may explain this dramatic difference: attention, interactive coordination that enhances information, a sense of relationship, and evolution of a socially adapted brain.

Many questions remain about the impact of social interaction on natural speech and language learning. The idea that social interaction is integral to language learning is admittedly not entirely new because it has been linked by many to word learning; however, and importantly, previous data and theorizing have not tied early *phonetic* learning to social factors, and doing so suggests a more fundamental connection between the motivation to learn socially and the mechanisms that enable language learning. Moreover, linking early speech acquisition

to social cognition places human communication solidly within the neurobiology of communication learning in nonhuman animals. Neuroscience studies over the next decade will help advance our understanding of how brain systems underlying language and social cognition interact during the earliest phases of language learning.

Acknowledgments

The author and research were supported by a grant from the National Science Foundation's Science of Learning Program to the LIFE Center (SBE-0354453), and by grants from the National Institutes of Health (HD37954, HD55782, HD02274, DC04661).

References

Abramson, A. S. & Lisker, L. (1970). Discriminability along the voicing continuum: Cross-language tests. *Proceedings of the Sixth International Conference of Phonetic Sciences in Prague* (pp. 569–573). Prague: Academia.

Adolphs, R. (2003). Cognitive neurosciences of human social behavior. *Nature Reviews Neuroscience*, 4, 165–178.

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

Akhtar, N. & Tomasello, M. (1998). Intersubjectivity in early language learning and use. In S. Bråten (Ed.), *Intersubjective communication and emotion in early ontogeny* (pp. 316–335). Cambridge: Cambridge University Press.

Baldwin, D. A. (1995). Understanding the link between joint attention and language. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 131–158). Hillsdale, NJ: Lawrence Erlbaum Associates.

Baptista, L. F. & Petrinovich, L. (1986). Song development in the white-crowned sparrow: Social factors and sex differences. *Animal Behavior*, 34, 1359–1371.

Best, C. & McRoberts, G. W. (2003). Infant perception of nonnative consonant contrasts that adults assimilate in different ways. *Language and Speech*, 46, 183–216.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development*, 70, 636–644.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. New York, NY: Cambridge University Press.

Bialystok, E. & Hakuta, K. (1994). *In other words: The science and psychology of second-language acquisition*. New York, NY: Basic Books.

Birdsong, D. & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisitions. *Journal of Memory and Language*, 44, 235–249.

Bloom, K. (1975). Social elicitation of infant vocal behavior. *Journal of Experimental Child Psychology*, 20, 51–58.

Bloom, K. & Esposito, A. (1975). Social conditioning and its proper control procedures. *Journal of Experimental Child Psychology*, 19, 209–222.

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

de Boer, B. & Kuhl, P. K. (2003). Investigating the role of infant-directed speech with a computer model. *Acoustic Research Letters Online (ARLO)*, 4, 129–134.

Bortfeld, H., Wruck, E., & Boas, D. A. (2007). Assessing infants' cortical response to speech using near-infrared spectroscopy. *NeuroImage*, 34, 407–415.

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

Brainard, M. S. & Knudsen, E. I. (1998). Sensitive periods for visual calibration of the auditory space map in the barn owl optic tectum. *The Journal of Neuroscience*, 18, 3929–3942.

Brooks, R. & Meltzoff, A. N. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology*, 38, 958–966.

Brooks, R. & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, 8, 535–543.

Bruer, J. T. (2008). Critical periods in second language learning: Distinguishing phenomena from explanation. In M. Mody and E. Silliman (Eds.), *Brain, behavior and learning in language and reading disorders* (pp.72–96). New York: The Guilford Press.

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

Chomsky, N. (1959). A review of B. F. Skinner's "Verbal Behavior". *Language*, 35, 26–58.

Conboy, B. T. & Kuhl, P. K. (2010a). Impact of second-language experience in infancy: Brain measures of first- and second-language speech perception. *Developmental Science*, 1–7. Advance online publication. doi:10.1111/j.1467-7687.2010.00973x.

Conboy, B. T. & Kuhl, P. K. (2010b). Brain responses to words in 11-month-old infants after exposure to a second language. Paper presented at the American Speech, Language, and Hearing Association Convention, Philadelphia, PA, 18–20 November.

Conboy, B. T., Brooks, R., Meltzoff, A. N., & Kuhl, P. K. (submitted). Infants' social behaviors during exposure to a second language predict phonetic learning in that language.

Conboy, B., Rivera-Gaxiola, M., Klarman, L., Aksoylu, E., & Kuhl, P. K. (2005). Associations between native and nonnative speech sound discrimination and language development at the end of the first year. In A. Brugos, M. R. Clark-Cotton, & S. Ha (Eds.), *Supplement to the Proceedings of the 29th Boston University Conference on Language Development*; <http://www.bu.edu/linguistics/APPLIED/BUCLD/supp29.html>.

Conboy, B., Sommerville, J., & Kuhl, P. K. (in preparation). Infants' phonetic second-language learning is linked to executive control.

Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Meriaux, S., & Roche, A. (2006). Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proceedings of the National Academy of Sciences* 103, 14240–14245.

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

Dunbar, R. J. M. (1998). The social brain hypothesis. *Evolutionary Anthropology*, 6, 178–190.

- Eales, L. (1989). The influences of visual and vocal interaction on song learning in zebra finches. *Animal Behavior*, 37, 507–508.
- Eimas, P. D. (1975). Auditory and phonetic coding of the cues for speech: Discrimination of the /r-/l/ distinction by young infants. *Perception and Psychophysics*, 18, 341–347.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171, 303–306.
- Englund, K. T. (2005). Voice onset time in infant directed speech over the first six months. *First Language*, 25, 219–234.
- Evans, C. S. & Marler, P. (1995). Language and animal communication: Parallels and contrasts. In H. L. Roitblat & J. -A. Meyer (Eds.), *Comparative approaches to cognitive science: Complex adaptive systems* (pp. 341–382). Cambridge, MA: MIT Press.
- Ferguson, C. A., L. Menn, & C. Stoel-Gammon. (Eds.) (1992). *Phonological development: Models, research, implications*. Timonium, MD: York Press.
- Fernald, C. A. & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10, 279–293.
- Fiser, J. & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Science, USA*, 99, 15822–15826.
- Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America*, 89, 395–411.
- Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, 41, 78–104.
- Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, 14, 3–28.
- Gallese, V. (2003). The manifold nature of interpersonal relations: The quest for a common mechanism. *Philosophical Transactions of the Royal Society B*, 358, 517–528.
- 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, 830–835.
- Goldstein, M. H., Schwade, J., Briesch, J., & Syal, S. (2010). Learning while babbling: Prelinguistic object-directed vocalizations indicate a readiness to learn. *Infancy*, 1–30. Retrieved February 1, 2010 from <http://www3.interscience.wiley.com/cgi-bin/fulltext/123237333/HTMLSTART>.
- Greiser, D. L. & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology*, 24, 14–20.
- Hari, R. & Kujala, M. (2009). Brain basis of human social interaction: From concepts to brain imaging. *Physiological Reviews*, 89, 453–479.
- Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a nonhuman primate: Statistical learning in cotton-top tamarins. *Cognition*, 78, B53–B64.
- Imada, T., Zhang, Y., Cheour, M., Taulu, S., Ahonen, A., & Kuhl, P. K. (2006). Infant speech perception activates Broca's area: A developmental magnetoencephalography study. *NeuroReport*, 17, 957–962.
- Immelmann, K. (1969). Song development in the zebra finch and other estrildid finches. In R. Hinde (Ed.), *Bird vocalizations* (pp. 61–74). London: Cambridge University Press.
- Johnson, J. & Newport, E. (1989) Critical period effects in second language learning: The influence of maturation state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60–99.
- Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, 83, B35–B42.
- Knudsen, E. I. (2004). Sensitive periods in the development of the brain and behavior. *Journal of Cognitive Neuroscience*, 16, 1412–1225.
- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5, 831–843.
- Kuhl, P. K. (2007). Is speech learning “gated” by the social brain? *Developmental Science*, 10, 110–120.
- Kuhl, P. K. (2009). Early language acquisition: Neural substrates and theoretical models. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences, 4th Edition* (pp. 837–854). Cambridge, MA: MIT Press.
- Kuhl, P. K. (2009). Early language acquisition: Phonetic and word learning, neural substrates, and theoretical model. In B. C. Moore, L. K. Tyler, & W. D. Marslen-Wilson (Eds.), *The perception of speech* (pp. 103–131). New York: Oxford University Press Inc.
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277, 684–686.
- Kuhl, P. K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological evidence. *Developmental Science*, 8, 1–12.
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, P., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B*, 363, 979–1000.
- Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the “critical period.” *Language Learning and Development*, 1, 237–264.
- Kuhl, P. K. & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, 218, 1138–1141.
- 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., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show facilitation for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9, 13–21.
- Kuhl, P. K., Tsao, F. -M., & Liu, H. -M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, 100, 9096–9101.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255, 606–608.
- Kuhl, P. K., Williams, K. A., & Meltzoff, A. N. (1991). Cross-modal speech perception in adults and infants using nonspeech auditory stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 829–840.
- Ladefoged, P. (2001). *Vowels and consonants: An introduction to the sounds of language*. Oxford: Blackwell Publishers.

- Lasky, R. E., Syrdal-Lasky, A., & Klein, R. E. (1975). VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology*, 20, 215–225.
- Lenneberg, E. (1967). *Biological foundations of language*. New York: John Wiley & Sons.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431–461.
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1–36.
- Liu, H. -M., Kuhl, P. K., & Tsao, F. -M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, 6, F1–F10.
- Liu, H. -M., Tsao, F. -M., & Kuhl, P. K. (2007). Acoustic analysis of lexical tone in Mandarin infant-directed speech. *Developmental Psychology*, 43, 912–917.
- Liu, H. -M., Tsao, F. -M., & Kuhl, P. K. (submitted). Lexical diversity and phonetic clarity of maternal speech in infancy predicts 5-year-olds' language skills.
- Lotto, A. J., Sato, M., & Diehl, R. (2004). Mapping the task for the second language learner: The case of Japanese acquisition of /t/ and /l/. In J. Slitka, S. Manuel, & M. Matthies (Eds.), *From sound to sense* (pp. C181–C186). Cambridge, MA: MIT Press.
- Marler, P. (1991). The instinct to learn. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 37–66). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Mayberry, R. I. & Lock, E. (2003). Age constraints on first versus second language acquisition: Evidence for linguistic plasticity and epigenesis. *Brain Language*, 87, 369–384.
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82, B101–B111.
- Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, 31, 838–850.
- Meltzoff, A. N. (2007). The "like me" framework for recognizing and becoming an intentional agent. *Acta Psychologica*, 124, (26–43).
- Meltzoff, A. N. & Decety, J. (2003). What imitation tells us about social cognition: A rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society B*, 358, 491–500.
- Meltzoff, A. N., Kuhl, P. K., Movellan, J., & Sejnowski, T. (2009). Foundations for a new science of learning. *Science*, 17, 284–288.
- Meltzoff, A. N. & Moore, M. K. (1997). Imitation of facial and manual gestures by human neonates. *Science*, 198, 75–78.
- Mundy, P. & Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant Behavior and Development*, 21, 469–482.
- Nelson, D. & Marler, P. (1994). Selection-based learning in bird song development. *Proceedings of the National Academy of Sciences, USA*, 91, 10498–10501.
- Neville, H. J., Coffey, S.A., Lawson, D. S., Fischer, A., Emmorey, K., & Bellugi, U. (1997). Neural systems mediating American Sign Language: Effects of sensory experience and age of acquisition. *Brain and Language*, 57, 285–308.
- Newport, E. (1990). Maturation constraints on language learning. *Cognitive Science*, 14, 11–28.
- Newport, E. L., Bavelier, D., & Neville, H. J. (2001). Critical thinking about critical periods: Perspectives on a critical period for language acquisition. In E. Dupoux (Ed.), *Language, brain, and cognitive development: Essays in honor of Jacques Mehler* (pp. 481–502). Cambridge, MA: MIT Press.
- Newport, E. L. & Supalla, T. (1987). *A critical period effect in the acquisition of a primary language*. Unpublished manuscript.
- Pena, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., et al. (2003). Sounds and silence: An optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences*, 100, 11702–11705.
- Posner, M. I. (Ed.) (2004). *Cognitive neuroscience of attention*. New York: Guilford Press.
- Pulvermuller, F. (2005). Brain mechanisms linking language to action. *Nature Reviews Neuroscience*, 6, 574–582.
- Rabiner, L. R. & Huang, B. H. (1993). *Fundamentals of speech recognition*. Englewood Cliffs, NJ: Prentice Hall.
- Rivera-Gaxiola, M., Klarman, L., Garcia-Sierra, A., & Kuhl, P. K. (2005). Neural patterns to speech and vocabulary growth in American infants. *NeuroReport*, 16, 495–498.
- Rivera-Gaxiola, M., Silvia-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science*, 8, 162–172.
- Rizzolatti, G. (2005). The mirror neuron system and imitation. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From neuroscience to social science—Volume 1: Mechanisms of imitation and imitation in animals* (pp. 55–76). Cambridge, MA: MIT Press.
- Rizzolatti, G. & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–192.
- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3, 131–141.
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2002). Motor and cognitive functions of the ventral premotor cortex. *Current Opinion in Neurobiology*, 12, 149–154.
- Roseberry, S., Nash, B., Garcia-Sierra, A. & Kuhl, P. K. (in preparation). Can infants learn phonetically from interactive TV?
- Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month old infants. *Science*, 274, 1926–1928.
- Saffran, J. R., Johnson, E. K., Aslin, R. N. & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70, 27–52.
- Saffran, J. R., Werker, J. F., & Werner, L. A. (2006). The infant's auditory world: Hearing, speech, and the beginnings of language. In W. Damon & R. M. Lerner (Series Eds.) & R. Siegler & D. Kuhn (Vol. Eds.), *Handbook of child psychology: Vol. 2, Cognition, perception and language* (6th ed., pp. 58–108). New York: Wiley.
- Skinner, B. F. (1957). *Verbal Behavior*. Acton, MA: Copely Publishing Group.
- Stevens, R., Mehus, S., Mertl, V., Conboy, B., & Kuhl, P. K. (submitted). An Interaction analysis of joint attention and joint action in a laboratory experiment with infants: Results from an interdisciplinary collaboration across the ethnographic and experimental divide.
- Tanaka, F., Cicourel, A., & Movellan, J. (2007). Socialization between toddlers and robots at an early childhood education center. *Proceedings of the National Academy of Sciences, US*, 104, 17954–17958.
- Tchernichovski, O., Mitra, P., Lints, T., & Nottebohm, F. (2001). Dynamics of the vocal imitation process: How a Zebra Finch learns its song. *Science*, 291, 2564–2569.

- Tomasello, M. (2003a). *Constructing a language*. Cambridge, MA: Harvard University Press.
- Tomasello, M. (2003b). The key is social cognition. In D. Gentner & S. Kuczaj (Eds.), *Language and thought* (pp. 47–58). Cambridge, MA: MIT Press.
- Trevarthen, C. & Hubley, P. (1978). Secondary intersubjectivity: Confidence, confiding, and acts of meaning in the first year. In A. Lock (Ed.), *Action, gesture, and symbol* (pp. 183–229). London: Academic Press.
- Tsao, F.-M., Lui, H.-M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development, 75*, 1067–1084.
- Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *Journal of the Acoustical Society of America, 120*, 2285–2294.
- Ward, N., Sundara, M., Conboy, B., & Kuhl, P. K. (2009, October). *Consequences of short-term language exposure in infancy on babbling*. Poster presented at the 158th meeting of the Acoustical Society of America, San Antonio.
- Watson, J. S. (1979). Perception of contingency as a determinant of social responsiveness. In E. B. Thomas (Ed.), *The origins of social responsiveness* (pp. 33–64). New York: Erlbaum.
- Watson, J. S. (2005). The elementary nature of purposive behavior: Evolving minimal neural structures that display intrinsic intentionality. *Evolutionary Psychology, 3*, 24–48.
- Weber-Fox, C. M. & Neville, H. J. (1999). Functional neural subsystems are differentially affected by delays in second language immersion: ERP and behavioral evidence in bilinguals. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 23–38). Mahwah, NJ: Lawrence Erlbaum and Associates, Inc.
- Werker, J. F. & Lalonde, C. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology, 24*, 672–683.
- Werker, J.F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition, 103*, 147–162.
- 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. & King, A. (1988). Female visual displays affect the development of male song in the cowbird. *Nature, 334*, 244–246.
- Woolley, S. C. & Doupe, A. J. (2008). Social context-induced song variation affects female behavior and gene expression. *PLoS Biol* 6, e62.
- Yeni-Komshian, G. H., Flege, J. E., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean–English bilinguals. *Bilingualism: Language and Cognition, 3*, 131–149.
- Yoshida, K. A., Pons, F., Cady, J. C., & Werker, J. F. (2006). *Distributional learning and attention in phonological development*. Paper presented at International Conference on Infant Studies, Kyoto, Japan, 19–23 June.
- Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., & Tohkura, Y. (2005). Effects of language experience: Neural commitment to language-specific auditory patterns. *NeuroImage, 26*, 703–720.
- Zhang, Y., Kuhl, P. K., Imada, T., Iverson, P., Pruitt, J., Stevens, E., et al. (2009). Neural signatures of phonetic learning in adulthood: A magnetoencephalography study. *NeuroImage, 46*, 226–240.