Early Language Learning and the Social Brain

Patricia K. Kuhl

Institute for Learning & Brain Sciences, University of Washington, Seattle, Washington 98195

Correspondence: pkkuhl@u.washington.edu

Explaining how every typically developing child acquires language is one of the grand challenges of cognitive neuroscience. Historically, language learning provoked classic debates about the contributions of innately specialized as opposed to general learning mechanisms. Now, new data are being brought to bear from studies that employ magnetoencephalograph (MEG), electroencephalograph (EEG), magnetic resonance imaging (MRI), and diffusion tensor imaging (DTI) studies on young children. These studies examine the patterns of association between brain and behavioral measures. The resulting data offer both expected results and surprises that are altering theory. As we uncover what it means to be human through the lens of young children, and their ability to speak, what we learn will not only inform theories of human development, but also lead to the discovery of neural biomarkers, early in life, that indicate risk for language impairment and allow early intervention for children with developmental disabilities involving language.

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 language provoked classic debates on nature versus nurture by equally strong proponents of nativism (Chomsky 1959) and learning (Skinner 1957). Although 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 (Saffran et al. 2006; Kuhl 2010, 2011), we are only beginning to posit the mechanisms in the brain and mind of a child that enable language acquisition (e.g., Kuhl et al. 2014).

In this paper, I 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. Furthermore, I hypothesize that the social brain—in ways we have yet to understand—"gates" the computational mechanisms underlying learning in the domain of language. My goal here is to discuss recent results that advance our understanding of the mechanisms underlying the social influences on language learning in infants.

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, the gating hypothesis may explain why social factors play a far more significant role than previously realized in human learning beyond language and throughout our lifetimes (Meltzoff et al. 2009).

Using methods of systems neuroscience tailored for use with very young children, we can now explore how the integration of information across specialized brain systems such as language and social cognition might take place, and thus provide valuable data about the potent role social interaction plays in language learning. These approaches, as well as others described here, could lead us toward a view of language acquisition in the human child that would be transformational.

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 computation and social skills combine to form a very powerful learning mechanism. Interestingly, the new solution resembles neither Skinner's operant conditioning and reinforcement model of learning, nor Chomsky's nativist model. Instead, it is a view of the process that takes into account infants' innate skills to perceive all phonetic distinctions at birth, an ability that is not unique to humans (Kuhl and Miller 1975), and new data on children's extraordinary abilities to learn from computational and social mechanisms once exposed to speech (Kuhl et al. 2003).

SECOND LANGUAGE LEARNING EXHIBITS A "CRITICAL PERIOD" FOR LEARNING

A stage-setting concept for human language learning is the graph shown in Figure 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 with adults, despite adults' cognitively superiority. Language is one of the classic examples of a "critical" or "sensitive" period in neurobiology (Johnson and Newport 1989; Newport et al. 2001; Knudsen 2004; Kuhl 2004; Bruer 2008).

Scientists are in agreement that this curve represents data across a wide variety of language-learning studies (Johnson and Newport 1989; Bialystok and Hakuta 1994; Neville et al. 1997; Flege et al. 1999; Weber-Fox and Neville 1999; Yeni-Komshian et al. 2000; Birdsong and Molis 2001; Mayberry and Lock 2003; Kuhl et al. 2005, 2008; EL Newport and T Supalla, unpubl.). The learning function describes our current understanding of second language learning, and reflects 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 this phenomenon. What accounts for adults' inability to learn a new language with the facility of an infant?

One of the candidate explanations is Lenneberg's hypothesis that development of the corpus callosum changes the brain in a way that affects learning (Lenneberg 1967; Newport et al. 2001). More recent hypotheses take a different perspective. Newport raised the "less is more" hypothesis, which suggests that infants' limited cognitive capacities actually allow superior learning of the simpli-

fied language spoken to infants (Newport 1990). Work in my laboratory has led me to advance the concept of *neural commitment*, the idea that neural networks are established early in infancy to detect the phonetic and prosodic patterns of speech (Kuhl 2004; Kuhl et al. 2014). The neural architecture is designed to maximize processing of 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.

FOCAL EXAMPLE: PHONEME LEARNING

The world's languages contain approximately 600 consonants and 200 vowels (Ladefoged 2001). Each language uses a unique set of approximately 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*, that are functionally equivalent in the language. The infant's task is to make some progress in figuring out the composition of the 40-odd phonemic categories before trying to acquire words on which these elementary units depend. For example, Japanese infants must learn to group the phonetic units *r* and *l* into a single phonemic category (Japanese *r*), whereas American infants must learn to uphold the distinction to separate *rake* from *lake*. Similarly, Spanish-learning infants must distinguish phonetic units critical to Spanish words (*bano* and *pano*), 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 yr (Ferguson et al. 1992). Yet, by 10 mo 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 wk (Kuhl and 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 auditory representations guide infants' successive motor approximations until a match is achieved (Kuhl and 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 and Kujala 2009), and recent brain data obtained on infants support this view (Kuhl et al. 2014). We revisit the brain's sensorimotor coupling for speech later in this paper.
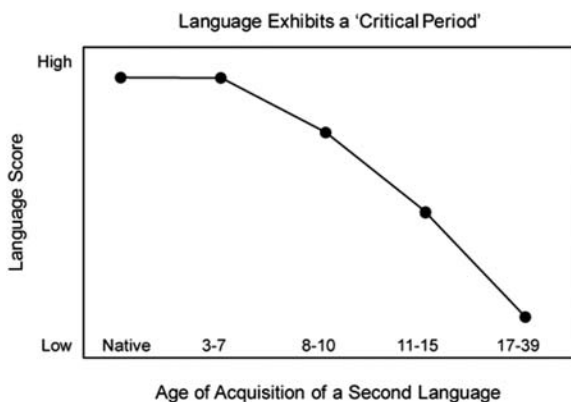
Figure 1. The relationship between age of acquisition of a second language and language skill. (Reprinted from Kuhl 2010; originally adapted from Johnson and Newport 1989, with permission from Elsevier.)

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 and Huang 1993), although models improve when exposed to "motherese," the linguistically simplified and acoustically exaggerated speech that adults universally use when speaking to infants (de Boer and Kuhl 2003; see also Kuhl et al. 1997). 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 et al. 2004; Werker et al. 2007). How do Japanese infants learn that these two sounds do not distinguish words, and that their differences should be ignored? Similarly, adult English speakers produce the Spanish *b* and *p*, exposing American infants to both categories of sound (Abramson and Lisker 1970). How do infants learn which sounds are important in distinguishing words in English? An important discovery in the 1970s showed that infants initially hear all these phonetic differences, but that this universal perceptual ability declines before the age of 1 yr (Eimas et al. 1971; Eimas 1975; Lasky et al. 1975; Werker and Lalonde 1988).

THE TIMING OF INFANT LEARNING

The transition from an early universal phonetic ability to language specific phonetic perception occurs between 6 and 12 mo of age, when nonnative phonetic perception declines (Werker and Tees 1984; Best and McRoberts 2003; Rivera-Gaxiola et al. 2005b; Tsao et al. 2006) and there is increased sophistication in the processing of native speech (Kuhl et al. 1992). Work in my laboratory showed that at the same time that nonnative perception for consonants declines, native-language speech perception shows a significant increase. We demonstrated that Japanese infants' discrimination of English *r*–*l* declined between 8 and 10 mo of age, whereas at the same time in development, American infants' discrimination of the same sounds showed an increase (Fig. 2; Kuhl et al. 2006).

Our work also demonstrated that native-language discrimination between 6 and 7 mo predicts the rate of lan-

guage growth between 11 and 30 mo (Tsao et al. 2004; Conboy et al. 2005; Kuhl et al. 2005, 2008; Rivera-Gaxiola et al. 2005a). Intriguingly, although better performance on *native* contrast discrimination predicts more rapid growth in later language abilities, better performance on *non*native contrasts predicts slower language growth (Fig. 3; Kuhl et al. 2005, 2008). In other words, phonetic learning does not simply depend on auditory acuity, but also on the ability to learn from exposure to language and to attend to some acoustic features and not others.

Bilingual learners acquire two languages, using the same principles, although our data suggest that the sensitive period for learning is extended (García-Sierra et al. 2011). We are currently investigating the extended sensitive period in bilinguals: The neural system may remain "open" longer due to variability in overall language input or due to reduction in input per unit time for each of the native languages, thus extending the learning process (see Kuhl et al. 2008, for discussion).

Both monolingual and bilingual infants benefit from hearing "motherese" speech. Produced at a higher pitch, with exaggerated intonation contours and slower articulation, motherese is an ideal signal for infant learning. Produced by parents and caregivers all over the world when speaking to infants, motherese exaggerates the acoustic differences between phonetic units, making phonemes easier to differentiate (Kuhl et al. 1997). Our studies link exposure to motherese to advanced language development. We demonstrated that 11- and 14-mo-old infants exposed to a greater proportion of motherese utterances at home, and particularly during face-to-face social interactions, have advanced language development at the age of 2 yr (Ramírez-Esparza et al. 2014, 2015).

A COMPUTATIONAL SOLUTION TO PHONETIC LEARNING

What explains infant learning? Studies in the decade of the 1990s demonstrated that infants are capable of a new form of learning, referred to as "statistical learning" (Saffran et al. 1996). Statistical learning 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 studies show that both phonetic learning (Maye et al. 2002) and early word learning (Saffran et al. 1996) are strongly influenced by this form of learning.

Recent studies show that infants are sensitive to the statistical distributions of sounds that they hear in ambient language, and that 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 five—but speakers of these languages produce a wide range of sounds. The mere presence of a particular
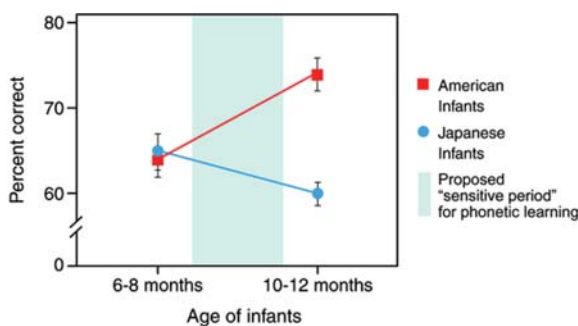
Figure 2. Effects of age on discrimination of the American English /ra-la/ phonetic contrast by American and Japanese infants at 6–8 and 10–12 mo of age. Mean percent correct scores are shown with standard errors indicated. (Reprinted from Kuhl 2010, with permission from Elsevier; originally adapted from Kuhl et al. 2006.)
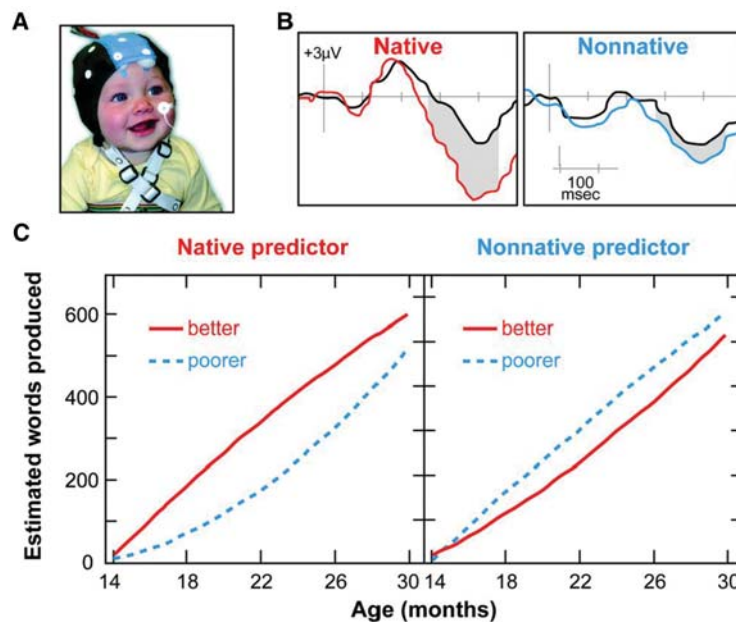
Figure 3. (*A*) A 7.5-mo-old infant wearing an ERP electrocap. (*B*) Infant ERP waveforms at one sensor location (CZ) for one infant are shown in response to a native (English) and nonnative (Mandarin) phonetic contrast at 7.5 mo. The mismatch negativity (MMN) is obtained by subtracting the standard 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 mo for MMN values of +1 S.D. and −1 S.D. on the native contrast at 7.5 mo (*C, left*) and vocabulary growth for MMN values of +1 S.D. and −1 S.D. on the nonnative contrast at 7.5 mo (*C, right*). Analyses show that both contrasts predict vocabulary growth and that the effects of better discrimination are reversed for the native and nonnative contrasts. (*A*, Reprinted from Kuhl 2010, with permission from Elsevier; *C*, adapted from Kuhl et al. 2008.)

sound, therefore, does not account for infant learning. However 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 mo of age show a unique response to the "prototypical" vowels in their language—those that are not only the distributional mean in productions of adults but are also perceptually preferred by adults; this response is not shown to foreign-language vowel prototypes (Kuhl et al. 1992). Moreover, when tested with very simple stimuli in the laboratory, infants have been shown to learn from distributional patterns in language input after short-term exposure to phonetic stimuli (Maye et al. 2002). Word learning is also supported by the detection of statistical regularities in speech, before infants understand word meanings (Saffran et al. 1996). Statistical learning is not limited to humans (Hauser et al. 2001), nor to speech; it operates for musical and visual patterns in the same way (Saffran et al. 1999; Fiser and Aslin 2002; Kirkham et al. 2002).

EFFECTS OF A SOCIAL CONTEXT ON SPEECH LEARNING

Although statistical learning provides a potential explanation for phonetic learning, studies in my laboratory suggest that learning from complex natural language requires more. Our studies indicate that statistical learning is not sufficient for phonetic learning from complex natural language input: Social interaction is necessary. In other words, infants appear to gain something critical from social context in natural language-learning situations that is not available in the absence of another human being (Kuhl et al. 2003).

We posed a simple question—can infants learn phonetically from first-time natural exposure to a foreign language at 9 mo? We wondered whether statistical learning at 9 mo required a long-term history of listening to that language—we reasoned that infant learning at this age might depend on the cumulative statistical distributions over the initial 9 mo of life. Alternatively, the transition might occur at 9 mo because a learning process initially became available to the child at the age of 9 mo.

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 a 4- to 5-wk period. The foreign language "tutors" read books and played with toys in sessions that

were unscripted. A control group was also exposed to language 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 with 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 mo (Kuhl et al. 2003).

Learning was robust and durable. Infants returned to the laboratory for their behavioral discrimination tests between 2 and 12 d after the final exposure session, with a median of 6 d, and for their ERP measures between 8 and 33 d following the last exposure session, with a median of 15 d. These delays allowed us to examine whether longer periods between exposure and test resulted in poorer Mandarin speech 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 ERP brain measures. No "forgetting" of the Mandarin contrast occurred during the 2- to 33-d delay.

Infants' remarkable learning led us to test two additional conditions. In one, infants were exposed to the same information on the same schedule in the absence of a human being, via television. In another, infants heard an audiotaped presentation of the same information. The results showed that 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 audiotape only, showed no learning—their performance equaled that of infants in the control group who had not been exposed to Mandarin at all (Fig. 4).

Thus, the presence of a human being interacting with the infant during language exposure, although not required for simpler statistical-learning tasks (Saffran et al. 1996; Maye et al. 2002), is critical for learning in
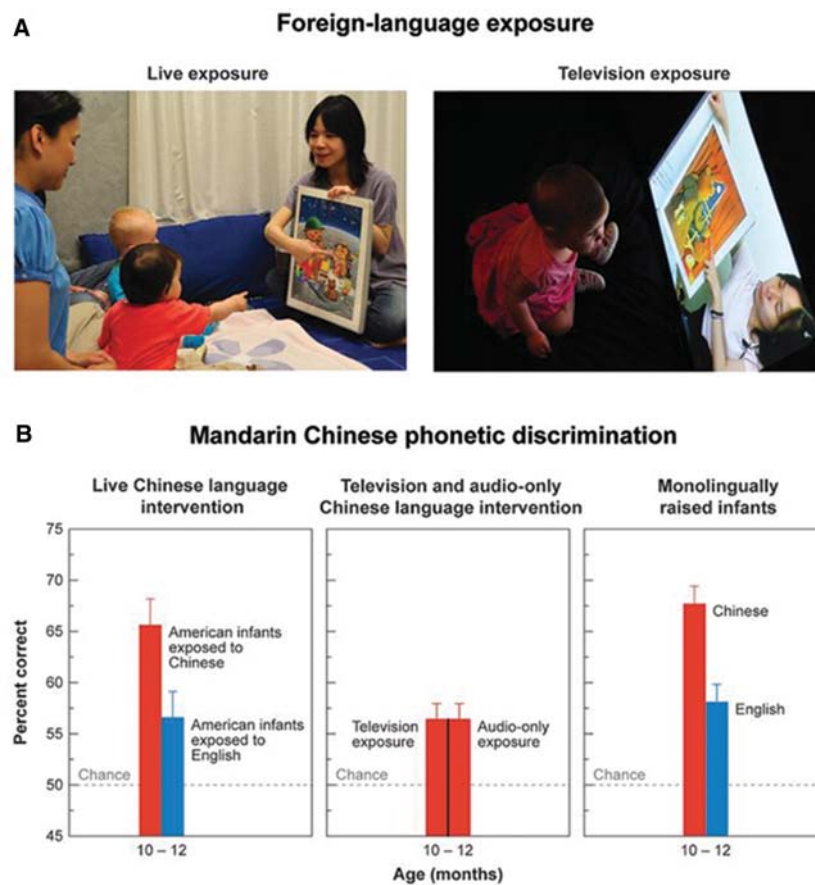
Figure 4. The need for social interaction in language acquisition is shown by foreign-language learning experiments. Nine-month-old infants experienced 12 sessions of Mandarin Chinese through (*A*) natural interaction with a Chinese speaker (*left*) or the identical linguistic information delivered via television (*right*) or audiotape (not shown). (*B*) Natural interaction resulted in significant learning of Mandarin phonemes when compared with a control group who experience interaction but heard only 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*). (*A,B*, Reprinted from Kuhl 2010, with permission from Elsevier; *B*, originally adapted from Kuhl et al. 2003.)

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- to 5-wk period (Kuhl et al. 2003).

EXPLAINING THE EFFECTS OF SOCIAL INTERACTION ON LANGUAGE LEARNING

I proposed a *Social Gating Hypothesis* (Kuhl 2007) to explain how social interaction might influence language learning. Two broad mechanisms—motivational and informational—were cited as possible vehicles for social learning. They are not mutually exclusive. Recent evidence suggests that both play a role.

Regarding motivation 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). In recent tests, we examined whether increasing infants' attention to the screen would result in increased learning. The new experiment used stimulus recordings from the same television presentations we used in the original study, and a touch-screen TV to give infants control over the presentation of the Mandarin speaker—infants turned on a 20-sec video presentation simply by touching the screen (SR Lytle, A García-Sierra, and PK Kuhl, in prep.). Other details of the experiment were identical to the original—12 sessions were conducted between the ages of 9.5 and 10.5 mo of age. We expected that the contingent presentation of a 20-sec clip of speech would increase learning, although not equal to that produced with social exposure.

The experiment was also designed to test a second hypothesis, that the presence of another baby while learning—in other words, two-baby learning—would enhance performance over one-baby learning. The critical manipulation was that half the infants in the touch-screen experiment were tested as solo learners, whereas the other half were tested in two-baby pairs in which one of the infants was designated the "target" learner. We expected that the target babies in the two-baby situation would show higher learning than infants tested in the solo learning condition. As in our original experiments, infant learning was tested both neurally, using event-related potential (ERP) methods (see Kuhl 2011), and using our standard behavioral method, the conditioned head-turn (HT) technique (see Kuhl et al. 2003 for details).

The results show that contingent presentation of the recorded information does not produce behavioral evidence of learning. Neither the one- or two-baby conditions produced significant HT performance that exceeded chance.

However, the neural data show very clearly that two are better than one—infants tested in pairs rather than individually showed more negative ERP responses to the Mandarin contrasts, consistent with those of an adult native speaker. Moreover, detailed examination of ERP data

from the one- versus two-baby learning sessions revealed an additional important result. ERP data from infants tested in the one- versus two-baby learning sessions showed two different patterns of ERP responses to phonetic stimuli that have been previously reported in the literature. One-baby learners showed a positive mismatch response (MMR) that our laboratory and others interpret as increased auditory attention to an acoustic stimulus, but not one associated with phonetic learning (Morr et al. 2002). Our data show that positive MMR responses become a more negative with increased exposure to a phonetic contrast (García-Sierra et al. 2011). Infants in the two-baby learning condition showed a negative MMR in response to the phonetic contrast, a response that resembles the negative MMR shown in phonetic tests on monolingual adults and infants (Näätänen et al. 1997). We interpret this to mean that infants tested in the two-baby condition are learning at a faster rate as a function of social interaction.

But what kind of social interaction produces this accelerated learning? Do infants in the two-baby learning condition collaborate, showing each other how to press the screen to produce the video? Or is it related to social arousal, produced by the mere presence of another infant? Detailed coding of the video and audio records by trained assistants who were blind to the learning outcomes support the social arousal explanation. We tracked the number of novel partners each target baby experienced during the paired sessions. We found a statistically significant relationship between the number of novel partners a baby experienced in paired sessions and the strength of that infant's negative MMR response to the Mandarin contrasts—the greater the number of novel partners, the stronger the negative MMR, indicating greater evidence of learning. Enhanced learning was not a function of the number of videos viewed by the infants, the duration of exposure to the videos, the number of screen touches, or infants' attention to the screen.

Infants paired with other babies while learning not only show enhanced brain measures when reacting to the foreign-language phonemes, paired babies show significantly higher numbers of speech vocalizations, supporting the idea that the social presence of another infant induces arousal, social communication, and increased learning. Our data suggest that infants are socially aroused in the presence of a peer baby, that their arousal increases when paired with a novel peer, and that social arousal increases learning in infants.

Social interaction also provides additional information, which helps explain learning enhancement in the social condition. During live exposure, tutors focus their visual gaze on pictures in the books or on the toys they talk about, and beginning at this point in development, infants' gaze tends to follow the speaker's gaze (Baldwin 1995; Brooks and Meltzoff 2002, 2015). 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. We hypothesized that 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-mo-old infants to Spanish, and measured gaze following in the infants during the learning sessions. Spanish exposure confirmed learning for the phonetic units of Spanish (Conboy and Kuhl 2011), and also revealed that infants' eye-gaze behavior during Spanish exposure predicted the degree of language learning as shown by the ERP brain measures (Conboy et al. 2015). 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 at the toy.

BRAIN MECHANISMS UNDERLYING PHONETIC LEARNING

I hypothesized that social signals may activate brain mechanisms of social understanding that link perception and action (Hari and 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 (Rizzolatti and Craighero 2004; Pulvermuller 2005; Rizzolatti 2005) have reinvigorated this tradition.

Our recent MEG brain imaging work on infants indicates that the brain systems involved in speech production are activated when infants listen to speech (Kuhl et al. 2014). We tested infants who straddled the change from universal speech perception to language specific speech perception, at 7 and 11 mo (Fig. 5). We found that listening to speech at both ages activated not only auditory regions in the brain, but also regions involved in motor planning, Broca's area and the cerebellum. Developmental change was also observed—7-mo-old infants activated both auditory and motor areas equally for native and nonnative speech, whereas 11-mo-old infants show greater activation in auditory areas for native speech, and greater activation for nonnative speech in motor areas.

The study advanced our understanding of infant speech processing, and raised additional questions. First, at the earliest age tested (7 mo of age), both native and nonna-

Figure 5. A 7-mo-old infant in the MEG machine listening to language. (Courtesy of P.K. Kuhl, UW Institute for Learning & Brain Sciences ©2011.)

tive speech activated motor brain areas, and equivalently, so the activation of motor brain areas in response to speech at 7 mo of age is not limited to sounds that infants hear in ambient language. At this early age, infants' motor brain areas appear to be reacting to all sounds with speech-like qualities. We argue that infants are engaged in something like "analysis by synthesis," a concept with origins in the field of artificial intelligence (see Stevens and Halle 1967). Analysis by synthesis proposed that speech perception involves a dual "hypothesize and test" process. Bottom-up *analysis* and top-down *synthesis* jointly and actively constrain perceptual interpretation. On this account, listeners generate an *internal model* of the motor commands needed to produce a given auditory signal, in essence, a "guess" or prediction about the auditory input. On this view, an internally generated motor hypothesis is based on a listener's experience producing speech, and then tested against incoming auditory data. We argue that infants' brain responses are also indicative of the creation of internal motor models for speech.

How are 7-mo-old infants capable of creating motor models for speech? We argue that speech production experience—the early "cooing" that occurs by 12 wk of age and allows infants to imitate vowels by 20 wk (Kuhl and Meltzoff 1996)—yields a nascent auditory–articulatory map, an emergent "schema" that specifies generative rules relating articulatory movements to sound. This emerging auditory–articulatory map is likely highly abstract, but it allows infants to generate internal motor models when they listen to sounds with speech-like qualities, regardless of whether they have experienced those specific speech sounds or not. Thus, we posit that infants' nascent speech motor experience is the catalyst for the effects observed in infants as young as 7 mo of age.

Our MEG data also demonstrate developmental change. By the end of the first year, infants' auditory and motor brain areas pattern differently for native and nonnative sounds: Auditory areas show greater activation for native sounds, whereas motor areas show greater activation for nonnative sounds, suggesting that these two brain systems are coding different kinds of information in response to speech. We offer the following tentative explanation: Hearing native speech throughout the first year increases infants' auditory sensitivity to native speech-sound differences. At the same time, hearing speech strengthens sensorimotor pairings, because incoming native speech signals prompt internally generated speech motor models based on infants' prior speech production experience. Language experience thus serves to strengthen knowledge of native-language speech—both perceptual and motor. By the end of the first year, the emergent schema relating speech motor movements to sound would describe native speech quite well. However, nonnative speech would not be well described by the developing schema, making it more difficult to generate internal models for nonnative speech. Hence our finding that nonnative speech elicits greater cortical activation (indicating greater effort) in Broca's area and the cerebellum than native speech.

Early models of developmental speech perception, such as my native language neural commitment (NLNC)

concept (Kuhl 2004), described a process of "neural commitment" to the auditory patterns of native speech. Revisions in the model, named native language magnet-expanded (NLM-e) described emergent links between speech perception and production (Kuhl et al. 2008). The new data will allow further refinement of the model by suggesting *how* speech perception and speech production become linked early in development: Infants' brains respond to hearing speech by activating motor brain areas, co-registering perceptual experience and motor brain patterns. When infants generate vocalizations, they build a schema that predicts how moving their tongues and lips during vocalizations will modify the sounds they create. By 20 wk, infants can imitate pitch patterns (Kuhl and Meltzoff 1982) and vowels (Kuhl and Meltzoff 1996) that acoustically and perceptually correspond to the ones they hear, indicating knowledge of speech production. This articulatory knowledge, we argue, allows infants to generate internal motor models of speech as they listen to us talk. On this view, both auditory and motor components contribute to the developmental transition in speech perception that occurs at the end of the first year of life.

Theories of social understanding in adults (Hari and Kujala 2009) and infants (Marshall and Meltzoff 2014), suggest that humans evolved brain mechanisms to detect and interpret human actions, behaviors, movements, and sounds. The current data contribute to these views by demonstrating that auditory speech activates motor areas in the infant brain. Motor brain activation in response to others' communicative signals could assist broader development of social understanding in humans.

Our findings offer an opportunity to test children with developmental disabilities such as autism spectrum disorder (Kuhl et al. 2013; Stavropoulos and Carver 2013) whose social and language deficits are potentially associated with a decreased ability to activate motor brain systems in response to human signals. The acquisition of language requires social learning, and therefore speech signals may provide excellent "biomarkers" of early risk for children with autism spectrum disorder (ASD), and our data show promise in this regard (Kuhl et al. 2013).

NEUROBIOLOGICAL FOUNDATIONS OF SOCIAL LEARNING

Humans are not the only species in which communicative learning is affected by social interaction, and the parallels between song learning in birds and speech learning in infants has been a rich area for theory (Doupe and Kuhl 1999). The work of Allison Doupe and her colleagues has beautifully illustrated the role of social contexts on vocal learning in the zebra finch (Woolley and Doupe 2008; Kojima and Doupe 2011). Moreover, studies on the zebra finch have also shown that motor learning occurs earlier and is more sophisticated than previously thought (Raghav and Doupe 2013). Theorizing in these two closely related fields is likely to advance based on these kinds of cross-species comparisons.

CONCLUSIONS

Studies on the development of human speech and language offer a unique view of the human mind and brain, one made tractable through the design of behavior and brain experiments on learning. Innate predisposition, rapid and robust learning during "sensitive" periods, and brain systems that link action and perception all play a role in the developing story that connects the dots regarding infant acquisition of one or more languages. Studies at the phonetic level of speech—designed to evaluate infant learning of the sounds used to convey meaning in their language environment—are tractable and allow cross-species and cross-cultural developmental experiments. These show strong auditory predispositions that exist at birth, allowing infants to discriminate all possible phonetic variations used in the world's languages. Equally impressive is the demonstration that infants learn rapidly and robustly from linguistic exposure, and that the timing and the context are both critical. The power of social interaction to induce language learning in infants is new, and raises intriguing questions. Recent MEG brain imaging results on infants listening to speech show that hearing speech activates motor planning areas (Brocas and the cerebellum) in the infant brain. The emerging picture is complex. Infants begin life prepared for any language, and have an extraordinary capacity to learn from experience, especially optimized experience—motherese in a social context during sensitive periods, perhaps because action–perception systems in the human brain enable learning from other humans in a special way. Experiments on speech may advance theoretical explanations for the strong influence of the "social brain" on human learning more generally.

ACKNOWLEDGMENTS

REFERENCES

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

Baldwin DA. 1995. Understanding the link between joint attention and language. In *Joint attention: Its origins and role in development* (ed. Moore C, Dunham PJ), pp. 131–158. Lawrence Erlbaum Associates, Hillsdale, NJ.

Best C, McRoberts GW. 2003. Infant perception of nonnative consonant contrasts that adults assimilate in different ways. *Lang Speech* **46:** 183–216.

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

Birdsong D, Molis M. 2001. On the evidence for maturational constraints in second-language acquisitions. *J Mem Lang* **44:** 235–249.

Brooks R, Meltzoff AN. 2002. The importance of eyes: How infants interpret adult looking behavior. *Dev Psychol* **38:** 958–966.

Brooks R, Meltzoff AN. 2015. Connecting the dots from infancy to childhood: A longitudinal study connecting gaze following, language, and explicit theory of mind. *J Exp Child Psychol* **130:** 67–78.

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

Chomsky N. 1959. Review of "Verbal behavior" by B.F. Skinner. *Language* **35:** 26–58.

Conboy BT, Kuhl PK. 2011. Impact of second-language experience in infancy: Brain measures of first- and second-language speech perception. *Dev Sci* **14:** 242–248.

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

Conboy BT, Brooks R, Meltzoff AN, Kuhl PK. 2015. Social interaction in infants' learning of second-language phonetics: An exploration of brain-behavior relations. *Dev Neuropsychol* **40:** in press.

de Boer B, Kuhl PK. 2003. Investigating the role of infant-directed speech with a computer model. *Acoust Res Lett Online (ARLO)* **4:** 129–134.

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

Doupe AJ, Kuhl PK. 1999. Birdsong and human speech: Common themes and mechanisms. *Annu Rev Neurosci* **22:** 567–631.

Eimas PD. 1975. Auditory and phonetic coding of the cues for speech: Discrimination of the /r–l/ distinction by young infants. *Percept Psychophys* **18:** 341–347.

Eimas PD, Siqueland ER, Jusczyk P, Vigorito J. 1971. Speech perception in infants. *Science* **171:** 303–306.

Ferguson CA, Menn L, Stoel-Gammon C, ed. 1992. *Phonological development: Models, research, implications*. York Press, Timonium, MD.

Fiser J, Aslin RN. 2002. Statistical learning of new visual feature combinations by infants. *Proc Natl Acad Sci* **99:** 15822–15826.

Flege JE. 1991. Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *J Acoust Soc Am* **89:** 395–411.

Flege JE, Yeni-Komshian GH, Liu S. 1999. Age constraints on second-language acquisition. *J Mem Lang* **41:** 78–104.

García-Sierra A, Rivera-Gaxiola M, Percaccio CR, Conboy BT, Romo H, Klarman L, Ortiz S, Kuhl PK. 2011. Bilingual language learning: An ERP study relating early brain responses to speech, language input, and later word production. *J Phon* **39:** 546–557.

Hari R, Kujala M. 2009. Brain basis of human social interaction: From concepts to brain imaging. *Physiol Rev* **89:** 453–479.

Hauser MD, Newport EL, Aslin RN. 2001. Segmentation of the speech stream in a nonhuman primate: Statistical learning in cotton-top tamarins. *Cognition* **78:** B53–B64.

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. *Cogn Psychol* **21:** 60–99.

Kirkham NZ, Slemmer JA, Johnson SP. 2002. Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition* **83:** B35–B42.

Knudsen EI. 2004. Sensitive periods in the development of the brain and behavior. *J Cogn Neurosci* **16:** 1412–1425.

Kojima S, Doupe AJ. 2011. Social performance reveals unexpected vocal competency in young songbirds. *Proc Natl Acad Sci* **108:** 1687–1692.

Kuhl PK. 2004. Early language acquisition: Cracking the speech code. *Nat Rev Neurosci* **5:** 831–843.

Kuhl PK. 2007. Is speech learning 'gated' by the social brain? *Dev Sci* **10:** 110–120.

Kuhl PK. 2010. Brain mechanisms in early language acquisition. *Neuron* **67:** 713–722.

Kuhl PK. 2011. Social mechanisms in early language acquisition: Understanding integrated brain systems supporting language. In *The Oxford handbook of social neuroscience* (ed. Decety J, Cacioppo J), pp. 649–667. Oxford University Press, Oxford.

Kuhl PK, Meltzoff AN. 1982. The bimodal perception of speech in infancy. *Science* **218:** 1138–1141.

Kuhl PK, Meltzoff AN. 1996. Infant vocalizations in response to speech: Vocal imitation and developmental change. *J Acoust Soc Am* **100:** 2425–2438.

Kuhl PK, Miller JD. 1975. Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science* **190:** 69–72.

Kuhl PK, Williams KA, Lacerda F, Stevens KN, Lindblom B. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* **255:** 606–608.

Kuhl PK, Andruski JE, Chistovich IA, Chistovich LA, Kozhevnikova EV, Ryskina VL, Stolyarova EI, Sundberg U, Lacerda F. 1997. Cross-language analysis of phonetic units in language addressed to infants. *Science* **277:** 684–686.

Kuhl PK, Tsao F-M, Liu H-M. 2003. Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proc Natl Acad Sci* **100:** 9096–9101.

Kuhl PK, Conboy BT, Padden D, Nelson T, Pruitt J. 2005. Early speech perception and later language development: Implications for the 'critical period'. *Lang Learn Dev* **1:** 237–264.

Kuhl PK, Stevens E, Hayashi A, Deguchi T, Kiritani S, Iverson P. 2006. Infants show facilitation for native language phonetic perception between 6 and 12 months. *Dev Sci* **9:** 13–21.

Kuhl PK, Conboy BT, Coffey-Corina S, Padden D, Rivera-Gaxiola M, Nelson T. 2008. Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philos Trans R Soc B* **363:** 979–1000.

Kuhl PK, Coffey-Corina S, Padden D, Munson J, Estes A, Dawson G. 2013. Brain responses to words in 2-year-olds with autism predict developmental outcomes at age 6. *PLoS ONE* **8:** e64967.

Kuhl PK, Ramirez RR, Bosseler A, Lin J-F, Imada T. 2014. Infants' brain responses to speech suggest analysis by synthesis: Developmental MEG studies. *Proc Natl Acad Sci* **111:** 11238–11245.

Ladefoged P. 2001. *Vowels and consonants: An introduction to the sounds of language*. Blackwell Publishers, Oxford.

Lasky RE, Syrdal-Lasky A, Klein RE. 1975. VOT discrimination by four to six and a half month old infants from Spanish environments. *J Exp Child Psychol* **20:** 215–225.

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

Lotto AJ, Sato M, Diehl R. 2004. Mapping the task for the second language learner: The case of Japanese acquisition of /r/ and /l/. In *From sound to sense* (ed. Slitka J, Manuel S, Matthies M), pp. C181–C186. MIT Press, Cambridge, MA.

Marshall PJ, Meltzoff AN. 2014. Neural mirroring mechanisms and imitation in human infants. *Philos Trans R Soc B* **369:** 20130620.

Mayberry RI, Lock E. 2003. Age constraints on first versus second language acquisition: Evidence for linguistic plasticity and epigenesis. *Brain Lang* **87:** 369–384.

Maye J, Werker JF, Gerken L. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* **82:** B101–B111.

Meltzoff AN, Kuhl PK, Movellan J, Sejnowski T. 2009. Foundations for a new science of learning. *Science* **17:** 284–288.

Morr ML, Shafer VL, Kreuzer JA, Kurtzberg D. 2002. Maturation of mismatch negativity in typically developing infants and preschool children. *Ear Hear* **23:** 118–136.

Näätänen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A, Vainio M, Alku P, Ilmoniemi RJ, Luuk A, et al. 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* **385:** 432–434.

Neville HJ, Coffey SA, Lawson DS, Fischer A, Emmorey K, Bellugi U. 1997. Neural systems mediating American Sign Language: Effects of sensory experience and age of acquisition. *Brain Lang* **57:** 285–308.

Newport EL. 1990. Maturational constraints on language learning. *Cogn Sci* **14:** 11–28.

Newport EL, Bavelier D, Neville HJ. 2001. Critical thinking about critical periods: Perspectives on a critical period for language acquisition. In *Language, brain, and cognitive development: Essays in honor of Jacques Mehlter* (ed. Dupoux E), pp. 481–502. MIT Press, Cambridge, MA.

Posner MI, ed. 2004. *Cognitive neuroscience of attention*. Guilford Press, New York.

Pulvermuller F. 2005. Brain mechanisms linking language to action. *Nat Rev Neurosci* **6:** 574–582.

Rabiner LR, Huang BH. 1993. *Fundamentals of speech recognition*. Prentice Hall, Englewood Cliffs, NJ.

Raghav RR, Doupe AJ. 2013. Behavioral and neural signatures of readiness to initiate a learned motor sequence. *Curr Biol* **23:** 87–93.

Ramírez-Esparza N, García-Sierra A, Kuhl PK. 2014. Look who's talking: Speech style and social context in language input to infants are linked to concurrent and future speech development. *Dev Sci* **17:** 880–891.

Ramírez-Esparza N, García-Sierra A, Kuhl PK. 2015. The language you hear, the language you speak: A study with Spanish–English bilingual infants. *Child Dev* **8:** in press.

Rivera-Gaxiola M, Klarman L, García-Sierra A, Kuhl PK. 2005a. Neural patterns to speech and vocabulary growth in American infants. *Neuroreport* **16:** 495–498.

Rivera-Gaxiola M, Silva-Pereyra J, Kuhl PK. 2005b. Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Dev Sci* **8:** 162–172.

Rizzolatti G. 2005. The mirror neuron system and imitation. In *Perspectives on imitation: From neuroscience to social science—Volume 1: Mechanisms of imitation and imitation in animals* (ed. Hurley S, Chater N), pp. 55–76. MIT Press, Cambridge, MA.

Rizzolatti G, Craighero L. 2004. The mirror-neuron system. *Annu Rev Neurosci* **27:** 169–192.

Saffran J, Aslin R, Newport E. 1996. Statistical learning by 8-month old infants. *Science* **274:** 1926–1928.

Saffran JR, Johnson EK, Aslin RN, Newport EL. 1999. Statistical learning of tone sequences by human infants and adults. *Cognition* **70:** 27–52.

Saffran JR, Werker JF, Werner LA. 2006. The infant's auditory world: Hearing, speech, and the beginnings of language. In *Handbook of child psychology: Vol. 2, cognition, perception and language*, 6th ed. (series ed. Damon W, Lerner RM; vol. ed. Siegler R, Kuhn D), pp. 58–108. Wiley, New York.

Skinner BF. 1957. *Verbal behavior*. Appleton-Century-Crofts, New York.

Stavropoulos K, Carver L. 2013. Research review: Social motivation and oxytocin in autism—Implications for joint attention development and intervention. *J Child Psychol Psychiatry* **54:** 603–618.

Stevens KN, Halle M. 1967. Remarks on analysis by synthesis and distinctive features. In *Models for the perception of speech and visual form* (ed. Walthem-Dunn W), pp. 88–102. MIT Press, Cambridge, MA.

Tsao F-M, Lui H-M, Kuhl PK. 2004. Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Dev* **75:** 1067–1084.

Tsao F-M, Liu H-M, Kuhl PK. 2006. Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *J Acoust Soc Am* **120:** 2285–2294.

Weber-Fox CM, Neville HJ. 1999. Functional neural subsystems are differentially affected by delays in second language immersion: ERP and behavioral evidence in bilinguals. In *Second language acquisition and the critical period hypothesis* (ed. Birdsong D). Lawerence Erlbaum, Mahwah, NJ.

Werker JF, Lalonde C. 1988. Cross-language speech perception: Initial capabilities and developmental change. *Dev Psychol* **24:** 672–683.

Werker JF, Tees RC. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behav Dev* **7:** 49–63.

Werker JF, 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.

Woolley SC, Doupe AJ. 2008. Social context-induced song variation affects female behavior and gene expression. *PLoS Biol* **6:** e62.

Yeni-Komshian GH, Flege JE, Liu S. 2000. Pronunciation proficiency in the first and second languages of Korean–English bilinguals. *Bilingualism: Lang Cogn* **3:** 131–149.

Cold Spring Harbor Symposia on Quantitative Biology

Early Language Learning and the Social Brain

Patricia K. Kuhl

Cold Spring Harb Symp Quant Biol published online May 5, 2015
Access the most recent version at doi:10.1101/sqb.2014.79.024802

P<P	Published online May 5, 2015 in advance of the print journal.
Creative Commons License	
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To subscribe to *Cold Spring Harbor Symposia on Quantitative Biology* go to:
http://symposium.cshlp.org/subscriptions