

Brain Mechanisms in Early Language Acquisition

Patricia K. Kuhl^{1,*}

¹Institute for Learning & Brain Sciences, University of Washington, Seattle, WA 98195, USA

*Correspondence: pkkuhl@u.washington.edu

DOI 10.1016/j.neuron.2010.08.038

The last decade has produced an explosion in neuroscience research examining young children's early processing of language. Noninvasive, safe functional brain measurements have now been proven feasible for use with children starting at birth. The phonetic level of language is especially accessible to experimental studies that document the innate state and the effect of learning on the brain. The neural signatures of learning at the phonetic level can be documented at a remarkably early point in development. Continuity in linguistic development from infants' earliest brain responses to phonetic stimuli is reflected in their language and prereading abilities in the second, third, and fifth year of life, a finding with theoretical and clinical impact. There is evidence that early mastery of the phonetic units of language requires learning in a social context. Neuroscience on early language learning is beginning to reveal the multiple brain systems that underlie the human language faculty.

Introduction

Neural and behavioral research studies show that exposure to language in the first year of life influences the brain's neural circuitry even before infants speak their first words. What do we know of the neural architecture underlying infants' remarkable capacity for language and the role of experience in shaping that neural circuitry?

The goal of the review is to explore this topic, focusing on the data and arguments about infants' neural responses to the consonants and vowels that make up words. Infants' responses to these basic building blocks of speech—the phonemes used in the world's languages—provide an experimentally tractable window on the roles of nature and nurture in language acquisition. Comparative studies at the phonetic level have allowed us to examine the uniqueness of humans' language processing abilities. Moreover, infants' responses to native and nonnative phonemes have documented the effects of experience as infants are bathed in a specific language. We are also beginning to discover how exposure to two languages early in infancy produces a bilingual brain. We focus here on when and how infants master the sound structure of their language(s), and the role of experience in explaining this important developmental change. As the data attest, infants' neural commitment to the elementary units of language begins early, and the review showcases the extent to which the tools of modern neuroscience are advancing our understanding of infants' uniquely human capacity for language.

Humans' capacity for speech and language provoked classic debates on nature versus nurture by strong proponents of nativism (Chomsky, 1959) and learning (Skinner, 1957). While we are far beyond these debates and 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 et al., 2006), we are still just breaking ground with regard to the neural mechanisms that underlie language development (see Friederici and Wartenburger, 2010; Kuhl and Rivera-Gaxiola, 2008). This decade may represent the dawn of

a golden age with regard to the developmental neuroscience of language in humans.

Windows to the Young Brain

The last decade has produced rapid advances in noninvasive techniques that examine language processing in young children (Figure 1). They include Electroencephalography (EEG)/Event-related Potentials (ERPs), Magnetoencephalography (MEG), functional Magnetic Resonance Imaging (fMRI), and Near-Infrared Spectroscopy (NIRS).

Event-related Potentials (ERPs) have been widely used to study speech and language processing in infants and young children (for reviews, see Conboy et al., 2008a; Friederici, 2005; Kuhl, 2004). ERPs, a part of the EEG, reflect electrical activity that is time-locked to the presentation of a specific sensory stimulus (for example, syllables or words) or a cognitive process (recognition of a semantic violation within a sentence or phrase). By placing sensors on a child's scalp, the activity of neural networks firing in a coordinated and synchronous fashion in open field configurations can be measured, and voltage changes occurring as a function of cortical neural activity can be detected. ERPs provide precise time resolution (milliseconds), making them well suited for studying the high-speed and temporally ordered structure of human speech. ERP experiments can also be carried out in populations who cannot provide overt responses because of age or cognitive impairment. Spatial resolution of the source of brain activation is, however, limited.

Magnetoencephalography (MEG) is another brain imaging technique that tracks activity in the brain with exquisite temporal resolution. The SQUID (superconducting quantum interference device) sensors located within the MEG helmet measure the minute magnetic fields associated with electrical currents that are produced by the brain when it is performing sensory, motor, or cognitive tasks. MEG allows precise localization of the neural currents responsible for the sources of the magnetic fields. Cheour et al. (2004) and Imada et al. (2006) used new head-tracking methods and MEG to show phonetic discrimination in

NEUROSCIENCE TECHNIQUES USED WITH INFANTS

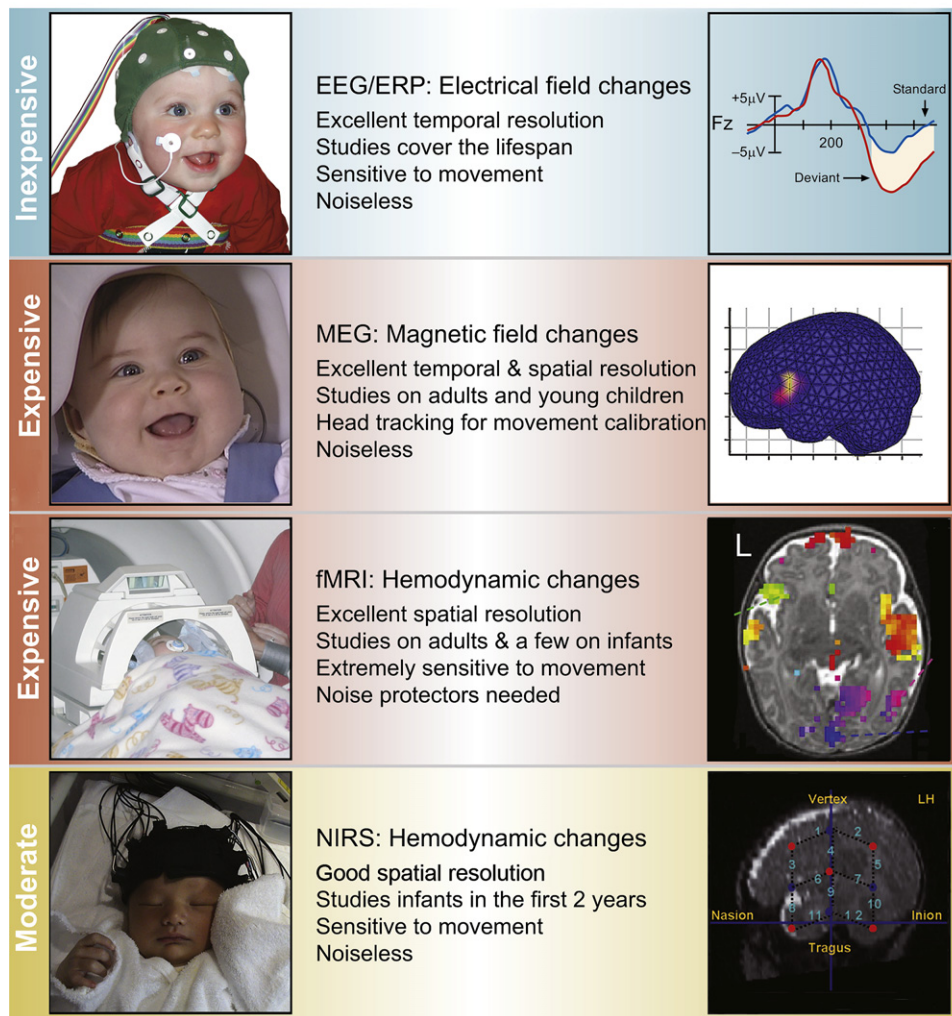


Figure 1. Four Techniques Now Used Extensively with Infants and Young Children to Examine Their Responses to Linguistic Signals (From Kuhl and Rivera-Gaxiola, 2008).

newborns and infants in the first year of life. Sophisticated head-tracking software and hardware enables investigators to correct for infants' head movements, and allows the examination of multiple brain areas as infants listen to speech (Imada et al., 2006). MEG (as well as EEG) techniques are completely safe and noiseless.

Magnetic resonance imaging (MRI) can be combined with MEG and/or EEG, providing static structural/anatomical pictures of the brain. Structural MRIs show anatomical differences in brain regions across the lifespan, and have recently been used to predict second-language phonetic learning in adults (Golestani and Pallier, 2007). Structural MRI measures in young infants identify the size of various brain structures and these measures have been shown to be related to language abilities later in childhood (Ortiz-Mantilla et al., 2010). When structural MRI images are superimposed on the physiological activity detected by MEG or EEG, the spatial localization of brain activities recorded by these methods can be improved.

Functional magnetic resonance imaging (fMRI) is a popular method of neuroimaging in adults because it provides high spatial-resolution maps of neural activity across the entire brain (e.g., Gernsbacher and Kaschak, 2003). Unlike EEG and MEG, fMRI does not directly detect neural activity, but rather the changes in blood-oxygenation that occur in response to neural activation. Neural events happen in milliseconds; however, the blood-oxygenation changes that they induce are spread out over several seconds, thereby severely limiting fMRI's temporal resolution. Few studies have attempted fMRI with infants because the technique requires infants to be perfectly still, and because the MRI device produces loud sounds making it necessary to shield infants' ears. fMRI studies allow precise localization of brain activity and a few pioneering studies show remarkable similarity in the structures responsive to language in infants and adults (Dehaene-Lambertz et al., 2002, 2006).

Near-Infrared Spectroscopy (NIRS) also measures cerebral hemodynamic responses in relation to neural activity, but utilizes

the absorption of light, which is sensitive to the concentration of hemoglobin, to measure activation (Aslin and Mehler, 2005). NIRS measures changes in blood oxy- and deoxy-hemoglobin concentrations in the brain as well as total blood volume changes in various regions of the cerebral cortex using near infrared light. The NIRS system can determine the activity in specific regions of the brain by continuously monitoring blood hemoglobin level. Reports have begun to appear on infants in the first two years of life, testing infant responses to phonemes as well as longer stretches of speech such as “motherese” and forward versus reversed sentences (Bortfeld et al., 2007; Homae et al., 2006; Peña et al., 2002; Taga and Asakawa, 2007). As with other hemodynamic techniques such as fMRI, NIRS typically does not provide good temporal resolution. However, event-related NIRS paradigms are being developed (Gratton and Fabiani, 2001). One of the most important potential uses of the NIRS technique is possible co-registration with other testing techniques such as EEG and MEG.

Neural Signatures of Early Learning

Perception of the phonetic units of speech—the vowels and consonants that make up words—is one of the most widely studied linguistic skills in infancy and adulthood. Phonetic perception and the role of experience in learning is studied in newborns, during development as infants are exposed to a particular language, in adults from different cultures, in children with developmental disabilities, and in nonhuman animals. Phonetic perception studies provide critical tests of theories of language development and its evolution. An extensive literature on developmental speech perception exists and brain measures are adding substantially to our knowledge of phonetic development and learning (see Kuhl, 2004; Kuhl et al., 2008; Werker and Curtin, 2005).

In the last decade, brain and behavioral studies indicate a very complex set of interacting brain systems in the initial acquisition of language, many of which appear to reflect adult language processing, even early in infancy (Dehaene-Lambertz et al., 2006). In adulthood, language is highly modularized, which accounts for the very specific patterns of language deficits and brain damage in adult patients following stroke (P.K.K. and A. Damasio, *Principles of Neuronal Science V* [McGraw Hill], in press, E.R. Kandel, J.H. Schwartz, T.M. Jessell, S. Siegelbaum, and J. Hudspeth, eds). Infants, however, must begin life with brain systems that allow them to acquire any and all languages to which they are exposed, and can acquire language as either an auditory-vocal or a visual-manual code, on roughly the same timetable (Petitto and Marentette, 1991). We are in a nascent stage of understanding the brain mechanisms underlying infants’ early flexibility with regard to the acquisition of language – their ability to acquire language by eye or by ear, and acquire one or multiple languages – and also the reduction in this initial flexibility that occurs with age, which dramatically decreases our capacity to acquire a new language as adults (Newport, 1990). The infant brain is exquisitely poised to “crack the speech code” in a way that the adult brain cannot. Uncovering why this is the case is a very interesting puzzle.

In this review I will also explore a current working hypothesis and its implications for brain development—that to crack the

speech code requires infants to combine a powerful set of domain-general computational and cognitive skills with their equally extraordinary social skills. Thus, the underlying brain systems must mutually influence one another during development. Experience with more than one language, for example, as in the case of people who are bilingual, is related to increases in particular cognitive skills, both in adults (Bialystok, 1991) and in children (Carlson and Meltzoff, 2008). Moreover, social interaction appears to be necessary for language acquisition, and an individual infant’s social behavior can be linked to their ability to learn new language material (Kuhl et al., 2003; B.T. Conboy et al., 2008, “Joint engagement with language tutors predicts learning of second-language phonetic stimuli,” presentation at the 16th International Conference on Infancy Studies, Vancouver).

Regarding the social effects, I have suggested that the social brain—in ways we have yet to understand—“gates” the computational mechanisms underlying learning in the domain of language (Kuhl, 2007). The assertion that social factors gate language learning 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 et al., 2009). Theories of social learning have traditionally emphasized the role of social factors in language acquisition (Bruner, 1983; Vygotsky, 1962; Tomasello, 2003a, 2003b). However, these models have emphasized the development of lexical understanding and the use of others’ communicative intentions to help understand the mapping between words and objects. The new data indicate that social interaction “gates” an even more basic aspect of language — learning of the elementary phonetic units of language — and this suggests a more fundamental connection between the brain mechanisms underlying human social understanding and the origins of language than has previously been hypothesized.

In the next decade, the methods of modern neuroscience will be used to explore how the integration of brain activity across specialized brain systems involved in linguistic, social, and cognitive analyses take place. These approaches, as well as others described here, will lead us toward a view of language acquisition in the human child that could be transformational.

The Learning Problem

Language learning is a deep puzzle that our theories and machines struggle to solve but children accomplish with ease. 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 a young child, merely one 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? A child’s budding ability to express a thought through words is a breathtaking feat of the human mind.

Research on infants’ phonetic perception in the first year of life shows how computational, cognitive, and social skills combine

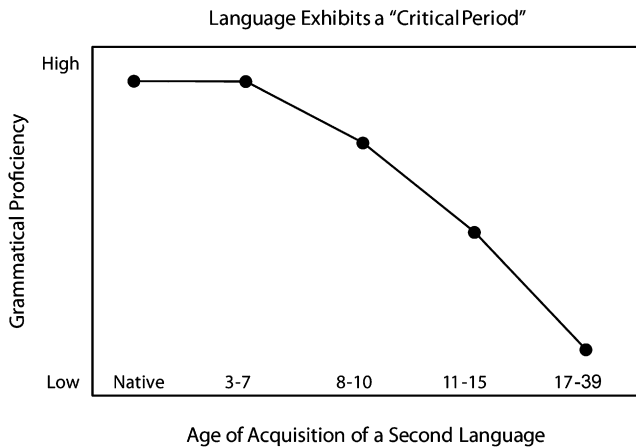


Figure 2. The Relationship between Age of Acquisition of a Second Language and Language Skill
Adapted from Johnson and Newport (1989).

to form a very powerful learning mechanism. Interestingly, this mechanism does not resemble Skinner's operant conditioning and reinforcement model of learning, nor Chomsky's detailed view of parameter setting. The learning processes that infants employ when learning from exposure to language are complex and multi-modal, but also child's play in that it grows out of infants' heightened attention to items and events in the natural world: the faces, actions, and voices of other people.

Language 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 native speakers of Korean learning English as a second language (1989). The graph as rendered shows a simplified schematic of second language competence as a function of the age of second language acquisition.

Figure 2 is surprising from the standpoint of more general human learning. 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 and Newport, 1989; Knudsen, 2004; Kuhl, 2004; Newport et al., 2001).

Scientists are generally in agreement that this learning curve is representative of data across a wide variety of second-language learning studies (Bialystok and Hakuta, 1994; Birdsong and Molis, 2001; Flege et al., 1999; Johnson and Newport, 1989; Kuhl et al., 2005a, 2008; Mayberry and Lock, 2003; Neville et al., 1997; Weber-Fox and Neville, 1999; Yeni-Komshian et al., 2000; though see Birdsong, 1992; White and Genesee, 1996). Moreover, not all aspects of language exhibit the same temporally defined critical "windows." The developmental timing of critical periods for learning phonetic, lexical, and syntactic levels of language vary, though studies cannot yet document the precise timing at each individual level. Studies indicate, for example, that the critical period for phonetic learning occurs prior to the end of the first year, whereas syntactic learning flour-

ishes between 18 and 36 months of age. Vocabulary development "explodes" at 18 months of age, but does not appear to be as restricted by age as other aspects of language learning—one can learn new vocabulary items at any age. One goal of future research will be to document the "opening" and "closing" of critical periods for all levels of language and understand how they overlap and why they differ.

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 affected language learning (Lenneberg, 1967; Newport et al., 2001). More recent hypotheses take a different perspective. Newport raised a "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 led me to advance the concept of *neural commitment*, the idea that neural circuitry and overall architecture develops early in infancy to detect the phonetic and prosodic patterns of speech (Kuhl, 2004; Zhang et al., 2005, 2009). This architecture is designed to maximize the efficiency of processing for the language(s) experienced by the infant. Once established, the neural architecture arising from French or Tagalog, for example, impedes learning of new patterns that do not conform. I will return to the concept of the critical period for language learning, and the role that computational, cognitive, and social skills may play in accounting for the relatively poor performance of adults attempting to learn a second language.

Focal Example: Phoneme 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* in English). But phonemes are actually groups of non-identical sounds, *phonetic units*, which are functionally equivalent in the language. 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 (*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, and have to derive the appropriate groupings used in their specific language. The baby's task in the first year of life, therefore, is to make some progress in figuring out the composition of the 40-odd phonemic categories in their language(s) before trying to acquire words that depend on these elementary units.

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 et al., 1992). Yet, by 10 months of age, differences can be

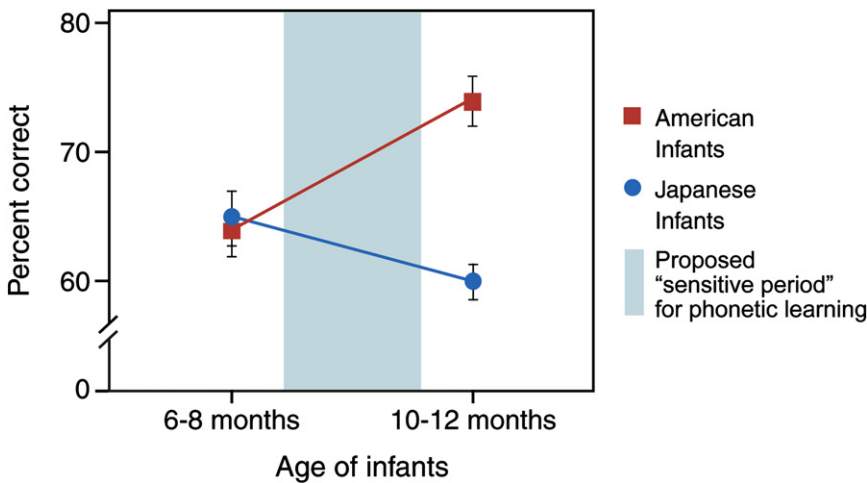


Figure 3. Effects of Age and Experience on Phonetic Discrimination

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 (adapted from Kuhl et al., 2006).

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 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 representations guide infants' successive motor approximations until a match is achieved (Kuhl and Meltzoff, 1996). This ability to imitate vocally may also depend on the brain's social understanding mechanisms which form a human mirroring system for seamless social interaction (Hari and Kujala, 2009), and we will revisit the impact of the brain's social understanding systems later in this review.

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), though 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). 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 in their language, and that these differences should be ignored? Similarly, English speakers produce Spanish *b* and *p*, exposing American infants to both categories of sound (Abramson and Lisker, 1970). How do American infants learn that these sounds do not distinguish words in English? An important discovery in the 1970s was that infants initially hear all these phonetic differences (Eimas, 1975; Eimas et al., 1971; Lasky et al., 1975; Werker and Lalonde, 1988). What we must explain is how infants learn to group phonetic units into phonemic categories that make a difference in their language.

The Timing of Phonetic Learning

Another important discovery in the 1980s identified the timing of a crucial change in infant perception. The transition from an early

universal perceptual ability to distinguish all the phonetic units of all languages to a more language specific pattern of perception occurred very early in development—between 6 and 12 months of age (Werker and Tees, 1984), and initial work demonstrated that infants' perception of nonnative distinctions declines during the second half of the first year of life (Best and McRoberts, 2003; Rivera-Gaxiola et al., 2005; Tsao et al., 2006; Werker and Tees, 1984). Work in this laboratory also established a new fact: At the same time that nonnative perception declines, native language speech perception shows a significant increase. Japanese infants' discrimination of English *r-l* declines between 8 and 10 months of age, while at the same time in development, American infants' discrimination of the same sounds shows an increase (Kuhl et al., 2006) (Figure 3).

Phonetic Learning Predicts the Rate of Language Growth

We argued that the increase observed in native-language phonetic perception represented a critical step in initial language learning and promoted language growth (Kuhl et al., 2006). To test this hypothesis, we designed a longitudinal study examining whether a measure of phonetic perception predicted children's language skills measured 18 months later. The study demonstrated that infants' phonetic discrimination ability at 6 months of age was significantly correlated with their success in language learning at 13, 16, and 24 months of age (Tsao et al., 2004). However, we recognized that in this initial study the association we observed might be due to infants' cognitive skills, such as the ability to perform in the behavioral task, or to sensory abilities that affected auditory resolution of the differences in formant frequencies that underlie phonetic distinctions.

To address these issues, we assessed both native and nonnative phonetic discrimination in 7-month-old infants, and used both a behavioral (Kuhl et al., 2005a) and an event-related potential measure, the mismatch negativity (MMN), to assess infants' performance (Kuhl et al., 2008). Using a neural measure removed potential cognitive effects on performance; the use of both native and nonnative contrasts addressed the sensory issue, since better sensory abilities would be expected to improve both native and nonnative speech discrimination.

The native language neural commitment (NLNC) view suggested that future language measures would be associated with early performance on both native and nonnative contrasts, but in opposite directions. The results conformed to this prediction. When both native and nonnative phonetic discrimination

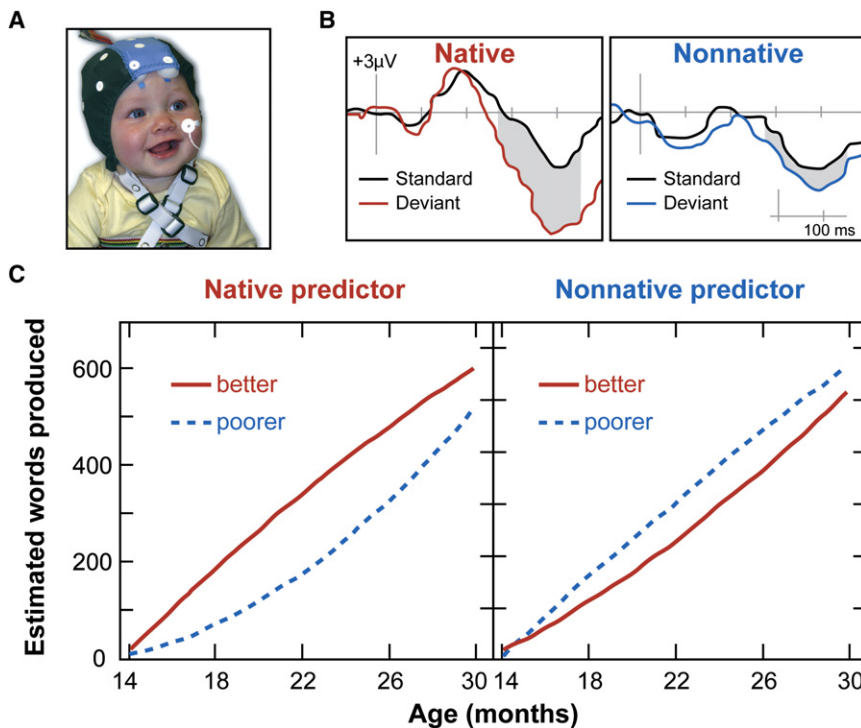


Figure 4. Speech Discrimination Predicts Vocabulary Growth

(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 +1 SD and -1 SD on the native contrast at 7.5 months (C, left) and vocabulary growth for MMN values of +1 SD and -1 SD on the nonnative contrast at 7.5 months (C, right) (adapted from Kuhl et al., 2008).

was measured in the same infants at 7.5 months of age, better *native* language perception predicted significantly higher language abilities between 18 and 30 months of age, whereas better *nonnative* phonetic perception at the same age predicted poorer language abilities at the same future points in time (Kuhl et al., 2005a, 2008). As shown in Figure 4, the ERP measure at 7.5 months of age (Figure 4A) provided an MMN measure of speech discrimination for both native and nonnative contrasts; greater negativity of the MMN reflects greater discrimination (Figure 4B). Hierarchical linear growth modeling of vocabulary between 14 and 30 months for MMN values of +1SD and -1SD (Figure 4C) revealed that both native and nonnative phonetic discrimination significantly predict future language, but in opposite directions with better native MMNs predicting advanced future language development and better nonnative MMNs predicting less advanced future language development.

The results are explained by NLNC: better native phonetic discrimination enhances infants' skills in detecting words and this vaults them toward language, whereas better nonnative abilities indicated that infants remained at an earlier phase of development – sensitive to all phonetic differences. Infants' ability to *learn* which phonetic units are relevant in the language(s) they are exposed to, while decreasing or inhibiting their attention to the phonetic units that do not distinguish words in their language, is the necessary step required to begin the path toward language. These data led to a theoretical argument that an implicit learning process 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).

Recent data indicate very long-term associations between infants' phonetic perception and future language and reading skills. Our studies show that the ability to discriminate two simple vowels

at 6 months of age predicts language abilities and pre-reading skills such as rhyming at the age of 5 years, an association that holds regardless of socio-economic status and the children's language skills at 2.5 years of age (Cardillo, 2010).

A Computational Solution to Phonetic Learning

A surprising new form of learning, referred to as “statistical learning” (Saffran et al., 1996), was discovered in the 1990s. 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 strongly influences both phonetic learning and early word learning.

For example, data show that the developmental change in phonetic perception between the ages of 6 and 12 months is supported by infants' sensitivity to the distributional frequencies of the sounds in the language(s) they hear, and that this affects perception. To illustrate, adult speakers of English and Japanese produce both English r- and l-like sounds, even though English speakers hear /r/ and /l/ as distinct and Japanese adults hear them as identical. Japanese infants are therefore exposed to both /r/ and /l/ sounds, even though they do not represent distinct categories in Japanese. The presence of a particular sound in ambient language, therefore, does not account for infant learning. However, distributional frequency analyses of English and Japanese show differential patterns of distributional frequency; in English, /r/ and /l/ occur very frequently; in Japanese, the most frequent sound of this type is Japanese /r/ which is related to but distinct from both the English variants. Can infants learn from this kind of distributional information in speech input?

A variety of studies show that infants' perception of phonetic categories is affected by distributional patterns in the sounds they hear. In one study using very simple stimuli and short-term exposure in the laboratory, 6- and 8-month-old infants were exposed for 2 min to 8 sounds that formed a continuum of sounds from /da/ to /ta/ (Maye et al., 2002; see also Maye et al., 2008). All infants heard all the stimuli on the continuum, but experienced different distributional frequencies of the sounds. 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 /da/ and /ta/ sounds, whereas those in the unimodal group did not. Furthermore, while previous studies show that infants integrate the auditory and visual instantiations of speech (Kuhl and Meltzoff, 1982; Patterson and Werker, 1999), more recent studies show that infants' detection of statistical patterns in speech stimuli, like those used by Maye and her colleagues, is influenced both by the auditory event and the sight of a face articulating the sounds. When exposed only to the ambiguous auditory stimuli in the middle of a speech continuum, infants discriminated the /da-ta/ contrast when each auditory stimulus was paired with the appropriate face articulating either /da/ or /ta/; discrimination did not occur if only one face was used with all auditory stimuli (Teinonen et al., 2008).

Cross-cultural studies also indicate that infants are sensitive to the statistical distribution of sounds they hear in natural language. Infants tested in Sweden and the United States at 6 months of age showed a unique response to vowel sounds that represent the distributional mean in productions of adults who speak the language (i.e., "prototypes"); this response was shown only for stimuli infants had been exposed to in natural language (native-vowel prototypes), not foreign-language vowel prototypes (Kuhl et al., 1992). Taken as a whole, these studies indicate infants pick up the distributional frequency patterns in ambient speech, whether they experience them during short-term laboratory experiments, or over months in natural environments, and can learn from them.

Statistical learning also supports word learning. 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 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 "tty," and between "ba" and "by," are higher than those between syllables that cross word boundaries, such as "tty" and "ba." 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). Recent findings show that even sleeping newborns detect this kind of statistical structure in speech, as shown in studies using event-related brain potentials (Teinonen et al., 2009). Statistical learning has been shown in nonhuman animals (Hauser et al., 2001), and in

humans for stimuli outside the realm of speech, operating for musical and visual patterns in the same way as speech (Fiser and Aslin, 2002; Kirkham et al., 2002; Saffran et al., 1999). Thus, a very basic implicit learning mechanism allows infants, from birth, to detect statistical structure in speech and in other signals. Infants' sensitivity to this statistical structure can influence both phoneme and word learning.

Effects of Social Interaction on Computational Learning

As reviewed, infants show robust learning effects in statistical learning studies when tested in the laboratory with very simple stimuli (Maye et al., 2002, 2008; Saffran et al., 1996). However, complex natural language learning may challenge infants in a way that these experiments do not. Are there constraints on statistical learning as an explanation for natural language learning? A series of later studies suggest that this is the case. Laboratory studies testing infant phonetic and word learning from exposure to a complex natural language suggest limits on statistical learning, and provide new information suggesting that social brain systems are integrally involved, and, in fact, may be necessary to explain natural language learning.

The new experiments tested infants in the following way: At 9 months of age, the age at which the initial universal pattern of infant perception has changed to one that is more language-specific, infants were exposed to a foreign language for the first time (Kuhl et al., 2003). Nine-month-old American infants listened to 4 different native speakers of Mandarin during 12 sessions scheduled over 4–5 weeks. 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-person" sessions – after exposure, 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 been listening to Mandarin for 10 months (Kuhl et al., 2003).

The study revealed that infants can learn from first-time natural exposure to a foreign language at 9 months, and answered what was initially the experimental question: can infants learn the statistical structure of phonemes in a new language given first-time exposure at 9 months of age? If infants required a long-term history of listening to that language—as would be the case if infants needed to build up statistical distributions over the initial 9 months of life—the answer to our question would have been no. However, the data clearly showed that infants are capable of learning at 9 months when exposed to a new language. Moreover, learning was durable. Infants returned to the laboratory for their behavioral discrimination tests between 2 and 12 days after the final language exposure session, and between 8 and 33 days for their ERP measurements. No "forgetting" of the Mandarin contrast occurred during the 2 to 33 day delay.

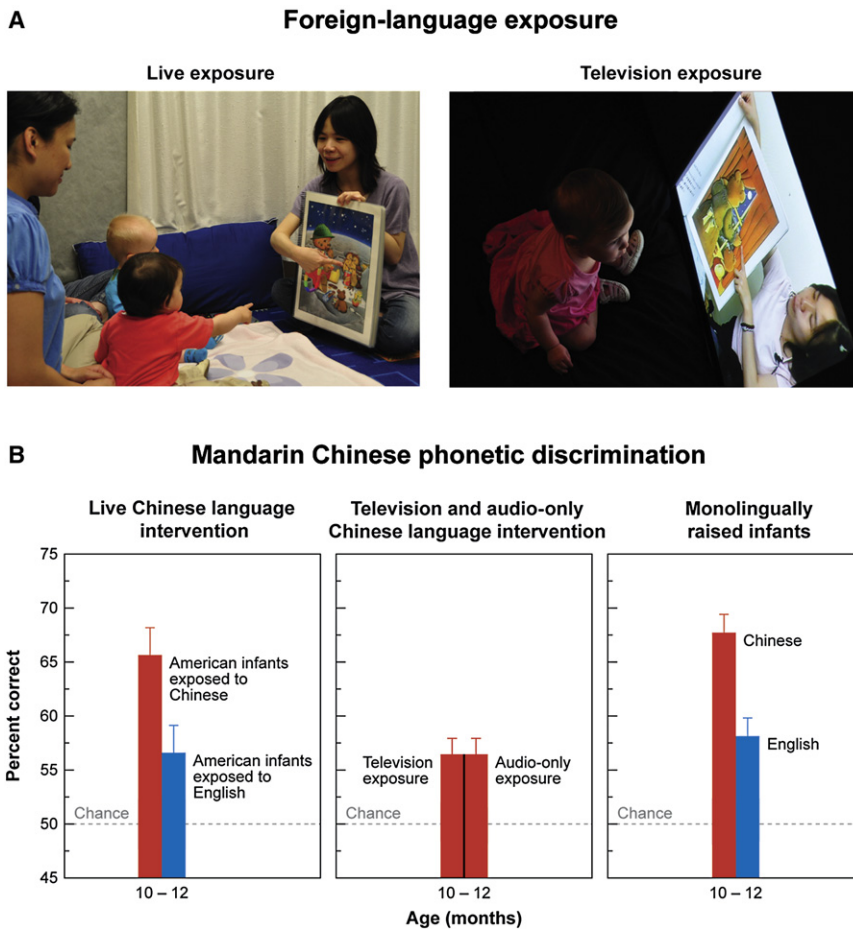


Figure 5. Social Interaction Facilitates Foreign Language Learning

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 (data 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) (adapted from Kuhl et al., 2003).

We were struck by the fact that infants exposed to Mandarin were socially very engaged in the language sessions and began to wonder about the role of social interaction in learning. Would infants learn if they were exposed to the same information in the absence of a human being, say, via television or an audiotape? If statistical learning is sufficient, the television and audio-only conditions should produce learning. Infants who were exposed to the same foreign-language material at the same time and at the same rate, but via standard television or audiotape only, showed no learning—their performance equaled that of infants in the control group who had not been exposed to Mandarin at all (Figure 5).

Thus, the presence of a human being interacting with the infant during language exposure, while not required for simpler statistical-learning tasks (Maye et al., 2002; Saffran et al., 1996), is critical for learning in complex natural language-learning situations in which infants heard an average of 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 impact of social interaction on language learning (Kuhl et al., 2003) led to the development of the *Social Gating Hypothesis*

(Kuhl, 2007). “Gating” suggested that social interaction creates a vastly different learning situation, one in which additional factors introduced by a social context influence learning. Gating could operate by increasing: (1) attention and/or arousal, (2) information, (3) a sense of relationship, and/or (4) activation of brain mechanisms linking perception and action.

Attention and arousal affect learning in a wide variety of domains (Posner, 2004), and could impact infant learning during exposure to a new language. Infant attention, measured in the original 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, measured as the amount of infant “looking time,” learned from bimodal stimulus distributions when “low-attenders” did not (Yoshida et al., 2006; see also Yoshida et al., 2010). Heightened attention and arousal could produce an overall increase in the quantity or quality of the speech information that infants encode and remember. Recent data suggest a role for attention in adult second-language phonetic learning as well (Guion and Pederson, 2007).

A second hypothesis was raised to explain the effectiveness of social interaction – the live learning situation allowed the infants and tutors to interact, and this added contingent and reciprocal social behaviors that increased information that could foster learning. During live exposure, tutors focused their visual gaze on pictures in the books or on the toys as they spoke, and the infants’ gaze tended to follow the speaker’s gaze, as previously observed in social learning studies (Baldwin, 1995; Brooks and Meltzoff, 2002). 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 and Meltzoff, 2005; Mundy and Gomes, 1998), and may help infants link the foreign speech

to the objects they see. 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 word-like units from ongoing speech, thus facilitating phonetic learning of the sounds contained in those words.

If this hypothesis is correct, then the degree to which infants interact and engage socially with the tutor in the social language-learning situation should correlate with learning. In studies testing this hypothesis, 9-month-old infants were exposed to Spanish (Conboy and Kuhl, 2010), extending the experiment to a new language. Other changes in method expanded the tests of language learning to include both Spanish phonetic learning and Spanish word learning, as well as adding measures of specific interactions between the tutor and the infant to examine whether interactive episodes could be related to learning of either phonemes or words.

The results confirmed Spanish language learning, both of the phonetic units of the language and the lexical units of the language (Conboy and Kuhl, 2010). 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 ERP measures of Spanish phoneme discrimination? Our results (Figure 6) show that they do (Conboy et al., 2008a). Infants who shifted their gaze between the tutor's eyes and newly introduced toys during the Spanish exposure sessions showed a more negative MMN (indicating greater neural discrimination) in response to the Spanish phonetic contrast. Infants who simply gazed at the tutor or at the toy, showing fewer gaze shifts, produced less negative MMN responses. The degree of infants' social engagement during sessions predicted both phonetic and word learning—infants who were more socially engaged showed greater learning as reflected by ERP brain measures of both phonetic and word learning.

Language, Cognition, and Bilingual Language Experience

Specific cognitive abilities, particularly the executive control of attention and the ability to inhibit a pre-potent response (inhibitory control), are associated with exposure to more than one language. Bilingual adult speakers show enhanced executive control skills (Bialystok, 1999, 2001; Bialystok and Hakuta, 1994; Wang et al., 2009), a finding that has been extended to young school-aged bilingual children (Carlson and Meltzoff, 2008). In monolingual infants, the decline in discrimination of nonnative contrasts (which promotes more rapid growth in language, see Figure 4C) is associated with enhanced inhibitory control, suggesting that domain-general cognitive mechanisms underlying attention may play a role in enhancing performance on native and suppressing performance on nonnative phonetic contrasts early in development (Conboy et al., 2008b; Kuhl et al., 2008). In support of this view, it is noteworthy that in the Spanish exposure studies, a median split of the post-exposure MMN phonetic discrimination data revealed that infants showing greater phonetic learning had higher cognitive control scores post-exposure. These same infants did not differ in their pre-exposure cognitive control tests (Conboy, Sommerville, and

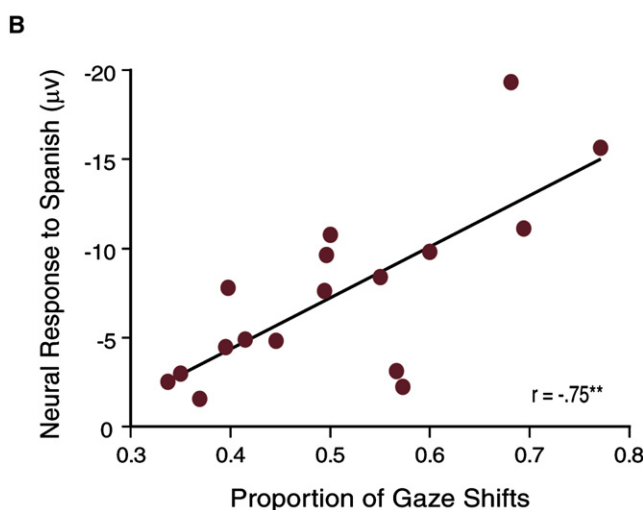


Figure 6. Social Engagement Predicts Foreign Language Learning
(A) Nine-month-old infants experienced 12 sessions of Spanish through natural interaction with a Spanish speaker.
(B) The neural response to the Spanish phonetic contrast (d-t) and the proportion of gaze shifts during Spanish sessions were significantly correlated (from Conboy et al., unpublished data).

P.K.K., unpublished data). Taken as a whole, the data are consistent with the notion that cognitive skills are strongly linked to phonetic learning at the initial stage of phonetic development (Kuhl et al., 2008).

The “Social Brain” and Language Learning Mechanisms

While attention and the information provided by interaction with another may help explain social learning effects for language, it is also possible that social contexts are connected to language learning through even more fundamental mechanisms. Social interaction may activate brain mechanisms that invoke a sense of relationship between the self and other, as well as social understanding systems 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 research (Liberman and Mattingly, 1985), and interest in “mirror systems” for social cognition have re-invigorated this

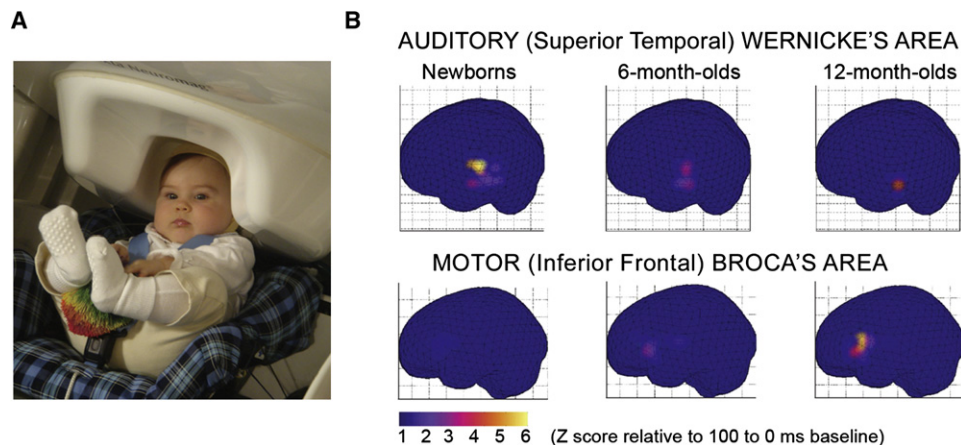


Figure 7. Perception-Action Brain Systems Respond to Speech in Infancy

(A) Neuromagnetic signals were recorded in newborns, 6-month-old infants (shown), and 12-month-old infants in the MEG machine while listening to speech and nonspeech auditory signals.

(B) Brain activation in response to speech recorded in auditory (B, top row) and motor (B, 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 in 6- and 12-month-old infants (from Imada et al., 2006).

tradition (Kuhl and Meltzoff, 1996; Meltzoff and Decety, 2003; Pulvermuller, 2005; Rizzolatti, 2005; Rizzolatti and Craighero, 2004). Might the brain systems that link perception and production for speech be engaged when infants experience social interaction during language learning?

The effects of Spanish language exposure extend to speech production, and provide evidence of an early coupling of sensory-motor learning in speech. The English-learning infants who were exposed to 12 sessions of Spanish (Conboy and Kuhl, 2010) showed subsequent changes in their patterns of vocalization (N. Ward et al., 2009, "Consequences of short-term language exposure in infancy on babbling," poster presented at the 158th meeting of the Acoustical Society of America, San Antonio). When presented with language from a Spanish speaker (but not from an English speaker), a new pattern of infant vocalizations was evoked, one that reflected the prosodic patterns of Spanish, rather than English. This only occurred in response to Spanish, and only occurred in infants who had been exposed to Spanish in the laboratory experiment.

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 toward 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 7). Dehaene-Lambertz and colleagues (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. (2006) 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 begin to produce vowel-like sounds.

Using the tools of modern neuroscience, we can now ask how the brain systems responsible for speech perception and speech production forge links early in development, and whether 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.

Brain Rhythms, Cognitive Effects, and Language Learning

MEG studies will provide an opportunity to examine brain rhythms associated with broader cognitive abilities during speech learning. Brain oscillations in various frequency bands have been associated with cognitive abilities. The induced brain rhythms have been linked to attention and cognitive effort, and are of primary interest since MEG studies with adults have shown that cognitive effort is increased when processing nonnative speech (Zhang et al., 2005, 2009). In the adult MEG studies, participants listened to their native- and to nonnative-language sounds. The results indicated that when listening to native language, the brain's activation was more focal, and faster, than when listening to nonnative-language sounds (Zhang et al., 2005). In other words, there was greater *neural efficiency* for native as opposed to nonnative speech processing. Training studies show that adults can improve nonnative phonetic perception when training occurs under more social learning conditions, and MEG measures before and after training indicate that neural efficiency increases after training (Zhang et al., 2009). Similar patterns of neural inefficiency occur as young children learn words. Young children's event-related brain potential

responses are more diffuse and become more focally lateralized in the left hemisphere's temporal regions as they develop (Conboy et al., 2008a; Durston et al., 2002; Mills et al., 1993, 1997; Tamm et al., 2002) and studies with young children with autism show this same pattern – more diffuse activation – when compared to typically developing children of the same age (Coffey-Corina et al., 2008).

Brain rhythms may be reflective of these same processes in infants as they learn language. Brain oscillations in four frequency bands have been associated with cognitive effects: *theta* (4–7 Hz), *alpha* (8–12 Hz), *beta* (13–30 Hz) and *gamma* (30–100 Hz). Resting gamma has been related to early language and cognitive skills in the first three years (Benasich et al., 2008). The induced theta rhythm has been linked to attention and cognitive effort, and will be of strong interest to speech researchers. Power in the theta band increases with memory load in adults tested in either verbal or nonverbal tasks (Gevins et al., 1997; Krause et al., 2000) and in 8-month-old infants tested in working memory tasks (Bell and Wolfe, 2007). Examining brain rhythms in infants using speech stimuli is now underway using EEG with high-risk infants (C.R. Percaccio et al., 2010, “Native and nonnative speech-evoked responses in high-risk infant siblings,” abstracts of the International Meeting for Autism Research, May 2010, Philadelphia) and using MEG with typically developing infants (A.N. Bosseler et al., 2010, “Event-related fields and cortical rhythms to native and nonnative phonetic contrasts in infants and adults,” abstracts of the 17th International Conference of Biomagnetism), as they listen to native and nonnative speech. Comparisons between native and nonnative speech may allow us to examine whether there is increased cognitive effort associated with processing nonnative language, across age and populations. We are also testing whether language presented in a social environment affects brain rhythms in a way that television and audiotape presentations do not. Neural efficiency is not observable with behavioral approaches—and one promise of brain rhythms is that they provide the opportunity to compare the higher-level processes that likely underlie humans' neural plasticity for language early in development in typical children as well as in children at risk for autism spectrum disorder, and in adults learning a second language. These kinds of studies may reveal the cortical dynamics underlying the “Critical Period” for language.

These results underscore the importance of a social interest in speech early in development in both typical and atypical populations. An interest in “motherese,” the universal style with which adults address infants across cultures (Fernald and Simon, 1984; Grieser and Kuhl, 1988) provides a good metric of the value of a social interest in speech. The acoustic stretching in motherese, observed across languages, makes phonetic units more distinct from one another (Burnham et al., 2002; Englund, 2005; Kuhl et al., 1997; Liu et al., 2003, 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 show significantly better performance in phonetic discrimination tasks when tested in the laboratory (Liu et al., 2003). New data show that the potential benefits of early motherese extend to the age of 5 years (Liu et al., 2009). Recent ERP studies indicate that infants' brain responses to the exag-

gerated patterns of motherese elicit an enhanced N250 as well as increased neural synchronization at frontal-central-parietal sites (Zhang et al., personal communication).

It is also noteworthy that children with Autism Spectrum Disorder (ASD) prefer to listen to non-speech rather than speech, when given a choice, and this preference is strongly correlated with the children's ERP brain responses to speech, as well as with the severity of their autistic symptoms (Kuhl et al., 2005b). Early speech measures may therefore provide an early biomarker of risk for ASD. Neuroscience studies in both typically developing and children with ASD that examine the coherence and causality of interaction between social and linguistic brain systems will provide valuable new theoretical data as well as potentially improving the early diagnosis and treatment of children with autism.

Neurobiological Foundations of Communicative Learning

Humans are not the only species in which communicative learning is affected by social interaction (see Fitch et al., 2010, for review). 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). More recent data indicate that male zebra finches vary their songs across social contexts; songs produced when singing to females vary from those produced in isolation, and females prefer these “directed” songs (Woolley and Doupe, 2008). Moreover, gene expression in high-level auditory areas is involved in this kind of social context perception (Woolley and Doupe, 2008). White-crowned sparrows, which reject the audiotaped songs of alien species, learn the same alien songs when a live tutor sings them (Baptista and Petrinovich, 1986). In barn owls (Brainard and Knudsen, 1998) and white-crowned sparrows (Baptista and 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 and King, 1988), and white-crowned sparrow tutors provide acoustic feedback that affects the repertoires of young birds (Nelson and Marler, 1994). Studies of the brain systems linking social and auditory-vocal learning in humans and birds may significantly advance theories in the near future (Doupe and Kuhl, 2008).

Neural Underpinnings of Cognitive and Social Influences on Language Learning

Our current model of neural commitment to language describes a significant role for cognitive processes such as attention in language learning (Kuhl et al., 2008). Studies of brain rhythms in infants and other neuroscience research in the next decade promise to reveal the intricate relationships between language and cognitive processes.

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; Kuhl, 2007; Pulvermuller, 2005). Social interaction appears to

be necessary for language learning in infants (Kuhl et al., 2003), and an individual infant's social behavior is linked to their ability to learn new language material (Conboy et al., 2008a). In fact, social "gating" may explain why social factors play a far more significant role than previously realized in human learning across domains throughout our lifetimes (Meltzoff et al., 2009). If social factors "gate" computational learning, as proposed, infants would be protected from meaningless calculations – learning would be restricted to signals that derive from live humans rather than other sources (Doupe and Kuhl, 2008; Evans and Marler, 1995; Marler, 1991). Constraints of this kind appear to exist for infant imitation: when infants hear nonspeech sounds with the same frequency components as speech, they do not attempt to imitate them (Kuhl et al., 1991).

Research has begun to appear on the development of the neural networks in humans that constitute the "social brain" and invoke a sense of relationship between the self and other, as well as on social understanding systems that link perception and action (Hari and Kujala, 2009). Neuroscience studies using speech and imaging techniques are beginning to examine links between sensory and motor brain systems (Pulvermüller, 2005; Rizzolatti and Craighero, 2004), and 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) provides exciting opportunities. MEG studies of brain activation in infants during social versus nonsocial language experience will allow us to investigate cognitive effects via brain rhythms and also examine whether social brain networks are activated differentially under the two conditions.

Many questions remain about the impact of cognitive skills and social interaction on natural speech and language learning. As reviewed, new data show the extensive interface between cognition and language and indicate that whether or not multiple languages are experienced in infancy affects cognitive brain systems. The idea that social interaction is integral to language learning has been raised previously for word learning; however, previous data and theorizing have not tied early phonetic learning to social factors. Doing so suggests a more fundamental connection between the motivation to learn socially and the mechanisms that enable language learning.

Understanding how language learning, cognition, and social processing interact in development may ultimately explain the mechanisms underlying the critical period for language learning. Furthermore, understanding the mechanism underlying the critical period may help us develop methods that more effectively teach second languages to adult learners. Neuroscience studies over the next decade will lead the way on this theoretical work, and also advance our understanding of the practical results of training methods, both for adults learning new languages, and children with developmental disabilities struggling to learn their first language. These advances will promote the science of learning in the domain of language, and potentially, shed light on human learning mechanisms more generally.

ACKNOWLEDGMENTS

The author and research reported here were supported by a grant from the National Science Foundation's Science of Learning Program to the University

of Washington LIFE Center (SBE-0354453), and by grants from the National Institutes of Health (HD37954, HD55782, HD02274, DC04661).

REFERENCES

- Abramson, A.S., and Lisker, L. (1970). Discriminability along the voicing continuum: cross-language tests. *Proc. Int. Congr. Phon. Sci.* 6, 569–573.
- Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nat. Rev. Neurosci.* 4, 165–178.
- Aslin, R.N., and Mehler, J. (2005). Near-infrared spectroscopy for functional studies of brain activity in human infants: promise, prospects, and challenges. *J. Biomed. Opt.* 10, 11009.
- Baldwin, D.A. (1995). Understanding the link between joint attention and language. In *Joint Attention: Its Origins and Role in Development*, C. Moore and P.J. Dunham, eds. (Hillsdale, NJ: Lawrence Erlbaum Associates), pp. 131–158.
- Baptista, L.F., and Petrinovich, L. (1986). Song development in the white-crowned sparrow: social factors and sex differences. *Anim. Behav.* 34, 1359–1371.
- Bell, M.A., and Wolfe, C.D. (2007). Changes in brain functioning from infancy to early childhood: evidence from EEG power and coherence during working memory tasks. *Dev. Neuropsychol.* 31, 21–38.
- Benasich, A.A., Gou, Z., Choudhury, N., and Harris, K.D. (2008). Early cognitive and language skills are linked to resting frontal gamma power across the first 3 years. *Behav. Brain Res.* 195, 215–222.
- Best, C.C., and McRoberts, G.W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Lang. Speech* 46, 183–216.
- Bialystok, E. (1991). *Language Processing in Bilingual Children* (Cambridge University Press).
- Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Dev.* 70, 636–644.
- Bialystok, E. (2001). *Bilingualism in Development: Language, Literacy, and Cognition* (New York: Cambridge University Press).
- Bialystok, E., and Hakuta, K. (1994). *Other Words: The Science and Psychology of Second-Language Acquisition* (New York: Basic Books).
- Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Ling. Soc. Am.* 68, 706–755.
- Birdsong, D., and Molis, M. (2001). On the evidence for maturational constraints in second-language acquisitions. *J. Mem. Lang.* 44, 235–249.
- Bortfeld, H., Wruck, E., and Boas, D.A. (2007). Assessing infants' cortical response to speech using near-infrared spectroscopy. *Neuroimage* 34, 407–415.
- Brainard, M.S., and Knudsen, E.I. (1998). Sensitive periods for visual calibration of the auditory space map in the barn owl optic tectum. *J. Neurosci.* 18, 3929–3942.
- Brooks, R., and Meltzoff, A.N. (2002). The importance of eyes: how infants interpret adult looking behavior. *Dev. Psychol.* 38, 958–966.
- Brooks, R., and Meltzoff, A.N. (2005). The development of gaze following and its relation to language. *Dev. Sci.* 8, 535–543.
- Bruer, J.T. (2008). Critical periods in second language learning: distinguishing phenomena from explanation. In *Brain, Behavior and Learning in Language and Reading Disorders*, M. Mody and E. Silliman, eds. (New York: The Guilford Press), pp. 72–96.
- Bruner, J. (1983). *Child's Talk: Learning to Use Language* (New York: W.W. Norton).
- Burnham, D., Kitamura, C., and Vollmer-Conna, U. (2002). What's new, pussycat? On talking to babies and animals. *Science* 296, 1435.
- Cardillo, G.C. (2010). Predicting the predictors: Individual differences in longitudinal relationships between infant phoneme perception, toddler vocabulary,

and preschooler language and phonological awareness. Doctoral Dissertation, University of Washington.

Carlson, S.M., and Meltzoff, A.N. (2008). Bilingual experience and executive functioning in young children. *Dev. Sci.* 11, 282–298.

Cheour, M., Imada, T., Taulu, S., Ahonen, A., Salonen, J., and Kuhl, P.K. (2004). Magnetoencephalography is feasible for infant assessment of auditory discrimination. *Exp. Neurol.* 190 (Suppl 1), S44–S51.

Chomsky, N. (1959). Review of Skinner's Verbal Behavior. *Language* 35, 26–58.

Coffey-Corina, S., Padden, D., Kuhl, P.K., and Dawson, G. (2008). ERPs to words correlate with behavioral measures in children with Autism Spectrum Disorder. *J. Acoust. Soc. Am.* 123, 3742–3748.

Conboy, B.T., and Kuhl, P.K. (2010). Impact of second-language experience in infancy: brain measures of first- and second-language speech perception. *Dev. Sci.*, in press. Published online June 28, 2010. 10.1111/j.1467-7687.2010.00973.x.

Conboy, B.T., Rivera-Gaxiola, M., Silva-Pereyra, J., and Kuhl, P.K. (2008a). Event-related potential studies of early language processing at the phoneme, word, and sentence levels. In *Early Language Development: Volume 5. Bridging Brain and Behavior, Trends in Language Acquisition Research*, A.D. Friederici and G. Thierry, eds. (Amsterdam, The Netherlands: John Benjamins), pp. 23–64.

Conboy, B.T., Sommerville, J.A., and Kuhl, P.K. (2008b). Cognitive control factors in speech perception at 11 months. *Dev. Psychol.* 44, 1505–1512.

de Boer, B., and Kuhl, P.K. (2003). Investigating the role of infant-directed speech with a computer model. *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*, B. de Boysson-Bardies, S. de Schonen, P. Juszczyk, P. McNeillage, and J. Morton, eds. (Dordrecht, Netherlands: Kluwer), pp. 353–363.

Dehaene-Lambertz, G., Dehaene, S., and Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science* 298, 2013–2015.

Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Mériaux, S., Roche, A., Sigman, M., and Dehaene, S. (2006). Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proc. Natl. Acad. Sci. USA* 103, 14240–14245.

Doupe, A.J., and Kuhl, P.K. (2008). Birdsong and human speech: common themes and mechanisms. In *Neuroscience of Birdsong*, H.P. Zeigler and P. Marler, eds. (Cambridge, England: Cambridge University Press), pp. 5–31.

Dunbar, R.I.M. (1998). The social brain hypothesis. *Evol. Anthropol.* 6, 178–190.

Durston, S., Thomas, K.M., Worden, M.S., Yang, Y., and Casey, B.J. (2002). The effect of preceding context on inhibition: an event-related fMRI study. *Neuroimage* 16, 449–453.

Eales, L.A. (1989). The influence of visual and vocal interaction on song learning in zebra finches. *Anim. Behav.* 37, 507–508.

Eimas, P.D. (1975). Auditory and phonetic coding of the cues for speech: discrimination of the /r-/ distinction by young infants. *Percept. Psychophys.* 18, 341–347.

Eimas, P.D., Siqueland, E.R., Juszczyk, P., and 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 Lang.* 25, 219–234.

Evans, C.S., and Marler, P. (1995). Language and animal communication: parallels and contrasts. In *Comparative Approaches to Cognitive Science: Complex Adaptive Systems*, H.L. Roitblat and J.-A. Meyer, eds. (Cambridge, MA: MIT Press), pp. 341–382.

Ferguson C.A., Menn L., and Stoel-Gammon C., eds. (1992). *Phonological Development: Models, Research, Implications* (Timonium, MD: York Press).

Fernald, A., and Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Dev. Psychol.* 20, 104–113.

Fiser, J., and Aslin, R.N. (2002). Statistical learning of new visual feature combinations by infants. *Proc. Natl. Acad. Sci. USA* 99, 15822–15826.

Fitch, W.T., Huber, L., and Bugnyar, T. (2010). Social cognition and the evolution of language: constructing cognitive phylogenies. *Neuron* 65, 795–814.

Flege, J.E. (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, J.E., Yeni-Komshian, G.H., and Liu, S. (1999). Age constraints on second-language acquisition. *J. Mem. Lang.* 41, 78–104.

Friederici, A.D. (2005). Neurophysiological markers of early language acquisition: from syllables to sentences. *Trends Cogn. Sci.* 9, 481–488.

Friederici, A.D., and Wartenburger, I. (2010). Language and Brain. *Wiley Interdisciplinary Reviews: Cognitive Science* 1, 150–159.

Gernsbacher, M.A., and Kaschak, M.P. (2003). Neuroimaging studies of language production and comprehension. *Annu. Rev. Psychol.* 54, 91–114.

Gevens, A., Smith, M.E., McEvoy, L., and Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cereb. Cortex* 7, 374–385.

Golestani, N., and Pallier, C. (2007). Anatomical correlates of foreign speech sound production. *Cereb. Cortex* 17, 929–934.

Gratton, G., and Fabiani, M. (2001). Shedding light on brain function: the event-related optical signal. *Trends Cogn. Sci.* 5, 357–363.

Grieser, D.L., and Kuhl, P.K. (1988). Maternal speech to infants in a tonal language: support for universal prosodic features in motherese. *Dev. Psychol.* 24, 14–20.

Guion, S.G., and Pederson, E. (2007). Investigating the role of attention in phonetic learning. In *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, O.-S. Bohn and M. Munro, eds. (Amsterdam: John Benjamins), pp. 55–77.

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

Hauser, M.D., Newport, E.L., and Aslin, R.N. (2001). Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition* 78, B53–B64.

Homae, F., Watanabe, H., Nakano, T., Asakawa, K., and Taga, G. (2006). The right hemisphere of sleeping infant perceives sentential prosody. *Neurosci. Res.* 54, 276–280.

Imada, T., Zhang, Y., Cheour, M., Taulu, S., Ahonen, A., and 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 *Bird Vocalizations*, R. Hinde, ed. (London: Cambridge University Press), pp. 61–74.

Johnson, J.S., and Newport, E.L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognit. Psychol.* 21, 60–99.

Kirkham, N.Z., Slemmer, J.A., and 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. *J. Cogn. Neurosci.* 16, 1412–1425.

Krause, C.M., Sillanmäki, L., Koivisto, M., Saarela, C., Häggqvist, A., Laine, M., and Hämäläinen, H. (2000). The effects of memory load on event-related EEG desynchronization and synchronization. *Clin. Neurophysiol.* 111, 2071–2078.

Kuhl, P.K. (2004). Early language acquisition: cracking the speech code. *Nat. Rev. Neurosci.* 5, 831–843.

- Kuhl, P.K. (2007). Is speech learning 'gated' by the social brain? *Dev. Sci.* 10, 110–120.
- Kuhl, P.K. (2009). Early language acquisition: neural substrates and theoretical models. In *The Cognitive Neurosciences IV*, M.S. Gazzaniga, ed. (Cambridge, MA: MIT Press), pp. 837–854.
- Kuhl, P.K., and Meltzoff, A.N. (1982). The bimodal perception of speech in infancy. *Science* 218, 1138–1141.
- Kuhl, P.K., and Meltzoff, A.N. (1996). Infant vocalizations in response to speech: vocal imitation and developmental change. *J. Acoust. Soc. Am.* 100, 2425–2438.
- Kuhl, P.K., and Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *Annu. Rev. Neurosci.* 31, 511–534.
- Kuhl, P.K., Williams, K.A., and Meltzoff, A.N. (1991). Cross-modal speech perception in adults and infants using nonspeech auditory stimuli. *J. Exp. Psychol. Hum. Percept. Perform.* 17, 829–840.
- Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., and Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255, 606–608.
- Kuhl, P.K., Andruski, J.E., Chistovich, I.A., Chistovich, L.A., Kozhevnikova, E.V., Ryskina, V.L., Stolyarova, E.I., Sundberg, U., and Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science* 277, 684–686.
- Kuhl, P.K., Tsao, F.-M., and Liu, H.-M. (2003). Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proc. Natl. Acad. Sci. USA* 100, 9096–9101.
- Kuhl, P.K., Conboy, B.T., Padden, D., Nelson, T., and Pruitt, J. (2005a). Early speech perception and later language development: implications for the 'critical period'. *Lang. Learn. Dev.* 1, 237–264.
- Kuhl, P.K., Coffey-Corina, S., Padden, D., and Dawson, G. (2005b). Links between social and linguistic processing of speech in preschool children with autism: behavioral and electrophysiological measures. *Dev. Sci.* 8, F1–F12.
- Kuhl, P.K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., and Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Dev. Sci.* 9, F13–F21.
- Kuhl, P.K., Conboy, B.T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., and Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 979–1000.
- Ladefoged, P. (2001). *Vowels and Consonants: An Introduction to the Sounds of Language* (Oxford: Blackwell Publishers).
- Lasky, R.E., Syrdal-Lasky, A., and Klein, R.E. (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* (New York: John Wiley and Sons).
- Lieberman, A.M., and Mattingly, I.G. (1985). The motor theory of speech perception revised. *Cognition* 21, 1–36.
- Liu, H.-M., Kuhl, P.K., and Tsao, F.-M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Dev. Sci.* 6, F1–F10.
- Liu, H.-M., Tsao, F.-M., and Kuhl, P.K. (2007). Acoustic analysis of lexical tone in Mandarin infant-directed speech. *Dev. Psychol.* 43, 912–917.
- Liu, H.-M., Tsao, F.-M., and Kuhl, P.K. (2009). Age-related changes in acoustic modifications of Mandarin maternal speech to preverbal infants and five-year-old children: a longitudinal study. *J. Child Lang.* 36, 909–922.
- Lotto, A.J., Sato, M., and 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*, J. Slitka, S. Manuel, and M. Matthies, eds. (Cambridge, MA: MIT Press), pp. C181–C186.
- Marler, P. (1991). The instinct to learn. In *The Epigenesis of Mind: Essays on Biology and Cognition*, S. Carey and R. Gelman, eds. (Hillsdale, NJ: Lawrence Erlbaum Associates), pp. 37–66.
- Mayberry, R.I., and 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, J.F., and Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82, B101–B111.
- Maye, J., Weiss, D.J., and Aslin, R.N. (2008). Statistical phonetic learning in infants: facilitation and feature generalization. *Dev. Sci.* 11, 122–134.
- Meltzoff, A.N., and Decety, J. (2003). What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 491–500.
- Meltzoff, A.N., Kuhl, P.K., Movellan, J., and Sejnowski, T. (2009). Foundations for a new science of learning. *Science* 17, 284–288.
- Mills, D., Coffey-Corina, S., and Neville, H. (1993). Language acquisition and cerebral specialization in 20 month old infants. *J. Cogn. Neurosci.* 5, 317–334.
- Mills, D., Coffey-Corina, S., and Neville, H. (1997). Language comprehension and cerebral specialization from 13 to 20 months. *Dev. Neuropsychol.* 13, 397–445.
- Mundy, P., and Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant Behav. Dev.* 21, 469–482.
- Nelson, D.A., and Marler, P. (1994). Selection-based learning in bird song development. *Proc. Natl. Acad. Sci. USA* 91, 10498–10501.
- Neville, H.J., Coffey, S.A., Lawson, D.S., Fischer, A., Emmorey, K., and Bellugi, U. (1997). Neural systems mediating American sign language: effects of sensory experience and age of acquisition. *Brain Lang.* 57, 285–308.
- Newport, E. (1990). Maturation constraints on language learning. *Cogn. Sci.* 14, 11–28.
- Newport, E.L., Bavelier, D., and Neville, H.J. (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 Mehler*, E. Dupoux, ed. (Cambridge, MA: MIT Press), pp. 481–502.
- Ortiz-Mantilla, S., Choe, M.-S., Flax, J., Grant, P.E., and Benasich, A.A. (2010). Associations between the size of the amygdala in infancy and language abilities during the preschool years in normally developing children. *Neuroimage* 49, 2791–2799.
- Patterson, M.L., and Werker, J.F. (1999). Matching phonetic information in lips and voice is robust in 4.5-month-old infants. *Infant Behav. Dev.* 22, 237–247.
- Peña, M., Bonatti, L.L., Nespor, M., and Mehler, J. (2002). Signal-driven computations in speech processing. *Science* 298, 604–607.
- Petitto, L.A., and Marentette, P.F. (1991). Babbling in the manual mode: evidence for the ontogeny of language. *Science* 251, 1493–1496.
- Posner, M.I., ed. (2004). *Cognitive Neuroscience of Attention* (New York: Guilford Press).
- Pulvermuller, F. (2005). Brain mechanisms linking language to action. *Nat. Rev. Neurosci.* 6, 574–582.
- Rabiner, L.R., and Huang, B.H. (1993). *Fundamentals of Speech Recognition* (Englewood Cliffs, NJ: Prentice Hall).
- Rivera-Gaxiola, M., Silva-Pereyra, J., and Kuhl, P.K. (2005). 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 – I: Mechanisms of Imitation and Imitation in Animals*, S. Hurley and N. Chater, eds. (Cambridge, MA: MIT Press), pp. 55–76.
- Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192.

- Saffran, J.R., Aslin, R.N., and Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science* 274, 1926–1928.
- Saffran, J.R., Johnson, E.K., Aslin, R.N., and Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition* 70, 27–52.
- Saffran, J.R., Werker, J.F., and Werner, L.A. (2006). The infant's auditory world: hearing, speech, and the beginnings of language. In *Handbook of Child Psychology: Volume 2, Cognition, Perception and Language VI*, W. Damon and R.M. Lerner, series eds., R. Siegler and D. Kuhn, volume eds. (New York: Wiley), pp. 58–108.
- Skinner, B.F. (1957). *Verbal Behavior* (New York: Appleton-Century-Crofts).
- Taga, G., and Asakawa, K. (2007). Selectivity and localization of cortical response to auditory and visual stimulation in awake infants aged 2 to 4 months. *Neuroimage* 36, 1246–1252.
- Tamm, L., Menon, V., and Reiss, A.L. (2002). Maturation of brain function associated with response inhibition. *J. Am. Acad. Child Adolesc. Psychiatry* 41, 1231–1238.
- Teinonen, T., Aslin, R.N., Alku, P., and Csibra, G. (2008). Visual speech contributes to phonetic learning in 6-month-old infants. *Cognition* 108, 850–855.
- Teinonen, T., Fellman, V., Nääätänen, R., Alku, P., and Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neurosci.* 10, 21.
- Tomasello, M. (2003a). *Constructing A Language: A Usage-Based Theory of Language Acquisition* (Cambridge, MA: Harvard University Press).
- Tomasello, M. (2003b). The key is social cognition. In *Language and Thought*, D. Gentner and S. Kuczaj, eds. (Cambridge, MA: MIT Press), pp. 47–51.
- Tsao, F.-M., Liu, H.-M., and Kuhl, P.K. (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., and Kuhl, P.K. (2006). Perception of native and non-native affricate-fricative contrasts: cross-language tests on adults and infants. *J. Acoust. Soc. Am.* 120, 2285–2294.
- Vygotsky, L.S. (1962). *Thought and Language* (Cambridge, MA: MIT Press).
- Wang, Y., Kuhl, P.K., Chen, C., and Dong, Q. (2009). Sustained and transient language control in the bilingual brain. *Neuroimage* 47, 414–422.
- Weber-Fox, C.M., and Neville, H.J. (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*, D. Birdsong, ed. (Mahwah, NJ: Lawrence Erlbaum and Associates, Inc), pp. 23–38.
- Werker, J.F., and Curtin, S. (2005). PRIMIR: a developmental framework of infant speech processing. *Lang. Learn. Dev.* 1, 197–234.
- Werker, J.F., and Lalonde, C. (1988). Cross-language speech perception: initial capabilities and developmental change. *Dev. Psychol.* 24, 672–683.
- Werker, J.F., and Tees, R.C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behav. Dev.* 7, 49–63.
- Werker, J.F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., and Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition* 103, 147–162.
- West, M.J., and King, A.P. (1988). Female visual displays affect the development of male song in the cowbird. *Nature* 334, 244–246.
- White, L., and Genesee, F. (1996). How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Lang. Res.* 12, 233–265.
- Woolley, S.C., and 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., and Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean–English bilinguals. *Bilingualism: Lang. Cogn.* 3, 131–149.
- Yoshida, K.A., Pons, F., Cady, J.C., and Werker, J.F. (2006). Distributional learning and attention in phonological development. Paper presented at International Conference on Infant Studies, Kyoto, Japan, 19–23 June.
- Yoshida, K.A., Pons, F., Maye, J., and Werker, J.F. (2010). Distributional phonetic learning at 10 months of age. *Infancy* 15, 420–433.
- Zhang, Y., Kuhl, P.K., Imada, T., Kotani, M., and 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.B., Kawakatsu, M., Tohkura, Y., and Nemoto, I. (2009). Neural signatures of phonetic learning in adulthood: a magnetoencephalography study. *Neuroimage* 46, 226–240.