

# Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e)

Patricia K. Kuhl<sup>1,2,\*</sup>, Barbara T. Conboy<sup>1</sup>, Sharon Coffey-Corina<sup>1</sup>,  
Denise Padden<sup>1</sup>, Maritza Rivera-Gaxiola<sup>1,2</sup> and Tobey Nelson<sup>1</sup>

<sup>1</sup>*Institute for Learning and Brain Sciences, and* <sup>2</sup>*Department of Speech and Hearing Sciences, University of Washington, Seattle, WA 98195, USA*

Infants' speech perception skills show a dual change towards the end of the first year of life. Not only does non-native speech perception decline, as often shown, but native language speech perception skills show improvement, reflecting a facilitative effect of experience with native language. The mechanism underlying change at this point in development, and the relationship between the change in native and non-native speech perception, is of theoretical interest. As shown in new data presented here, at the cusp of this developmental change, infants' native and non-native phonetic perception skills predict later language ability, but in opposite directions. Better *native* language skill at 7.5 months of age predicts faster language advancement, whereas better *non-native* language skill predicts slower advancement. We suggest that native language phonetic performance is indicative of *neural commitment* to the native language, while non-native phonetic performance reveals *uncommitted* neural circuitry. This paper has three goals: (i) to review existing models of phonetic perception development, (ii) to present new event-related potential data showing that native and non-native phonetic perception at 7.5 months of age predicts language growth over the next 2 years, and (iii) to describe a revised version of our previous model, the native language magnet model, expanded (NLM-e). NLM-e incorporates five new principles. Specific testable predictions for future research programmes are described.

**Keywords:** theories of speech perception; language development; event-related potentials; effects of linguistic experience; neural commitment

## 1. INTRODUCTION

From babbling at six months of age to full sentences by the age of 3, young children learn their mother tongue rapidly and effortlessly, following similar developmental paths regardless of culture (figure 1). How infants accomplish the task has become the focus of debate on the nature and origins of language (Hauser *et al.* 2002a; Kuhl 2004; Newport & Aslin 2004; Fitch *et al.* 2005; Pinker & Jackendoff 2005).

Research and theory are now aimed at elucidating the mechanisms underlying developmental change in infants' perception of speech, and the emerging picture is complex. Young infants learn 'statistically' (Saffran 2002) at many levels, including phonetic (Kuhl *et al.* 1992; Maye *et al.* 2002, *in press*), phonotactic (Jusczyk *et al.* 1993, 1994), lexical (Goodsitt *et al.* 1993; Saffran *et al.* 1996) and syntactic (Gomez & Gerken 1999, 2000; Marcus *et al.* 1999; Peña *et al.* 2002; Bortfeld *et al.* 2005; Gerken *et al.* 2005). However, learning requires more than computation. In experimental interventions mimicking natural language learning situations, social factors play an important role (Kuhl *et al.* 2003; Yu *et al.* 2005; Yoshida *et al.* 2006). When

exposed to a new language for the first time at nine months of age, for example, infants learn phonetically from a live interacting human being, but not from a disembodied source, even though the acoustic information remains the same in the two situations (Kuhl *et al.* 2003). While social factors in language learning have often been discussed (Bruner 1983; Baldwin 1995; Tomasello 2003), the potential role that social factors may play in speech perception development has only recently been explored (Kuhl 2007).

Non-linguistic cognitive factors also appear to play a role in phonetic learning. Previous research (Lalonde & Werker 1995) suggested a link between general cognitive abilities and reductions in non-native perception towards the end of the first year. The increasing ability to inhibit attention to irrelevant information was offered as a possible mechanism underlying performance on both phonetic discrimination and non-linguistic tasks (Diamond *et al.* 1994). More recent work (Conboy *et al.* submitted) has replicated and extended that finding using a different set of non-linguistic tasks that tap cognitive control skills. These findings, along with research showing that children raised bilingually from birth have an advantage on non-linguistic tasks requiring control of attention (Bialystok 1999, 2001), suggest that responses to speech stimuli are linked to cognitive control, with inhibitory control

\* Author for correspondence (pkkuhl@u.washington.edu).

One contribution of 13 to a Theme Issue 'The perception of speech: from sound to meaning'.

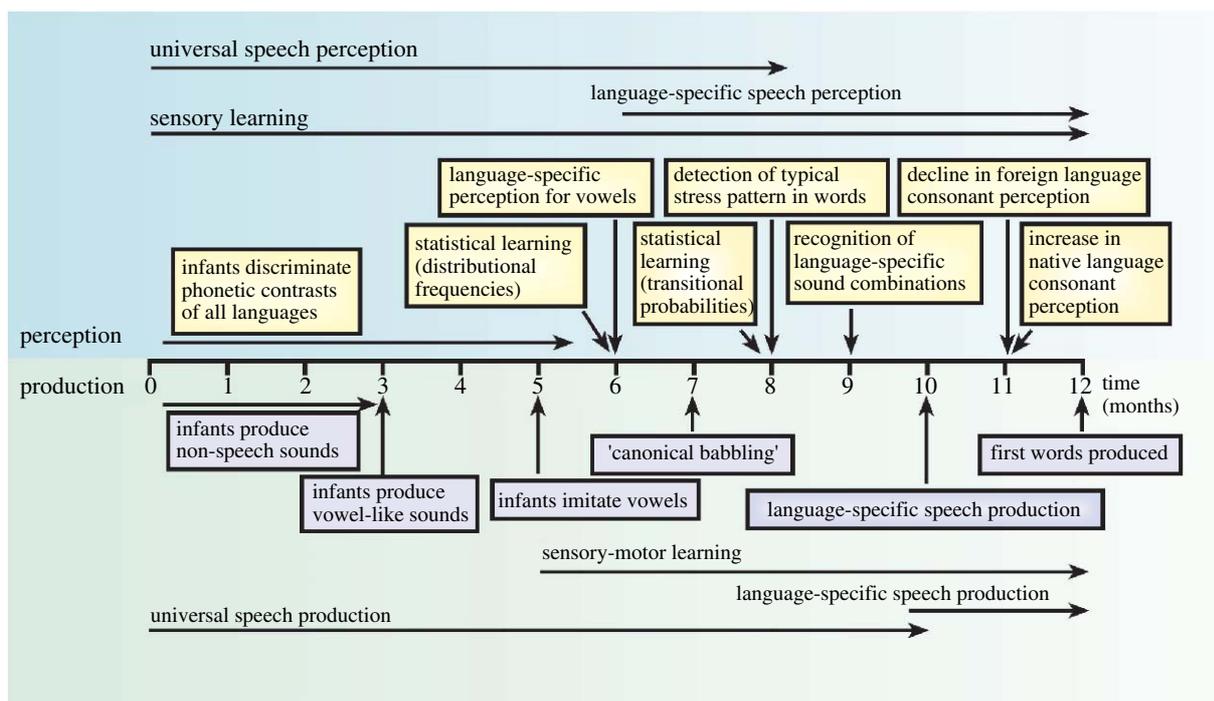


Figure 1. Universal timeline of infants' perception and production of speech in the first year of life. Modified from Kuhl (2004).

playing a role in non-native speech perception in monolingual infants (Diamond *et al.* 1994; Conboy *et al.* submitted).

Moreover, there is continuity between early measures of phonetic perception and later language skills (Molfese & Molfese 1997; Molfese 2000; Tsao *et al.* 2004; Conboy *et al.* 2005; Kuhl *et al.* 2005b; Rivera-Gaxiola *et al.* 2005a). A new study will be presented here that utilizes an event-related potential (ERP) brain measure of infants' native and non-native speech perception in infancy to predict language in the second and third year of life. The new findings allow further development of an explanation for the difference between native and non-native contrasts as predictors of later language. The data provide some support for the neural commitment hypothesis as a potential contributor to the 'critical period' in phonetic learning.

This paper has three goals: (i) to examine theoretical perspectives related to the earliest phases of language acquisition, (ii) to present a new ERP experiment that explores the linkage between early speech perception and later language development, and (iii) to elaborate on our original theoretical model, the native language magnet (NLM) model (Kuhl 1992, 1994), producing a revised version, *native language magnet theory, expanded (NLM-e)*. NLM-e incorporates five new principles and makes specific, empirically testable, predictions.

## 2. THE DEVELOPMENTAL PROBLEM AND EARLY THEORY

To acquire a language, infants have to discover which phonetic distinctions will be utilized in the language of their culture. For example, English is different from Japanese; the phonemes /r/ and /l/ create different words in English ('rake' and 'lake'), but do not change the meaning of a word in Japanese. Our understanding of this process is anchored by two well-established

facts. Early in life, infants discriminate among virtually all the phonetic units of the world's languages (Eimas *et al.* 1971; Streeter 1976; Trehub 1976; Werker & Tees 1984a; Best & McRoberts 2003; Kuhl *et al.* 2006). By adulthood, this universal phonetic capacity is no longer in place and non-native phonetic discrimination is much more difficult (Miyawaki *et al.* 1975; Werker & Lalonde 1988; Best *et al.* 2001; Iverson *et al.* 2003). Our focus is the mechanism underlying this developmental transition.

Historical models of developmental speech perception were based on *selection*. Infants' phonetic abilities were argued to stem from an innate specification of all possible phonetic units, which were subsequently maintained or lost as a function of linguistic experience. Eimas' phonetic feature detector account (Eimas 1975) and Liberman's motor theory (Liberman *et al.* 1967; Liberman & Mattingly 1985) are classic examples. The primitives differ in the two accounts—Eimas' phonetic feature detectors were responsive to the acoustic events underlying phonetic distinctions, whereas Liberman's motor theory held that infants initially detect all phonetically relevant gestures (Liberman & Mattingly 1985); but both views were based on selection and the maintenance/loss view. Early data on developmental change in infants' perception of speech (Werker & Tees 1984a) supported this view; native abilities appeared to be maintained and non-native abilities lost.

In the 1970s, the discovery of 'categorical perception' for speech in non-human animals (Kuhl & Miller 1975, 1978; Kuhl & Padden 1982, 1983; see also Dooling *et al.* 1995), and demonstrations of categorical perception for non-speech stimuli in infants (Jusczyk *et al.* 1977, 1980), undermined the selection model and provided an alternative explanation for infants' abilities which was rooted in neurobiology and evolution. The argument was that phonetic contrasts in speech capitalized on existing, more general properties of

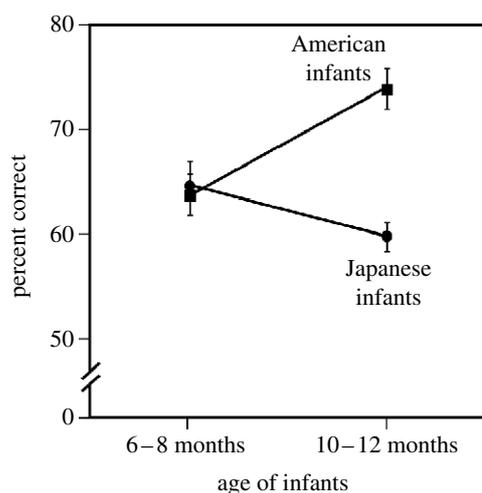


Figure 2. The effects of age on speech perception performance in a cross-language study of the perception of American English /r-l/ sounds by American and Japanese infants. From Kuhl *et al.* (2006).

auditory perceptual mechanisms rather than specially evolved phonetic feature detectors. These data suggested that infants' initial abilities were more primitive—the base on which language builds—rather than domain-specific mechanisms evolved for language (Kuhl & Miller 1978; Kuhl 1991a). Additional comparative studies subsequently revealed many more similarities in human–animal speech perception (Kluender *et al.* 1987; Hauser *et al.* 2001, 2002b), and of equal importance, differences between human and animal perception and learning (Kuhl 1991b; Fitch & Hauser 2004; Newport & Aslin 2004).

The idea that infants began life with phonetic feature detectors that were either maintained or lost was further undermined by adult (Carney *et al.* 1977; Pisoni *et al.* 1982; Werker & Tees 1984b, Werker & Logan 1985; Werker 1995; Rivera-Gaxiola *et al.* 2000) and infant results (Cheour *et al.* 1998; Rivera-Gaxiola *et al.* 2005b, 2007; Kuhl *et al.* 2006; Tsao *et al.* 2006). Both show that we remain capable of discriminating non-native phonetic contrasts, though at a reduced level when compared with native contrasts.

However, the idea that more than selection is involved in developmental phonetic perception was most clearly demonstrated by experimental findings showing that native language phonetic perception shows a significant improvement between 6 and 12 months of age. American infants tested on the English /r-l/ contrast showed a statistically significant improvement between 6 and 12 months of age (Kuhl *et al.* 2006; figure 2). Also, both Mandarin-learning and English-learning infants showed native language phonetic improvement on affricate–fricative contrasts between 6 and 12 months of age (Tsao *et al.* 2006). Brain measures in the form of electrophysiological ERPs in response to syllable changes between 6 and 12 months of age also showed an increase in native consonant perception (Rivera-Gaxiola *et al.* 2005b, 2007); the same pattern in ERP data has been shown for vowels (Cheour *et al.* 1998). Previous studies had shown native language improvement *after* 12 months of age and before adulthood (Polka *et al.* 2001; Sundara

*et al.* 2006), but the new studies establish a pattern of improvement in native language phonetic perception in the first year, which we consider significant for theory. The improvement suggested that selection—a process of maintenance or loss—could not account for the transition in phonetic perception between 6 and 12 months of life.

Models of the earliest phases of language acquisition required an explanation of (i) the facilitation pattern seen for native language contrasts between 6 and 12 months of age, (ii) the typical decline seen for many non-native contrasts at the same point in time, (iii) the variability observed across contrasts, and (iv) the relationship between changes in native and non-native speech perception and their differential prediction of future language.

### 3. BEYOND MODELS OF SELECTION

Three newer models went beyond selection in explaining developmental change in infants' perception of speech; we will review models offered by Werker, Best and Kuhl.

Werker and colleagues focused on the decline in non-native perception and described six possible explanations (Werker & Pegg 1992), one of which focused on cognitive abilities. Infants' performance on non-native phonetic contrasts was associated with performance on visual categorization and the A-not-B task; 8- to 10-month-old infants with better performance on either non-linguistic task had poor discrimination of a non-native (Hindi-voiced, unaspirated dental versus retroflex stop) contrast (Lalonde & Werker 1995). The findings of that study were consistent with the view that age-related changes in non-native speech discrimination are influenced by broad, domain-general cognitive processes. Diamond *et al.* (1994) further suggested a link between inhibitory control skills and discrimination of non-native contrasts. A recent study with 11-month-old infants indicates that reduction in non-native discrimination skills at this age is linked to better performance on a detour-reaching task that requires inhibition of prepotent responses, and more goal-directed behaviours on a means-end task (Conboy *et al.* submitted). The finding that native language perceptual abilities are not associated with non-linguistic skills in either study suggests that cognitive control abilities may play a specific role in the ability to disregard irrelevant phonetic information while maintaining attention to relevant information.

Best's perceptual assimilation model (PAM; Best 1994, 1995; Best & McRoberts 2003) also focused on the decline in non-native speech perception, arguing that infants' recognition of the articulatory gestures underlying speech explains the change in non-native perception at 10–12 months of age. PAM indicates that infants' difficulty in discriminating non-native contrasts is predicted by the articulatory similarity between specific native and non-native categories. In an update of the model, Best describes PAM/AO (articulatory organ) and suggests that non-native discrimination declines when phonetic contrasts involve the same articulatory organ (/s-z/), as opposed to different articulatory organs (/b-t/). Tests on non-native perception in 6–8 and 10–12 months old American infants

supported its predictions (Best & McRoberts 2003). Recent tests on additional different organ contrasts, such as liquids (Kuhl *et al.* 2006) and affricate–fricatives (Tsao *et al.* 2006), confirm PAM's predictions when they are non-native contrasts for the infants tested, but not when they are native contrasts for the infants—both showed facilitation effects.

Kuhl offered a model of early speech perception termed the NLM model, which focused on infants' native phonetic categories and how they could be structured through ambient language experience (Kuhl 1994, 2000*a,b*). NLM specified three phases in development. In phase 1, the initial state, infants are capable of differentiating all the sounds of human speech, and these abilities derive from their general auditory processing mechanisms rather than from a speech-specific mechanism (Kuhl 1991*b*). In phase 2, infants' sensitivity to the distributional properties of linguistic input produces phonetic representations based on the distributional 'modes' in ambient speech input (Kuhl 1993). Experience is described as 'warping' perception, producing a distortion that decreases perceptual sensitivity near category modes and increases perceptual sensitivity near the boundaries between categories (Kuhl 1991*a*; Kuhl *et al.* 1992; Iverson *et al.* 2003). As experience accumulates, the representations most often activated (*prototypes*) begin to function as *perceptual magnets* for other members of the category, increasing the perceived similarity between members of the category (Kuhl 1991*a*). In phase 3, this distortion of perception, termed the *perceptual magnet effect*, produces facilitation in native and a reduction in foreign language phonetic abilities.

Accounts of infant speech perception, such as those of Jusczyk (1997), Kuhl (1992, 2000*a*) and more recently Werker & Curtin (2005), increasingly resemble more general developmental cognitive learning models (Karmiloff-Smith 1992; Elman *et al.* 1996; Gopnik & Meltzoff 1997). Studies on animals using speech and on infants using stimuli across domains suggest that aspects of infant speech perception, at least initially, are domain general and available to non-human species as well as human infants (Kuhl & Miller 1975; Marcus *et al.* 1999; Gomez & Gerken 2000; Hauser *et al.* 2002*b*; Saffran 2003; Newport & Aslin 2004; Newport *et al.* 2004), but that phonetic *learning* is unique to humans (Kuhl 1991*a,b*, 2000*a*). There is an increasing consensus for the idea, proposed following the initial animal studies (Kuhl & Miller 1975; Kuhl 1991*a*), that language evolved to match a set of general perceptual and learning abilities—a notion that is at odds with Skinner's (1957) learning-through-explicit-reinforcement view, but also at some variance with Chomsky's original notion of an innately specified universal phonetics and universal grammar, leading to a revision in the theory that Chomsky acknowledges (Hauser *et al.* 2002*a*).

#### 4. NATIVE LANGUAGE MAGNET THEORY, EXPANDED

The principles, learning account and framework described by the original NLM (Kuhl 1992, 1994) are the starting point for this revision of the theory. In this section, we (i) review the five general principles

guiding the model, (ii) present the results of a new experiment, and (iii) describe NLM-e.

##### (a) *Basic principles of NLM-e*

(i) *Distributional patterns and infant-directed speech are agents of change*

We describe a model in which two agents of change drive the transition from a universal pattern of phonetic perception to one that is language specific: (i) detection of distributional frequencies in the patterns of phonetic units in ambient speech and (ii) the exaggerated acoustic cues to phonetic units contained in infant-directed (ID) speech, often referred to as 'motherese'.

Distributional differences in the patterns of native language input were first suggested as an explanation of infants' language-specific perception of vowels at six months of age based on the results of a cross-language study (Kuhl *et al.* 1992). The interpretation of the study was that distributional differences in native language speech heard by infants in two different countries during the first six months of life caused language-specific representations (*prototypes*) to develop, which altered infants' perception (Kuhl 1993). Recently, Maye and her colleagues conducted direct tests that examined whether infants are sensitive to distributional frequency differences in speech input; they presented 6- and 8-month-old infants with syllables from a continuum for 2 min (Maye *et al.* 2002). Infants experienced stimuli from the entire continuum, but with different distributional frequencies. 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 were tested on the endpoints of the continuum; infants in the bimodal group discriminated the sounds, whereas those in the unimodal group did not. Taken together with data showing infants' sensitivity to distributional patterns and the role they play in word recognition (McMurray & Aslin 2005), these data support the view that infants' sensitivity to distributional properties can induce the changes observed in early phonetic perception.

A second agent of change proposed in the NLM-e model is the exaggeration of relevant phonetic differences by adult speakers during ID as opposed to adult-directed (AD) speech. Studies show that mothers addressing infants acoustically 'stretch' the acoustic cues of phonetic units, exaggerating their differences and making them more discriminable (Bernstein-Ratner 1984; Kuhl *et al.* 1997; Burnham *et al.* 2002; Liu *et al.* 2003, *in press*). For example, stretching the formant frequencies of vowels makes them more distinct and creates more intelligible speech (see Liu *et al.* (2003) for review). Importantly, the exaggeration of vowels in ID speech has been shown to be unique to language addressing infants as opposed to that used when addressing pets (Burnham *et al.* 2002). The ID speech effect is not limited to vowels; Mandarin mothers expand the frequency differences among tones that are phonemic in Mandarin, exaggerating their distinctiveness (Liu *et al.* *in press*). Consonantal contrasts involving voice onset time (VOT) have also been compared in ID and AD speech, but with somewhat

mixed results (Baran *et al.* 1977; Malsheen 1980; Sundberg & Lacerda 1999). A recent, more comprehensive study investigated stop consonants in Norwegian ID and AD speech throughout the first six months of life, showing that differences in VOT in stop consonants were exaggerated in ID as opposed to AD speech, a difference that was stable across the six-month period (Englund 2005). Increasing the differences among phonetic units makes their contrastive features easier to learn, and this should assist in typically developing infants' language learning; it could be especially helpful for children with auditory perceptual deficits (Merzenich *et al.* 1996; Tallal *et al.* 1996).

Liu *et al.* (2003) provided the first evidence that mothers' increased intelligibility in ID speech may be helpful for infants. They measured individual mother's vowels during ID speech and subsequently measured her infant's speech perception skills in the laboratory using computer-synthesized consonant sounds. The degree to which an individual mother exaggerated the acoustic cues during ID speech was significantly correlated with her infant's speech perception abilities. This was replicated in two independent samples of mother–infant pairs, in mothers with 6- to 8-month-old infants and in those with 10- to 12-month-old infants.

Acoustic stretching of phonetic cues in ID speech would be expected to exaggerate the distributional cues to phonetic units, and there is some evidence to validate this (Werker *et al.* 2007). Computer-learning models also indicate that ID speech improves the robustness of category learning achieved over that obtained with AD speech (de Boer & Kuhl 2003). In order for ID speech to support phonetic learning, infants have to be interested in listening to it. Studies of typically developing infants, even newborns, suggest a preference for speech, and especially ID speech (Fernald 1984; Fernald & Kuhl 1987; Vouloumanos & Werker 2004). And when a social interest in speech is absent, as is the case for children with autism, our tests show that non-speech analogue signals are preferred over ID speech, and that the strength of preference for non-speech predicts severity of autism symptoms as well as the degree to which neural responses to speech are aberrant (Kuhl *et al.* 2005a). Thus, research on both typical children and those with developmental disabilities suggest that social factors play an important role in the earliest phases of language acquisition.

(ii) *Language exposure produces neural commitment that affects future learning*

A growing number of studies confirm the effects of language experience on an adult brain (Sanders *et al.* 2002; Koyama *et al.* 2003; Perani *et al.* 2003; Wang *et al.* 2003; Callan *et al.* 2004; Golestani & Zatorre 2004; Zhang *et al.* 2005). Neural imaging techniques have also increasingly been applied to infants and young children (Dehaene-Lambertz & Gliga 2004; Mills *et al.* 2004, 2005a,b; Friederici 2005; Rivera-Gaxiola *et al.* 2005a,b, 2007; Silva-Pereyra *et al.* 2005, 2007; Conboy & Mills 2006; Imada *et al.* 2006).

To explain the effects of language experience on the brain, we proposed the concept of *native language neural commitment* (NLNC), arguing that the brain's early coding of language affects our subsequent abilities

to learn the phonetic scheme of a new language (Kuhl 2000a,b, 2004). NLNC describes a process in which initial language exposure causes physical changes in neural tissue and circuitry that reflect the statistical and perceptual properties of language input. Neural networks become committed to patterns of native language speech producing bi-directional effects. They reinforce the detection of higher-order patterns in language (morphemes, words) that capitalize on learned phonetic patterns, while at the same time reducing sensitivity to alternative phonetic schemes (Kuhl 2004). In development, the increase in native language phonetic perception thus reflects neural commitment. In contrast, infants' non-native phonetic abilities reflect a more immature state of uncommitted circuitry. Progress towards language thus requires committing neural circuitry to the patterns of native language speech (Kuhl 2004). NLNC thus predicts that native as opposed to non-native speech perception, measured at the cusp of phonetic learning, should produce differential patterns of association with later language abilities, a pattern we have now confirmed experimentally (see below).

In adults, a number of studies support the idea that linguistic experience 'interferes' with phonetic learning of a new language in adulthood (Flege 1995; McCandliss *et al.* 2002; Iverson *et al.* 2003; Zhang *et al.* 2005, submitted). Adult studies of /r-l/ stimuli varying in F2 and F3 using speakers of English, Japanese and German show that speakers attend to different dimensions of the same stimulus (Iverson *et al.* 2003). Japanese adults, for example, are sensitive to an acoustic cue (second formant) that is irrelevant to the categorization of English /r-l/, and this interferes with correct categorization. We argue that early exposure to language shapes these attentional networks, and that in adulthood, they make second language learning difficult. Early in infancy, neural commitment is a 'soft' constraint; infants' networks are not fully developed and therefore interference is weak and infants can acquire more than one language.

Studies using magnetoencephalography (MEG) can examine both the spatial location and time course of the brain's response to native versus non-native patterns. When processing non-native speech sounds, the adult brain is activated over a significantly longer duration and a significantly larger area than when processing native language sounds, showing that the processing of non-native sounds is neurally time consuming and requires additional brain resources (Zhang *et al.* 2005). MEG studies also indicate that training adults on second language contrasts can be successful and is enhanced by the exaggeration of phonetic cues in a manner similar to motherese (Pisoni & Lively 1995; Iverson *et al.* 2005; Vallabha & McClelland 2007; Zhang *et al.* submitted).

Computational neural modelling of second language phonetic learning supports the approach described by NLM-e. Models by McClelland and colleagues (McCandliss *et al.* 2002; Vallabha & McClelland 2007) and Guenther and colleagues (Guenther *et al.* 2004) describe phonetic learning as Hebbian unsupervised learning, shaping 'attractor' networks that function similarly to the perceptual magnet effect of

NLM-e. In the visual domain, a related formulation produces attractors that subsequently increase the perceived similarity of neighbouring stimuli (Rosenthal *et al.* 2001).

(iii) *Social interaction influences early language learning at the phonetic level*

Current studies show that young infants have computational skills that assist language acquisition; simple laboratory experiments indicate that statistical learning can occur with just 2 min exposure to novel speech material (Saffran *et al.* 1996; Maye *et al.* 2002). Nevertheless, social influences on computational learning were recently shown in a study investigating whether infants are capable of phonetic learning at nine months of age at natural first-time exposure to a foreign language.

Mandarin Chinese, a language with prosodic and phonetic structures very different from those in English, was used in a foreign language intervention experiment (Kuhl *et al.* 2003). Infants heard four native speakers of Mandarin (both male and female) during twelve 25 min sessions of book reading and play during a four to six week period. A control group of infants also came into the laboratory for the same number and variety of reading and play sessions, but heard only English. On average, infants heard approximately 33 000 Mandarin syllables during the course of the 12 language-exposure sessions, including approximately 1000 instances of each of the two Mandarin syllables (an affricate–fricative contrast not phonemic in English) that were used to test infants after exposure.

The results of both behavioural (Kuhl *et al.* 2003) and brain (Kuhl *et al.* in preparation) tests on infants after exposure demonstrated that Mandarin-exposed infants performed significantly better on the Mandarin test syllables than the English control group, indicating that phonetic learning from first-time exposure could occur at nine months of age. Learning was sufficiently durable to allow infants to perform in the behavioural tests 2–12 days (with a median of 6 days) after the final language-exposure session; no differences were seen in infant performance as a function of the delay.

In two additional conditions, Kuhl *et al.* (2003) examined whether social interaction played a significant role in this complex natural language learning situation. Infants experienced the exact same language material, but from a disembodied source, either a television or an audiotape (Kuhl *et al.* 2003). The auditory statistical cues available to the infants were identical in the media-delivered and live settings, as was the use of ID motherese (Fernald & Kuhl 1987; Kuhl *et al.* 1997). If exposure to language automatically prompts statistical learning, the presence of a live human being will not be essential. However, infants' Mandarin discrimination scores after exposure to televised or audiotaped speakers were no greater than those of the control infants who had not experienced Mandarin at all; both the TV-exposed and the audio-exposed groups differed significantly from the live-exposure group but not from the control group. The data suggest that in natural, complex language learning situations, infants may require a social tutor to learn, i.e. they are not computational automatons.

Similar second language exposure experiments are now underway using Spanish and they are exploring both phonetic and word learning and the extent to which social factors such as visual attention during the exposure sessions predict the degree to which individual infants learn (Conboy & Kuhl 2007). These experiments are suggesting relationships between infants' cognitive skills and/or their attention to language tutors and the ability to learn from second language exposure.

Many authors have discussed the role of social interaction in language learning (Bruner 1983; Baldwin 1995; Tomasello 2003), but the role of social interaction in phonetic learning has not previously been investigated. Does the finding that social interaction affects language learning in more natural settings invalidate studies showing that phonetic (Maye *et al.* 2002) and word (Saffran *et al.* 1996) learning can be demonstrated with very short-term laboratory exposure in the absence of a social context? Clearly, not all learning requires a social context; short-term exposures can result in learning in the absence of social interaction. Recent studies show that attention enhances simple distributional learning in the laboratory (Yoshida *et al.* 2006). Complex natural language learning may demand social interaction, because social cues from competent tutors can highlight relevant information for the learner. Language evolved in a social setting and the neurobiological mechanisms underlying it probably utilize interactional cues made available only in a social setting.

In other species, such as songbirds, communicative learning is also enhanced by social contact, and contingency plays a role. Visual interaction with a tutor bird is often necessary to learn song in the laboratory (Eales 1989), and, a live foster father of another species who feeds young birds can override an innate preference for conspecific song, even when conspecific adults can be heard nearby (Immelmann 1969). A live tutor allows learning of an alien song when audiotaped presentations of these alien songs are rejected (Baptista & Petrinovich 1986). Social interaction thus enhances learning in birds. It influences interpersonal cognition and 'theory of mind' in humans (Meltzoff 2005). Social interaction may affect language learners in similar ways.

(iv) *The perception–production link is forged developmentally*

NLM-e predicts strong linkages between the perceptual representations formed through experience with language and vocal imitation. In this respect, it is similar to earlier theoretical positions arguing for close interaction between speech perception and production, such as the motor theory (Lieberman *et al.* 1967) and direct realism (Fowler 1986). However, a distinction can be drawn between NLM-e and motor theories, and also between NLM-e and the hypothesized 'mirror neurone' system, a neural system that reacts to actions produced by others as identical to the same actions produced by oneself (Rizzolatti *et al.* 1996, 2001; Gallese 2003; Meltzoff & Decety 2003).

The difference is development. We view the connection as developmental in nature, i.e. infants forge a link

between speech perception and production based on perceptual experience and a learned mapping between perception and production (Kuhl & Meltzoff 1982, 1996). On this formulation, sensory learning occurs first, based on experience with language, and this guides the development of motor patterns. Infants' vocal play allows them to relate the auditory results of their own vocalizations to the articulatory movements that caused them, and this creates a learned mapping between the two. Infants strive to imitate the sounds they hear and are guided by the degree of 'match' between the sounds they produce and those stored in memory.

This formulation is similar to models developed for songbirds. Experience with conspecific song is argued to produce an auditory template that subsequently guides vocal production (Phan *et al.* 2006). As in the human imitation of action in infants and adults (Meltzoff & Moore 1997; Jackson *et al.* 2006), and similar to the bird model, we posit that infants store sensory information during the early months of life, when speech production remains primitive and highly variable. Eventually, the perceptual patterns stored in memory serve as guides for production, and this subsequently results in language-specific perception–production mapping.

Data supporting a developmental account stem from behavioural as well as brain studies on infants. Laboratory studies examining infants' capacity for vocal imitation show that listening to simple vowels in the laboratory alters infants' vocalizations, and that this ability emerges at approximately 20 weeks of age but is not present at 12 or 16 weeks of age (Kuhl & Meltzoff 1996). In general, language-specific patterns emerge in speech perception prior to speech production. Infants' vowels produced spontaneously in natural settings do not become language-specific until 10–12 months of age (de Boysson-Bardies 1993), although their perceptual systems show specificity earlier (Kuhl *et al.* 1992; Polka & Werker 1994). The difficult-to-produce /r/ and /l/ sounds will not appear in spontaneous productions until the age of 3 or 4 years (Ferguson *et al.* 1992), but infants in America and Japan show language-specific patterns of perception by 10 months of age (Kuhl *et al.* 2006).

Finally, a new brain study using MEG with newborns, 6- and 12-month-old infants indicates that when listening to syllables, auditory perceptual brain areas (superior temporal) are activated to an equal degree in the three age groups (Imada *et al.* 2006). However, the pure *perception* of speech syllables does not activate brain areas responsible for *production* (the inferior frontal, i.e. Broca's area) in newborns, but does so increasingly in 6- and 12-month-old infants; brain activation of the auditory and motor areas becomes temporally synchronized (see also Dehaene-Lambertz *et al.* 2006). This finding is consistent with the idea that the connection between auditory and motor-speech areas of the human brain requires experience with speech production to bind the two (Imada *et al.* 2006).

(v) *Early speech perception predicts language growth*

NLM-e predicts an association between infants' early perception of native language phonetic units and later language development, an association that differs for native and non-native perception. Retrospective

studies suggested a connection between early speech perception and later language (Molfese & Molfese 1997; Molfese 2000; Newman *et al.* 2006), but prospective studies measuring some aspect of early speech processing in typically developing infants and its relationship to language have appeared only recently (e.g. Fernald *et al.* 2006).

We conducted the first prospective studies investigating the relationship between early phonetic perception and later language. Tsao *et al.* (2004) tested 6-month-old infants' performance on a behavioural measure of speech perception using a simple vowel contrast (the vowels in 'tea' and 'two') and showed that infants' performance measures on the task at six months of age were significantly correlated with their language abilities measured at 13, 16 and 24 months of age. The findings demonstrated that a standard measure of native language speech perception at six months of age prospectively predicted language outcomes in typically developing infants over the next 18 months. As discussed by the authors, Tsao *et al.*'s results could be explained by infants' basic auditory or cognitive abilities. Infants with better auditory skills might perform better in tests of phonetic perception as well as in later language; the same could be argued for infants' cognitive skills—clever infants might respond more readily in the head-turn task and also acquire language more quickly.

To examine this question, Kuhl *et al.* (2005b) tested a more detailed hypothesis that both native and non-native speech perception skills would predict later language, but differentially. Differential effects for native and non-native contrasts would rule out simple auditory or cognitive explanations. The authors used head-turn conditioning to test 7-month-old infants using a Mandarin affricate–fricative (/tʃ–tʃʰ/) contrast and the /p–t/ native place contrast. The findings supported the hypothesis. Both native and non-native performances at seven months of age predicted future language abilities, but in opposite directions. Better native phonetic perception at seven months of age predicted accelerated language development at between 14 and 30 months of age, whereas better non-native performance at the same age predicted slower language development at the same future points in time (Kuhl *et al.* 2005b). The results supported the view that the ability to discriminate non-native phonetic contrasts reflects the degree to which the brain remains in the initial, more immature, state—'open' and uncommitted to native language speech patterns. A language-specific pattern of listening accelerates language growth. The generality of this finding was tested in a new experiment, described in §4a(vi).

(vi) *A new experiment: ERPs to native and non-native contrasts as early predictors of later language*

In this study, infants were tested with one native and two non-native consonant contrasts to examine the generality of the finding that native and non-native phonetic contrasts predict later language, but in opposing directions. ERPs were used in this experiment, which have been shown to provide discriminative responses to phonetic changes in the form of the mismatch negativity (MMN) in adults (Näätänen *et al.* 1997) and an MMN-like response in infants

(Cheour *et al.* 1998; Pang *et al.* 1998; Rivera-Gaxiola *et al.* 2005b, 2007). Electrophysiological measures reduce the potential for cognitive factors, such as attention, to affect discrimination.

### Methods

ERPs were recorded in 30 monolingual full-term infants (14 female) at 7.5 months of age ( $M=7.58$  months, range = 6.84–8.15 months) in response to the native and non-native phonetic contrasts, tested in counterbalanced order. The native contrast (/ta–pa/) varied only in the critical acoustic features (initial formant transitions) that distinguish them for English listeners (Kuhl *et al.* 2005b). A Mandarin Chinese affricate–fricative contrast (/tʃi–tʃʰi/) distinguished by amplitude rise time during the period of frication that has been previously used (Kuhl *et al.* 2003) served as one of the non-native contrasts. A Spanish voicing contrast that is not phonemic in English served as the other non-native contrast. The Spanish stimuli were synthesized versions of the prevoiced and voiceless unaspirated stops /ta–da/. The syllables differed only in their voice onset time (VOT), the primary acoustic cue for the voicing distinction. Total syllable durations were matched for each contrast, as were the fundamental frequency characteristics and overall amplitudes.

Infant ERPs were recorded while infants listened passively to stimuli presented by loudspeakers placed at 45° angles approximately 1 m in front of them; infants sat on a parent's lap while an experimenter entertained them with quiet toys. An oddball paradigm was used with 85% standards and 15% deviants. For the native English contrast, /ta/ served as the standard; for the non-native Mandarin contrast, /tʃi/ served as the standard; for the non-native Spanish contrast, /ta/ served as the standard. In all cases, the interstimulus interval was 700 ms and stimuli were played at 67 dBA.

EEG was collected continuously with a sampling rate of 250 Hz and was bandpass filtered from 0.1 to 60 Hz. EEGs were collected at 16 electrode sites using Electro-caps with standard international 10/20 placements. An additional electrode was placed below the left eye to record eye movements. All sites were referenced to the left mastoid. Data were processed off-line, using epochs of 50 ms pre-stimulus and 800 ms post-stimulus onset. Standards immediately following deviants were excluded from analysis. Trials were hand edited to ensure artefact-free data. Finally, data were filtered using a low-pass filter with a cut-off set at 25 Hz.

Language abilities were assessed at 14, 18, 24 and 30 months of age using the MacArthur–Bates communicative development inventories (CDI), a reliable and valid parent survey for assessing language and communication development from 8 to 30 months of age (Fenson *et al.* 1993). The infant form (CDI: words and gestures) assesses vocabulary comprehension, vocabulary production and gesture production in children ranging from 8 to 16 months of age. The vocabulary production section (checklist of 396 words) was used in this study. The toddler form (CDI: words and sentences) is designed to measure language production in children ranging from 16 to 30 months of age. It assesses vocabulary production using a 680-word checklist and assesses morphological and syntactic

development. Three sections were used in this study: vocabulary production, sentence complexity, and mean length of the longest three utterances (M3L). Parents were asked to complete the CDI on the day their child reached the target age and received \$10 for returning the form.

### Results

Data collected from eight lateral electrode sites (F7/8, F3/4, T3/4, C3/4) were used in the analysis. Participants heard approximately 500 syllables (s.d. = 39) and 60 deviant trials (s.d. = 14) for each contrast. Mean amplitude of the deviant minus standard difference wave (infant MMN) between 300 and 500 ms after the onset of the deviant was measured. Usable ERP data were obtained from 24 of the 30 participants for the native contrast and 22 of the 30 participants for the non-native contrast (15 to Mandarin and 7 to Spanish). Out of the 30 participants, 21 had acceptable ERP data for both the native and non-native contrasts (15 with the Mandarin non-native contrast and 6 with the Spanish non-native contrast).

Data analysis proceeded in three steps. First, average waveforms for the standards and deviants obtained for the native and non-native contrasts at the eight electrode sites were analysed for each child, and the mean amplitude of the mismatch negativity (MMN; Näätänen *et al.* 1997) was calculated for each site. The MMN is a difference wave, calculated by subtracting the average waveforms in response to the standard and deviant stimuli. Adults respond to a deviant stimulus with a negative wave that is observed at approximately 250 ms. The infant response is slightly later at approximately 300–500 ms (Cheour *et al.* 1998; Rivera-Gaxiola *et al.* 2005b).<sup>1</sup> Better discrimination is indicated by larger amplitudes of the negativity, which can be measured either as a peak value or a mean amplitude value (Kraus *et al.* 1996). Separate repeated measures ANOVAs, conducted for each contrast (native, Spanish and Mandarin), indicated no interactions of stimulus (standard versus deviant) by hemisphere (left versus right) by site ( $N=4$  sites) for the native ( $F(3,69)=0.264$ ,  $p=0.851$ ,  $n=24$ ); the Mandarin ( $F(3,42)=0.505$ ,  $p=0.649$ ,  $n=15$ ); or the Spanish contrast ( $F(3,18)=1.309$ ,  $p=0.305$ ,  $n=7$ ). Based on the broad distributions of the MMNs, a single MMN mean amplitude difference value (deviant – standard at 300–500 ms) was calculated for the native and non-native languages for each infant by averaging values across hemisphere and electrode site.

We observed a significant negative correlation between an infant's ERPs to the native and the non-native contrasts; this was the case whether the Mandarin non-native ( $r=-0.584$ ,  $p=0.011$ ,  $n=15$ ) or the Spanish non-native ( $r=-0.741$ ,  $p=0.046$ ,  $n=6$ ) contrast was tested (combined  $r=-0.631$ ,  $p=0.001$ ,  $n=21$ ). Infants showing more negative MMN effects (indicating greater discrimination at the neural level) for the native /p–t/ contrast showed less negative effects (i.e. lower discrimination) for the non-native contrast (either Mandarin or Spanish), and those with better non-native abilities showed comparably poorer native abilities. This relationship replicates and extends to the Spanish non-native contrast the

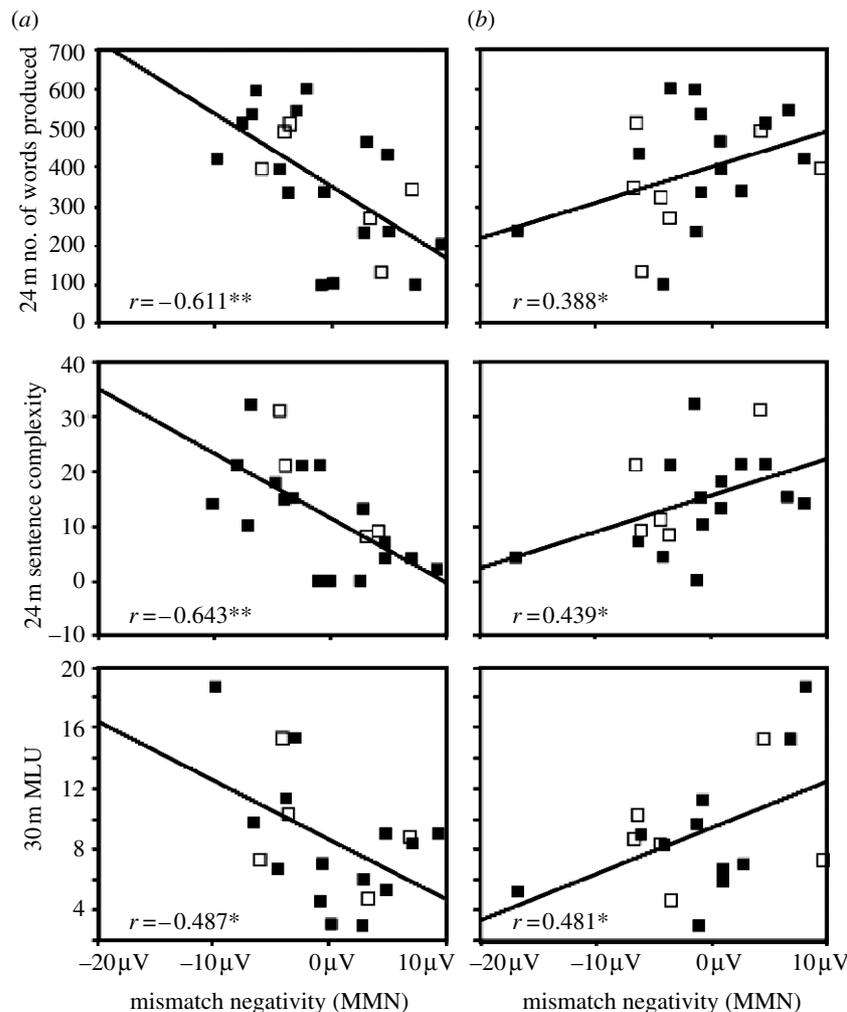


Figure 3. Scatter plots show significant correlations between infants' phonetic discrimination measures at 7.5 months for (a) native as opposed to (b) non-native phonetic contrasts and their later language abilities. Better performance on speech discrimination, as measured by the MMN, is indicated by a more negative value, producing a negative correlation between native speech discrimination and later language and a positive correlation between non-native speech discrimination and later language (closed square, Mandarin; open square, Spanish).

findings of our previous behavioural study (Kuhl *et al.* 2005b).

Infants' MMN values for the native and non-native contrasts were related to their CDI measures of language ability at 14, 18, 24 and 30 months of age. As predicted, both native and non-native neural measures predicted future language, but in opposing directions. The native language MMN at 7.5 months of age predicted the number of words produced at 18 months of age ( $r = -0.430$ ,  $p = 0.020$ ) and the number of words produced at 24 months of age ( $r = -0.611$ ,  $p = 0.001$ ) with greater negativity of the MMN associated with a larger number of words produced (figure 3a). More negative MMNs to native language sound contrasts also predicted higher sentence complexity at 24 months of age ( $r = -0.643$ ,  $p = 0.001$ ) and a longer M3L at 24 ( $r = -0.632$ ,  $p = 0.001$ ) and 30 months of age ( $r = -0.487$ ,  $p = 0.017$ ).

A very different pattern of prediction was observed when infants' MMN measures of non-native perception were used to predict future language skills (figure 3b). More negative MMNs to non-native phonetic contrasts at 7.5 months of age predicted fewer words produced at 24 months of age ( $r = 0.388$ ,

$p = 0.041$ ), lower sentence complexity at 24 months of age ( $r = 0.439$ ,  $p = 0.030$ ) and a shorter M3L at 30 months of age ( $r = 0.481$ ,  $p = 0.025$ ). This pattern of correlations replicates the findings of our previous behavioural study (Kuhl *et al.* 2005b).

The rate of language growth over time can be assessed using the number of words produced at each of the four ages studied. Acceleration in expressive vocabulary growth during the second year characterizes learning in many of the world's languages (Huttenlocher *et al.* 1991; Fenson *et al.* 1994; Bornstein & Cote 2005). To examine whether brain responses to speech sounds at 7.5 months of age predicted rates of expressive vocabulary development from 14 to 30 months of age, we used the hierarchical linear models programme (Raudenbush *et al.* 2005). In multi-level modelling, repeated measurements of vocabulary size are used to estimate growth functions for each child, and the resultant growth parameters for each individual are modelled as random with variance predicted by a between-subjects variable. Inspection of individual children's data indicated that variation across children was observed from 14 to 30 months of age. Of the 23 children for whom at least three data

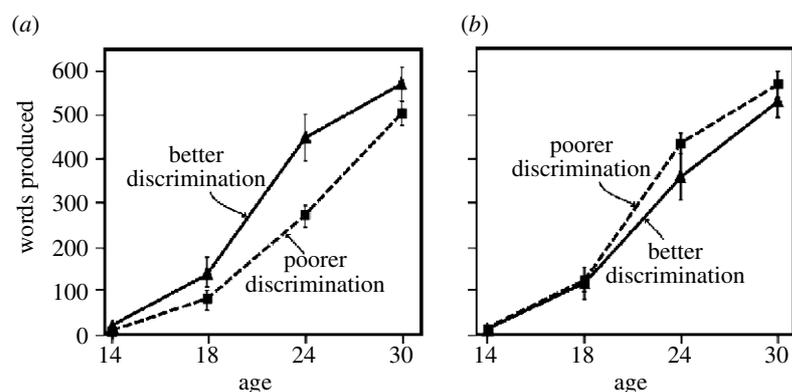


Figure 4. A median split of infants whose MMNs indicate better versus poorer discrimination of (a) native and (b) non-native phonetic contrasts is shown along with their corresponding longitudinal growth curve functions for the number of words produced between 14 and 30 months of age.

points were available, approximately half ( $n=12$ ) showed rapid initial growth, reaching close to 400 words or more by 24 months of age (range, 393–597). From 24 to 30 months of age, the slopes were flatter in these children, which may be at least partially an artefact of the vocabulary measure (their 30-month vocabulary sizes ranged from 555 to 673 and the CDI ceiling was 680 words). The remaining children evidenced lesser gains in vocabulary size up to 24 months of age, although their scores were still within the normal range at 24 (97–343 words) and 30 months of age (207–678 words).

Separate analyses were conducted for the children for whom we had obtained artefact-free native- and non-native-contrast ERP data at 7.5 months of age ( $n=24$  and 22, respectively, with 21 children participating in both analyses). At the first level of each analysis, we estimated individual growth curves for each child using a quadratic equation with the intercept centred at 18 months (due to the small sample sizes, we used restricted maximum likelihood). Several reports on expressive vocabulary development in this age range have indicated that quadratic models capture typical growth patterns, both a steady increase and acceleration (Huttenlocher *et al.* 1991; Ganger & Brent 2004; Fernald *et al.* 2006). Centring at 18 months allowed us to evaluate individual differences in vocabulary size at an age that has previously been associated with rapid growth ('vocabulary spurt') as well as differences in rates of growth across the whole period.

For each sample of children, unconditional models indicated individual variation in the random effects for the intercept (18-month vocabulary size), linear slope and quadratic component. Covariance estimates indicated high degrees of collinearity between the linear and quadratic components for each sample. For both the native and non-native MMN samples, the intercepts and linear slopes were highly positively correlated at 0.99, indicating that children with higher 18-month vocabulary sizes tended to have faster growth throughout the 14–30 months period. For both samples, the intercepts and quadratic components were highly negatively correlated (native  $\tau=-0.95$ ; non-native  $\tau=-0.99$ ) and the slopes and quadratic components were highly negatively correlated (native  $\tau=-0.89$ ; non-native  $\tau=-0.97$ ), which likely reflects the fact that children whose vocabulary sizes reached higher

levels by 18 months and had overall faster growth had a levelling-off function towards 30 months of age as they approached the CDI ceiling.

At the second level of analysis, child-level variations in the intercepts (i.e. 18-month vocabulary sizes) and in the slopes of the growth functions were modelled as a function of each of the 7.5-month MMN values. The quadratic growth curve model indicated that the average native contrast MMN was significantly related to the intercept (18-month vocabulary size;  $t(22)=-4.15$ ,  $p<0.001$ ) the linear slope component ( $t(22)=-4.07$ ,  $p<0.001$ ) and the quadratic component of the growth function ( $t(22)=3.32$ ,  $p=0.003$ ). To illustrate this relationship, figure 4a shows the growth patterns for children whose 7.5-month native MMNs were below and above the median. The children with 7.5-month MMN values that were more negative (indicating better discrimination) showed significantly faster initial vocabulary growth with a later levelling-off function (possibly due to a CDI ceiling effect). In contrast, children with 7.5-month native MMN values that were less negative (poorer discrimination) showed less rapid growth in the number of words.

The opposite pattern was obtained for the non-native language contrast (figure 4b). Children with 7.5-month non-native MMNs that were more negative (indicating better discrimination) showed significantly slower growth in the number of words produced, while those with less negative non-native MMN values showed faster vocabulary growth. The quadratic growth curve model showed that the average non-native contrast MMN values were significantly related to the intercept ( $t(20)=2.27$ ,  $p=0.03$ ) and the linear slope component ( $t(20)=2.63$ ,  $p=0.02$ ). There was a trend for the interaction between non-native MMN size and the quadratic component of the growth function ( $t(20)=-1.97$ ,  $p=0.06$ ). Thus, greater discrimination of the non-native contrast at 7.5 months of age was associated with slower vocabulary growth. In contrast, infants with non-native MMN values indicating poorer discrimination showed more rapid growth in vocabulary size.

#### Discussion

Infants' performance on both the native and non-native phonetic discrimination tasks, measured using ERPs at 7.5 months of age, significantly predict children's language abilities 2 years later, but differentially. Better

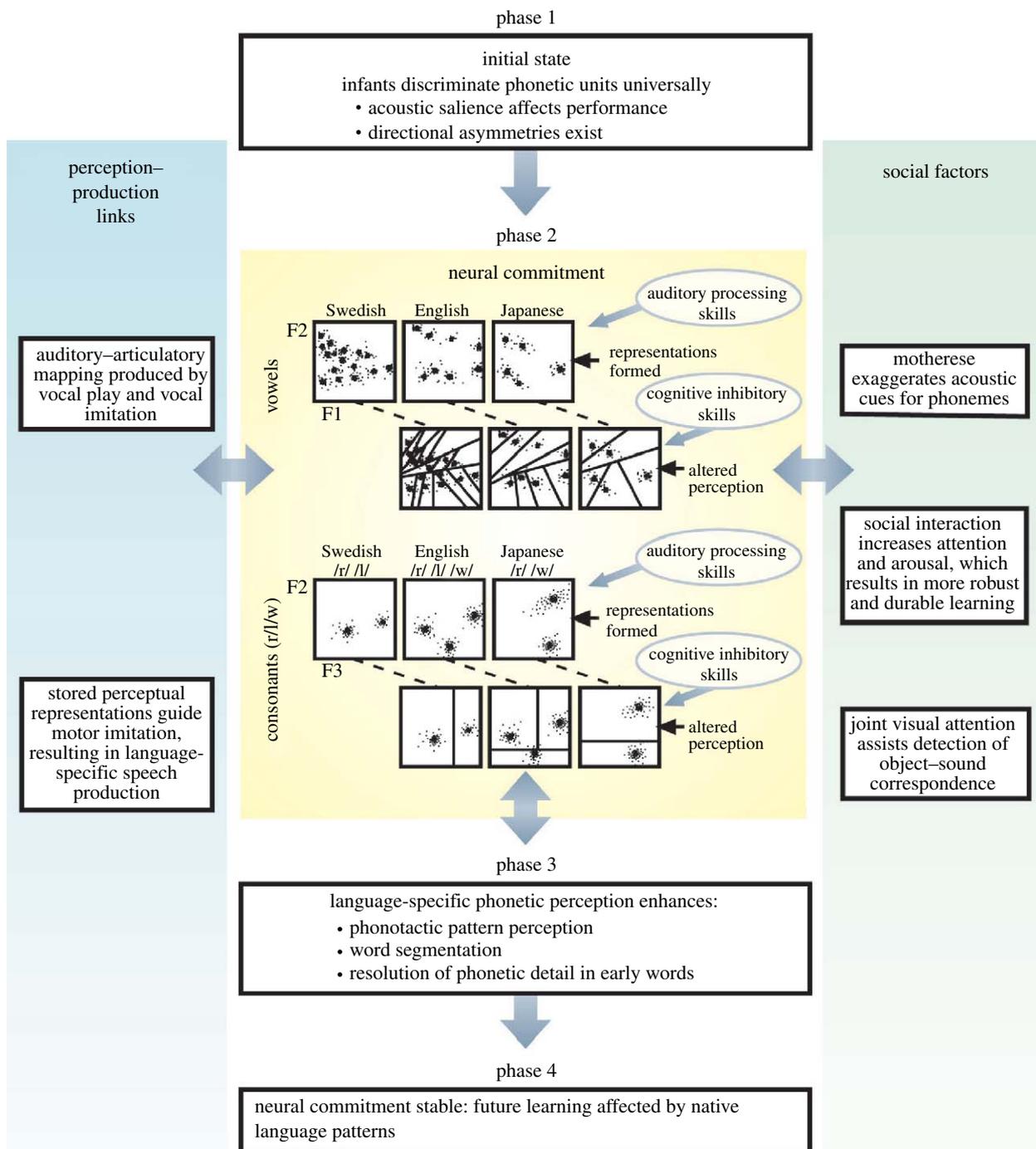


Figure 5. NLM-e is shown in four phases (see text for description). The representations of native language input for vowels and consonants are drawn roughly to reflect existing data for Swedish (Fant 1973; Lacerda in preparation), English (Dalston 1975; Flege et al. 1995; Hillenbrand et al. 1995) and Japanese (Iverson et al. 2003; Lotto et al. 2004).

native phonetic abilities predict faster advancement in language, whereas better non-native phonetic abilities predict slower linguistic advancement. Infants' early phonetic perception predicts language at many levels, including the number of words produced, the degree of sentence complexity and the mean length of utterance. The present data show that these predictive relations, observed previously in our behavioural tests (Kuhl et al. 2005b), generalize to ERP measures, and importantly, to a new non-native contrast.

Additional studies from our laboratory show a similar pattern of prediction for native and non-native phonetic abilities and later language. Rivera-Gaxiola et al. (2005a) measured ERPs in response to native and

non-native English and Spanish contrasts in 11-month-old infants and measured their language skills with the CDI at 18, 22, 25, 27 and 30 months of age. Infants were categorized into two groups depending on the latency and polarity of their non-native contrast responses; infants with prominent negativities between 250 and 600 ms after stimulus onset (good discrimination) at 11 months produced significantly fewer words at each age than infants who showed less negative responses to the non-native contrast.

Using the stimuli of Rivera-Gaxiola and colleagues (Rivera-Gaxiola et al. 2005a,b) and a new double-target behavioural measure to relate concurrent language abilities and speech perception in 11-month-old infants,

Conboy *et al.* (2005) showed that the degree to which infants' d' scores to native contrasts exceeded their performance on non-native contrasts predicted the number of words they had comprehended at that age.

Finally, a study of Finnish 7- and 11-month-old infants replicated this pattern (Silven *et al.* 2004). Monolingual infants tested on a native Finnish and a non-native Russian contrast at the two ages, and followed-up with the Finnish version of the CDI at 14 months of age, showed a significant positive correlation between native language scores at seven months of age and future language measures, and a significant negative correlation between non-native perception at 11 months of age and future language measures.

Thus, a number of studies of typically developing 7- and 11-month-old infants, ones using different phonetic contrasts in different countries, show consistent findings. Better native language speech perception skill, measured either behaviourally or neurally, predicts more rapid acquisition of language, whereas better non-native phonetic skill, measured in the same way at the same age, predicts slower language growth. It is important to note that these results are for monolingual infants who have not had any experience with the non-native language being tested. The pattern expected for bilingual infants should differ and is discussed in a later section of this paper (see §5a).

Taken together, the results support the argument that phonetic learning predicts the rate of language acquisition over the first 30 months of life, and that this result does not rely on general auditory or cognitive skills, or even on infants' abilities to track the kinds of acoustic cues characteristic of all phonemes. Rather, infants' abilities to *learn phonetically* from exposure to language predict the rate of language growth over the first 30 months of age.

We do not intend, however, to ascribe no role to basic auditory abilities in language acquisition; in fact, rapid auditory processing abilities early in development are predictive of language disabilities later (Benasich & Leavers 2002; Benasich & Tallal 2002), and the NLM-e model describes a specific role for the ability to resolve differences auditorially.

Similarly, NLM-e describes a specific role for cognitive skills. We know that cognitive abilities are linked to various aspects of communicative development (Tomasello & Farrar 1984; Bates & Snyder 1987; Gopnik & Meltzoff 1987; Thal 1991). Specific cognitive abilities, ones that tap attentional and/or inhibitory control, are related to performance on non-native speech perception tasks at the end of the first year (Diamond *et al.* 1994; Lalonde & Werker 1995; Conboy *et al.* submitted), and the NLM-e model depicts this specific relationship. Given the evidence that infants' capacity to discriminate non-native contrasts declines but remains above chance at the end of the first year (Rivera-Gaxiola *et al.* 2005b; Kuhl *et al.* 2006; Tsao *et al.* 2006), additional explanations are needed to account for how infants refrain from responding to these contrasts; cognitive control may provide an explanation. Bilingual children, whose language environments require them to 'switch' between languages, develop certain attentional control processes to a greater extent than monolingual children

(Bialystok 1999). Early speech perception may be one mechanism through which this 'bilingual advantage' emerges (Conboy *et al.* submitted).

Social factors, such as joint visual attention, also predict aspects of language, such as the number of words produced at 18 months of age (Tomasello & Farrar 1986; Baldwin & Markman 1989; Baldwin 1995; Brooks & Meltzoff 2005). It remains for future research to measure simultaneously these various skills in infants—basic auditory, cognitive and social skills, the ability to detect distributional patterns and phonetic learning—in a longitudinal study that examines the individual and joint effects of these factors on language development and brain development. Multiple factors are expected to play a role in language acquisition (Hollich *et al.* 2000). NLM-e takes multiple factors into account and makes predictions about the nature of their roles in the early developmental period. Additional data are needed to flesh out the intricate way in which multiple factors affect early language development.

Our claim is that the pattern we have observed between native and non-native speech perceptions—with native and non-native abilities predicting future language in opposite directions—requires a multi-factor explanation. We argue that the phonetic learning process relies on the distributional patterns in ambient language and the exaggerated cues provided by ID speech, i.e. infants' experience with these patterns produces neural commitment. Social factors play a vital role in this learning process in natural settings. Auditory abilities affect this process: if infants cannot resolve the basic differences between phonetic units at the beginning of life, native language phonetic learning will be reduced. Domain-general cognitive control abilities also play a role in infants' relative attention to native- and non-native language phonetic differences and in their suppression of non-native differences. All factors will be necessary to explain the patterns observed in developmental speech perception. In §4b, we describe the phases of the new model in detail.

### (b) *Description of NLM-e*

NLM-e is schematically described in figure 5.

#### (i) *NLM-e: phase 1*

Phase 1 of the model shows that early in life infants discriminate all phonetic units in the world's languages. Additional factors that explain the degree to which performance varies across phonetic contrasts are noted. For example, studies demonstrate that the acoustic salience of a phonetic contrast affects performance; fricatives, for example, have been shown to be more difficult to discriminate due to their low amplitude (Eilers *et al.* 1977; Kuhl 1980; Nitttrouer 2001). Burnham (1986) argues a theoretical position based on salience. Moreover, studies show that discrimination performance in infants and young children is above chance but far below that shown by adult native listeners (Nitttrouer 2001; Kuhl *et al.* 2006; Sundara *et al.* 2006). Infants' initial performance thus leaves room for substantial improvement, especially for those contrasts that are acoustically fragile. The model stipulates that, in phase 1, infants' phonetic abilities

are relatively crude, reflecting general auditory constraints and/or learning *in utero* (Moon *et al.* 1993). Testing infants' discrimination abilities for a greater variety of phonetic contrasts in the newborn period would be informative for theory.

In addition, phase 1 of the model recognizes that directional asymmetries, shown when a change in one direction results in significantly better performance than a change in the other direction, are observed. Directional effects across age and culture have been shown for vowels (Polka & Bohn 1996, 2003) as well as consonants (Kuhl *et al.* 2006), indicating that when these effects are observed, they are seen across cultures and age, at least in infancy. The origins of directional effects remain unclear, Polka & Bohn (2003) discuss the alternatives, and could either reflect general psychoacoustic factors or factors that reflect language-specific constraints (Miller & Eimas 1996).

In sum, the critical feature of the initial state stipulated by the model is that infants begin life with a capacity to discriminate the acoustic cues that code differences among phonetic units. The ability to discriminate the sounds, albeit crudely, assists development in phase 2.

#### (ii) *NLM-e: phase 2*

Phase 2 represents the core of the NLM-e model. At this stage in development, infants' sensitivity to the distributional patterns (Kuhl *et al.* 1992; Maye *et al.* 2002) and exaggerated cues of ID speech (Liu *et al.* 2003) cause phonetic learning. As depicted, learning occurs earlier for vowels than consonants (Werker & Tees 1984a; Kuhl *et al.* 1992; Polka & Werker 1994; Best & McRoberts 2003). This difference could reflect the availability of exaggerated cues in ID speech (consonants are not as easily exaggerated as vowels, because exaggeration can change the category, e.g. stretching the formant transitions of /b/ produces /w/). Alternatively, there may be differences in the availability and/or prominence of distributional differences for consonants (e.g. consonants like /th/ occur in function words, which are lower in energy and do not capture infant attention, see Sundara *et al.* 2006). Understanding how these two aspects of environmental input—exaggerated acoustics and distributional properties—interact to support infants' perception of categories will be important for future studies. In general, the model holds that the timing of perceptual change for various phonetic contrasts will vary depending on the availability of information about the contrast in language input.

In phase 2, NLM-e shows social interaction as playing a facilitative role in learning. Social interaction enhances phonetic learning as infants become more skilled at social understanding (Bruner 1983; Baldwin 1995; Tomasello 2003). Future studies will be required to determine whether the mechanism by which social interaction affects learning is the increased attention and arousal that occurs during social interaction, or whether the specific information provided during social interaction (such as joint visual attention to an object), or both, are responsible for the facilitative effect social interaction has on language learning. Either a general 'motivational' explanation involving attention or

arousal or a more specific 'informational' explanation could account for the effects of social interaction on learning, and both are likely to play a role (Kuhl *et al.* 2003). In complex natural communicative settings, social interaction may serve to 'gate' computational learning (Kuhl 2007). The greater complexity and naturalness of the language learning setting may make it more probable that social interaction will play a significant role.

Finally, NLM-e indicates a link to speech production that is forged during this period. Infants develop connections between speech production and the auditory signals it causes during development as they practice and play with vocalizations, and imitate those they hear. As speech production improves, imitation of the learned patterns stored in memory leads to language-specific speech production. It has been suggested that speech production itself plays a role by encouraging the use of learned motor patterns (DePaolis 2005), and NLM-e depicts bi-directional effects between perception and production in phase 2 as the connection between them is formed.

By the end of phase 2, infant perception is altered. The detection of native language phonetic cues is enhanced in the process, while detection of non-native-phonetic patterns is reduced. At this stage, infant perception has been warped by experience and begins to reflect attunement between infant perception and the language and culture in which the infants are being raised.

#### (iii) *NLM-e: phase 3*

In phase 3, enhanced speech perception abilities improve three independent skills that propel infants towards word acquisition: the detection of phonotactic patterns (Friederici & Wessels 1993; Mattys *et al.* 1999); the detection of transitional probabilities between segments and syllables (Goodsitt *et al.* 1993; Saffran *et al.* 1996; Newport & Aslin 2004); and the association between sound patterns and objects (Swingley & Aslin 2002; Werker *et al.* 2002; Ballem & Plunkett 2005). Each of these skills—detection of phonotactic patterns, detection of word-like units and the resolution of phonetic detail in early words—is likely to predict future language, though empirical studies have just begun to test these relationships (Newman *et al.* 2006). Bidirectional effects are indicated at this stage in that phonetic learning would assist the detection of word patterns, and the learning of phonetically close words would be expected to sharpen the awareness of phonetic distinctions.

#### (iv) *NLM-e: phase 4*

By phase 4, analysis of incoming language has produced relatively stable neural representations—new utterances do not cause shifts in the distributional properties coded neurally. In infancy, neural networks are not completely formed and do not restrict learning. Infants are thus capable of learning from multiple languages, as shown in everyday life, and also as shown by experimental interventions (Maye *et al.* 2002; Kuhl *et al.* 2003). In adults, representations are stable and are relatively unaffected by short periods of listening to a new language. Thus, exposure to a new language does not automatically create new neural structure.

The principle underlying the model is that the degree of ‘plasticity’ in learning the phonetics of a second language depends on the stability of the underlying perceptual representations, and therefore on the degree of neural commitment.

## 5. PREDICTIONS OF THE NLM-E MODEL

### (a) *Predictions regarding the effects of bilingual language experience*

What does the NLM-e model predict in the case of infants raised with bilingual exposure? NLM-e describes phonetic development in bilinguals as following the same principles for two languages as it does for one. Bilingual infants learn through the exaggerated acoustic cues provided by ID speech and through the distributional properties of the two languages, as do monolingual infants. It is not yet clear whether bilingual infants form two distinct representations, one for each language, in the early period. Phonetic exaggeration and the distributional properties of the two languages would differ, and these properties could provide infants with a means of separating the two streams of input. If infants do, representations for each language would be expected to follow the path described by NLM-e; in short, monolingual and bilingual infants learn in the same way.

Bilingual language experience could potentially have an impact, according to the model—the development of representations in phase 2 could require a longer period of time than for the monolingual case. Infants learning two first languages simultaneously might reach the developmental change in perception at a later point in development than infants learning either language monolingually (see Bosch & Sebastián-Gallés 2003a,b). Bilingual infants could remain in phase 2 for a longer period of time because it takes longer for sufficient data from both languages to be experienced and to reach sufficient stability; this could depend on factors such as the number of people in the infants’ environment producing the two languages in speech directed towards the child and the amount of input they provide. These factors could change the rate of development in bilingual infants.

Since neural commitment in the early period is incomplete, a second language introduced during infancy does not encounter as much ‘interference’ from commitment to the features of the first language as a second language introduced at a later point in life. The phonetic features of each language could be mapped onto separate perceptual spaces because their acoustic and statistical properties are sufficiently distinct. We do not know how much language input is required from two languages to produce this purported *dual mapping* in bilingual infants; the foreign language intervention experiment (Kuhl *et al.* 2003) required only 12 sessions to produce learning, and that learning was shown to be durable, but it would nonetheless be expected to show a ‘forgetting function’ (see below). We have no data to indicate how much exposure is necessary to produce long-term phonetic learning.

As in the case of monolingual exposure, social factors would be expected to play a role in bilingual learning, and could in fact be argued to assist learning. In some cases of simultaneous bilingualism, different

people speak the two languages to which the infant is exposed. If the social settings in which exposure to the two languages occurs also differ, greater separation of the two inputs would be achieved. At present, there is no evidence of an advantage for such ‘one person, one language’ approaches to bilingual language socialization over approaches in which infants hear both languages from the same people, or situations in which parents frequently code-switch between languages; this is clearly a matter for future research. Code-switching and mixing are common practices in many bilingual communities, and it has been shown that a strict separation of languages is difficult for many families (Goodz 1989). Even in mixed language situations, ID speech could exaggerate different aspects of the two languages, assisting infants’ mapping of features that are relevant for each of the two languages.

Predicting future language from infants’ early speech perception should also apply to bilingual infants, though to see the pattern of predictive correlations that we observed, a third language, to which the infants have not been exposed, would have to be tested. Phonetic contrasts from both languages to which bilingual infants are exposed should correlate positively with later language; a third language, to which infants are not exposed, would be expected to show the opposite pattern.

There is little data on speech perception in infants exposed to two languages simultaneously early in development. Some studies suggest that infants exposed to two languages show later acquisition of language-specific phonetic skills when compared with monolingual infants (Bosch & Sebastián-Gallés 2003a,b; but see Burns *et al.* 2007). This is especially the case when infants are tested on contrasts that are phonemic in only one of the two languages; this has been shown both for vowels (Bosch & Sebastián-Gallés 2003a) and consonants (Bosch & Sebastián-Gallés 2003b). A preliminary report of phonetic perception in bilingual English–Spanish learners tested at 6–8 and 10–12 months of age using a brain measure (ERPs) indicated that bilingual infants showed robust MMN-like responses to both English and Spanish contrasts (Rivera-Gaxiola & Romo 2006), which distinguished them from their English monolingual peers tested at the same ages with the same stimuli, who showed much stronger MMN-like responses to the English as opposed to the Spanish voicing contrast (Rivera-Gaxiola *et al.* 2005b). Additional data will be necessary to understand bilingual phonetic development.

Several studies have noted variations in the voice onset time of stop consonants produced by bilingual adults compared with monolingual speakers of the same languages (Flege 1988; MacLeod & Stoel-Gammon 2005). Thus, a monolingual reference may be inappropriate for bilingualism. Infants raised with two (or more) languages should not necessarily be expected to resemble monolingual learners of each of their languages; they may develop perceptual, cognitive and linguistic systems that are unique responses to the conditions and demands of their bilingual input. Given that bilingual infants are required to alternate attention to different linguistic features in their everyday speech processing, the cognitive component of the model

also predicts that attentional or inhibitory cognitive processes needed for such perceptual switching would be enhanced in bilingual infants (Bialystok 2001; Conboy & Mills 2006).

**(b) Predictions on the durability and robustness of learning**

The NLM-e model predicts that social interaction produces learning in natural settings which is more robust and durable; in other words, we suggest that learning in social settings is in some sense more potent and enduring. There are two reasons to suggest that social factors affect learning in this way. First, our own data suggest some degree of durability; infants in the Mandarin exposure studies (Kuhl *et al.* 2003) returned to the laboratory between 2 and 12 days after the final exposure session to complete their behavioural Mandarin discrimination tests. Analysis showed that the delay had no effect on the infants' performance. Moreover, data on these infants' ERP responses to the Mandarin contrast, gathered at even greater delays between the last exposure session and the test session (infants returned to the laboratory between 8 and 33 days after the last exposure session, with a median of 15 days) also indicated no effect of the delay (Kuhl *et al.* in preparation). Infants in the exposure experiment would nonetheless be expected to show a forgetting function eventually because 5 hours of listening experience would not be sufficient to undo the representations built up over the previous nine months of life. The memory of the early one month experience of Mandarin could, however, prompt more rapid learning later in life than would be the case if never exposed to Mandarin. Neural modelers suggest that short-term learning of new phonetic contrasts is initially perceptually separated, and therefore produces learning without undoing the representations formed by long-term listening to one's primary language (Vallabha & McClelland 2007).

Second, adopting the neurobiological framework, song learning in birds also indicates that social interaction extends the period of learning and produces learning that is more robust and durable. Richer social environments extend the duration of the sensitive period for learning in owls and songbirds (Baptista & Petrinovich 1986; Brainard & Knudsen 1998). Social contexts affect the rate, quality and retention of song elements in songbirds' repertoires (West & King 1988). The idea that social interaction affects learning in this way can be experimentally assessed by systematically measuring the forgetting function under conditions in which input complexity (conversational language from multiple talkers versus syllable presentations in the laboratory), as well as the social factors that the learning paradigm incorporates, are manipulated.

**(c) Predictions regarding the mechanism underlying the critical period**

Language and the critical period have long been associated and many language scientists have discussed the issue (Lenneberg 1967; Johnson & Newport 1989; Bialystok & Hakuta 1994; Flege *et al.*

1999; Weber-Fox & Neville 1999; Yeni-Komshian *et al.* 2000; Birdsong & Molis 2001; Newport *et al.* 2001; Werker & Tees 2005). As described in recent publications, our recent studies provide evidence concerning the mechanisms underlying a critical period at the phonetic level for language (Kuhl *et al.* 2005b). According to the model, phonetic learning causes a decline in neural flexibility, suggesting that *experience*, not simply *time*, is a critical factor driving phonetic learning and perception of a second language.

Bruer (in press) recently discussed the need to separate studies that focus on identifying the phenomena and optimum periods of learning in various domains (see also Lorenz 1957; Hess 1973; Bateson & Hinde 1987) from experimental tests that explore the explanatory causal mechanism that underlies a critical period for language. Thus far, our work has focused only on the mechanism question; we have not varied the timing of foreign language experience to identify the periods during which infants are most sensitive to a new language. The mechanism in question requires a different kind of experiment, one that differentiates the role of maturation from that of experience. Both the maturational view (Lenneberg 1967; Johnson & Newport 1989; Bialystok & Hakuta 1994; Flege *et al.* 1999; Weber-Fox & Neville 1999; Yeni-Komshian *et al.* 2000; Birdsong & Molis 2001; Newport *et al.* 2001) and the experience/interference view (Kuhl 1998, 2000a,b, 2004; Iverson *et al.* 2003; Seidenberg & Zevin in press) are supported by experimental data on first and second language learning. NLM-e highlights the role of neural commitment as a potential mechanistic cause of the critical period phenomenon. The data shown in the present study indicates that at the cusp of learning, phonetic perception of native and non-native contrasts is negatively correlated. This supports the idea that learning itself may play a role in reducing the future capacity to learn new phonetic patterns.

In most species, particular events open and close the critical period during which sensitivity to environmental input is increased. What opens and closes the period of optimal sensitivity to phonetic cues according to NLM-e? A variety of factors suggest that initial phonetic learning could be triggered on a maturational timetable, between 6 and 12 months of age. It is during this period that infants show an increase in native language speech perception (Rivera-Gaxiola *et al.* 2005b; Kuhl *et al.* 2006), a decline in non-native perception (Werker & Tees 1984a; Best & McRoberts 2003) and readily learn phonetically when exposed to new phonetic patterns for the first time (Maye *et al.* 2002; Kuhl *et al.* 2003). Studies of the maturation of the human auditory cortex show that between the middle of the first year of life and 3 years of age, there is maturation of axons entering the deeper cortical layers from the subcortical white matter; and neurofilament-expressing axons appear for the first time in the temporal lobe, with projections to the deep cortical layers of the brain. These axons would provide the first highly processed auditory input from the brainstem to higher auditory cortical areas (Moore & Guan 2001). The temporal coincidence between this cytoarchitectural change and infants' phonetic learning provides a

possible maturational cause of the ‘opening’ of a critical period for phonetic learning.

If maturation opens the critical period, what ‘closes’ the period of optimum sensitivity for phonetic learning? According to NLM-e, learning continues until stability is achieved. It has been argued elsewhere (Kuhl 2000a,b, 2004) that the closing of the critical period may be a statistical process whereby the underlying networks continue to change until the amount and variability of acoustic cues for phonetic categories reach stability. Neural networks stay flexible and continue to ‘learn’ until the number and variability of occurrences of a particular phonetic unit produce a distribution that predicts new instances of that unit and do not significantly shift the underlying distribution. Computational neural modelling experiments have produced findings that are consistent with this view (Vallabha & McClelland 2007).

Neural readiness for phonetic learning in humans is, of course, not well understood. Animal data indicate that architectural shifts change the patterns of conductivity among circuits during learning (Knudsen 2004). Regarding language, exposure to spoken (or signed, see Pettito *et al.* 2004) language during a critical period may be enabled by maturation, which instigates the mapping process described by NLM-e during which the brain’s circuits are altered. NLM-e offers an encompassing view of the multiple factors that play a role in infants’ early phonetic learning and provides a framework that offers specific hypotheses that are amenable to empirical investigation.

Funding for the research was provided by a National Science Foundation Science of Learning Center grant to the University of Washington’s LIFE Center (0354453) and by a grant to P.K.K. from the National Institutes of Health (HD37954). This work was facilitated by P30 HD02274 from the National Institute of Child Health and Human Development and an NIH UW Research Core Grant, University of Washington P30 DC04661. The authors thank Andrew Meltzoff and four anonymous reviewers for their very helpful comments on an earlier draft and Jeff Munson for his assistance in the hierarchical linear models analysis.

## ENDNOTE

<sup>1</sup>In infants, both positive and negative waves can occur in response to a syllable or tone change (see Pang *et al.* 1998; Morr *et al.* 2002; Martin *et al.* 2003; Rivera-Gaxiola *et al.* 2005a, 2007) and may indicate different levels of discrimination.

## REFERENCES

- Baldwin, D. A. 1995 Understanding the link between joint attention and language. In *Joint attention: its origins and role in development* (eds C. Moore & P. J. Dunham), pp. 131–158. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Baldwin, D. A. & Markman, E. M. 1989 Establishing word-object relations: a first step. *Child Dev.* **60**, 381–398.
- Ballem, K. D. & Plunkett, K. 2005 Phonological specificity in children at 1; 2. *J. Child Lang.* **32**, 159–173. (doi:10.1017/S0305000904006567)
- Baptista, L. F. & Petrinovich, L. 1986 Song development in the white-crowned sparrow: social factors and sex differences. *Anim. Behav.* **34**, 1359–1371. (doi:10.1016/S0003-3472(86)80207-X)
- Baran, J. A., Laufer, M. Z. & Daniloff, R. 1977 Phonological contrastivity in conversation: a comparative study of voice onset time. *J. Phonet.* **5**, 339–350.
- Bates, E. & Snyder, L. 1987 The cognitive hypothesis in language development. In *Research with scales of psychological development in infancy* (eds I. Uzgiris & J. M. Hunt), pp. 168–206. Champaign, IL: University of Illinois Press.
- Bateson, P. & Hinde, R. A. 1987 Developmental changes in sensitivity to experience. In *Sensitive periods in development* (eds M. H. Bornstein), pp. 19–34. Hillsdale, NJ: Erlbaum.
- Benasich, A. A. & Leevers, H. J. 2002 Processing of rapidly presented auditory cues in infancy: implications for later language development. In *Progress in infancy research*, vol. 3 (eds J. Fagan & H. Haynes), pp. 245–288. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Benasich, A. A. & Tallal, P. 2002 Infant discrimination of rapid auditory cues predicts later language impairment. *Behav. Brain Res.* **136**, 31–49. (doi:10.1016/S0166-4328(02)00098-0)
- Bernstein-Ratner, N. 1984 Patterns of vowel modification in mother–child speech. *J. Child Lang.* **11**, 557–578.
- Best, C. T. 1994 The emergence of native-language phonological influences in infants: a perceptual assimilation hypothesis. In *The development of speech perception: the transition from speech sounds to spoken words* (eds J. C. Goodman & H. C. Nusbaum), pp. 167–224. Cambridge, MA: MIT Press.
- Best, C. T. 1995 A direct realist view of cross-language speech perception. In *Speech perception and linguistic experience* (ed. W. Strange), pp. 171–206. Baltimore, MD: York Press.
- Best, C. T. & McRoberts, G. W. 2003 Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Lang. Speech* **46**, 183–216.
- Best, C. T., McRoberts, G. W. & Goodell, E. 2001 Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener’s native phonological system. *J. Acoust. Soc. Am.* **109**, 775–794. (doi:10.1121/1.1332378)
- Bialystok, E. 1999 Cognitive complexity and attentional control in the bilingual mind. *Child Dev.* **70**, 636–644. (doi:10.1111/1467-8624.00046)
- Bialystok, E. 2001 *Bilingualism in development: language, literacy, and cognition*. New York, NY: Cambridge University Press.
- Bialystok, E. & Hakuta, K. 1994 *In other words: the science and psychology of second-language acquisition*. New York, NY: Basic Books.
- Birdsong, D. & Molis, M. 2001 On the evidence for maturational constraints in second-language acquisitions. *J. Mem. Lang.* **44**, 235–249. (doi:10.1006/jmla.2000.2750)
- Bornstein, M. H. & Cote, L. R. 2005 Expressive vocabulary in language learners from two ecological settings in three language communities. *Infancy* **7**, 299–316. (doi:10.1207/s15327078in0703\_5)
- Bortfeld, H., Morgan, J. L., Golinkoff, R. M. & Rathbun, K. 2005 Mommy and me: familiar names help launch babies into speech-stream segmentation. *Psychol. Sci.* **16**, 298–304. (doi:10.1111/j.0956-7976.2005.01531.x)
- Bosch, L. & Sebastián-Gallés, N. 2003a Simultaneous bilingualism and the perception of language-specific vowel contrast in the first year of life. *Lang. Speech* **46**, 217–243.
- Bosch, L. & Sebastián-Gallés, N. 2003b Language experience and the perception of a voicing contrast in fricatives: infant and adult data. In *International congress of phonetic sciences* (eds M. J. Sole, D. Recasens & J. Romero), pp. 1987–1990. Barcelona, Spain: UAB.
- Brainard, M. S. & 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. & Meltzoff, A. N. 2005 The development of gaze following and its relation to language. *Dev. Sci.* **8**, 535–543. (doi:10.1111/j.1467-7687.2005.00445.x)
- Bruer, J. T. In press. Critical periods: phenomena versus theories. In *Language impairment and reading disability: interactions among brain, behavior, and experience* (eds M. Mody & E. Silliman). New York, NY: Guilford Press.
- Bruner, I. 1983 *Child's talk: learning to use language*. New York, NY: W. W. Norton.
- Burnham, D. K. 1986 Developmental loss of speech perception: exposure to and experience with a first language. *Appl. Psycholinguist.* **7**, 207–240.
- Burnham, D., Kitamura, C. & Vollmer-Conna, U. 2002 What's new pussycat? On talking to babies and animals. *Science* **296**, 1435. (doi:10.1126/science.1069587)
- Burns, T. C., Yoshida, K. A., Hill, K. & Werker, J. F. 2007 The development of phonetic representation in bilingual and monolingual infants. *Appl. Psycholinguist.* **28**, 455–474. (doi:10.1017/S0142716407070257)
- Callan, D. E., Jones, J. A., Callan, A. M. & Akahane-Yamada, R. 2004 Phonetic perceptual identification by native- and second language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory–auditory/orosensory internal models. *NeuroImage* **22**, 1182–1194. (doi:10.1016/j.neuroimage.2004.03.006)
- Carney, A. E., Widin, G. P. & Viemeister, N. F. 1977 Noncategorical perception of stop consonants differing in VOT. *J. Acoust. Soc. Am.* **62**, 961–970. (doi:10.1121/1.381590)
- Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K. & Näätänen, R. 1998 Development of language-specific phoneme representations in the infant brain. *Nat. Neurosci.* **1**, 351–353. (doi:10.1038/1561)
- Conboy, B. T. & Kuhl, P. K. 2007 ERP mismatch negativity effects in 11-month-old infants after exposure to Spanish. Poster presented at the *Biennial Meeting of the Society for Research in Child Development, Boston, March 29–April 1, 2007*.
- Conboy, B. T. & Mills, D. L. 2006 Two languages, one developing brain: event-related potentials to words in bilingual toddlers. *Dev. Sci.* **9**, F1–F12. (doi:10.1111/j.1467-7687.2005.00453.x)
- Conboy, B. T., Rivera-Gaxiola, M., Klarman, L., Aksoylu, E. & Kuhl, P. K. 2005 Associations between native and nonnative speech sound discrimination and language development at the end of the first year. In *Supplement to the Proc. 29th Boston University Conf. on Language Development* (eds A. Brugos, M. R. Clark-Cotton & S. Ha). See <http://www.bu.edu/linguistics/APPLIED/BUCLD/supp29.html>.
- Conboy, B. T., Sommerville, J. & Kuhl, P. K. Submitted. Cognitive control factors in speech perception at 11 months.
- Dalston, R. M. 1975 Acoustic characteristics of English /w,r,l/ spoken correctly by young children and adults. *J. Acoust. Soc. Am.* **57**, 462–469. (doi:10.1121/1.380469)
- de Boer, B. & Kuhl, P. K. 2003 Investigating the role of infant-directed speech with a computer model. *ARLO* **4**, 129–134. (doi:10.1121/1.1613311)
- de Boysson-Bardies, B. 1993 Ontogeny of language-specific syllabic productions. In *Developmental neurocognition: speech and face processing in the first year of life* (eds B. de Boysson-Bardies, S. de Schonen, P. Juszyk, P. McNeilage & J. Morton), pp. 353–363. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Dehaene-Lambertz, G. & Gliga, T. 2004 Common neural basis for phoneme processing in infants and adults. *J. Cogn. Neurosci.* **16**, 1375–1387. (doi:10.1162/0898929042304714)
- Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Mériaux, S., Roche, A., Sigman, M. & Dehaene, S. 2006 Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proc. Natl Acad. Sci. USA* **103**, 14 240–14 245. (doi:10.1073/pnas.0606302103)
- DePaolis, R. 2005 The influence of production on perception: output as input. In *Supplement to the Proc 29th Boston University Conf. on Language Development* (eds A. Brugos, M. R. Clark-Cotton & S. Ha). See <http://www.bu.edu/linguistics/APPLIED/BUCLD/supp29.html>.
- Diamond, A., Werker, J. F. & Lalonde, C. 1994 Toward understanding commonalities in the development of object search, detour navigation, categorization, and speech perception. In *Human behavior and the developing brain* (eds G. Dawson & K. W. Fischer), pp. 380–426. New York, NY: Guilford Press.
- Dooling, R. J., Best, C. T. & Brown, S. D. 1995 Discrimination of synthetic full-formant and sinewave /ra-la/ continua by budgerigars (*Melopsittacus undulatus*) and zebra finches (*Taeniopygia guttata*). *J. Acoust. Soc. Am.* **97**, 1839–1846. (doi:10.1121/1.412058)
- Eales, L. 1989 The influences of visual and vocal interaction on song learning in zebra finches. *Anim. Behav.* **37**, 507–508. (doi:10.1016/0003-3472(89)90097-3)
- Eilers, R. E., Wilson, W. R. & Moore, J. M. 1977 Developmental changes in speech discrimination in infants. *J. Speech Hear. Res.* **20**, 766–780.
- Eimas, P. D. 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, P. D., Siqueland, E. R., Juszyk, P. & Vigorito, J. 1971 Speech perception in infants. *Science* **171**, 303–306. (doi:10.1126/science.171.3968.303)
- Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. 1996 *Rethinking innateness: a connectionist perspective on development*. Cambridge, MA: MIT Press.
- Englund, K. T. 2005 Voice onset time in infant directed speech over the first six months. *First Lang.* **25**, 219–234. (doi:10.1177/0142723705050286)
- Fant, G. 1973 *Speech sounds and features*. Cambridge, MA: MIT Press.
- Fenson, L., Dale, P., Reznick, J. S., Thal, D., Bates, E., Hartung, J., Pethick, S. & Reilly, J. S. 1993 *MacArthur communicative development inventories: user's guide and technical manual*. San Diego, CA: Singular Publishing Group.
- Fenson, L., Dale, P., Reznick, J. S., Bates, E., Thal, D. & Pethick, S. 1994 Variability in early communicative development. *Monogr. Soc. Res. Child* **59**, 1–185.
- Ferguson, C., Menn, L. & Stoel-Gammon, C. (eds) 1992 *Phonological development: models, research, implications*. Timonium, MD: York Press.
- Fernald, A. 1984 The perceptual and affective salience of mothers' speech to infants. In *The origins and growth of communication* (eds L. Feagans, C. Garvey, R. Golinkoff, M. T. Greenberg, C. Harding & J. Bohannon), pp. 5–29. Norwood, NJ: Ablex.
- Fernald, A. & Kuhl, P. 1987 Acoustic determinants of infant preference for motherese speech. *Infant Behav. Dev.* **10**, 279–293. (doi:10.1016/0163-6383(87)90017-8)
- Fernald, A., Perfors, A. & Marchman, V. A. 2006 Picking up speed in understanding: speech processing efficiency and vocabulary growth across the 2nd year. *Dev. Psychol.* **42**, 98–116. (doi:10.1037/0012-1649.42.1.98)
- Fitch, W. T. & Hauser, M. D. 2004 Computational constraints on syntactic processing in a nonhuman primate. *Science* **303**, 377–380. (doi:10.1126/science.1089401)

- Fitch, W. T., Hauser, M. D. & Chomsky, N. 2005 The evolution of the language faculty: clarifications and implications. *Cognition* **97**, 179–210. (doi:10.1016/j.cognition.2005.02.005)
- Flege, J. E. 1988 The production and perception of foreign language speech sounds. In *Human communication and its disorders, a review—1988* (ed. H. Winitz), pp. 224–401. Norwood, NJ: Ablex.
- Flege, J. E. 1995 Second language speech learning: theory, findings, and problems. In *Speech perception and linguistic experience* (ed. W. Strange), pp. 233–277. Timonium, MD: York Press.
- Flege, J. E., Takagi, N. & Mann, V. 1995 Japanese adults can learn to produce English /r/ and /l/ accurately. *Lang. Speech* **38**, 25–55.
- Flege, J. E., Yeni-Komshian, G. H. & Liu, S. 1999 Age constraints on second-language acquisition. *J. Mem. Lang.* **41**, 78–104. (doi:10.1006/jmla.1999.2638)
- Fowler, C. A. 1986 An event approach to the study of speech perception from a direct-realist perspective. *J. Phonet.* **14**, 3–28.
- Friederici, A. D. 2005 Neurophysiological markers of early language acquisition: from syllables to sentences. *Trends Cogn. Sci.* **9**, 481–488. (doi:10.1016/j.tics.2005.08.008)
- Friederici, A. D. & Wessels, J. M. I. 1993 Phonotactic knowledge of word boundaries and its use in infant speech perception. *Percept. Psychophys.* **54**, 287–295.
- Gallese, V. 2003 The manifold nature of interpersonal relations: the quest for a common mechanism. *Phil. Trans. R. Soc. B* **358**, 517–528. (doi:10.1098/rstb.2002.1234)
- Ganger, J. & Brent, M. R. 2004 Reexamining the vocabulary spurt. *Dev. Psychol.* **40**, 621–632. (doi:10.1037/0012-1649.40.4.621)
- Gerken, L., Wilson, R. & Lewis, W. 2005 Infants can use distributional cues to form syntactic categories. *J. Child Lang.* **32**, 249–268. (doi:10.1017/S0305000904006786)
- Golestani, N. & Zatorre, R. J. 2004 Learning new sounds of speech: reallocation of neural substrates. *NeuroImage* **21**, 494–506. (doi:10.1016/j.neuroimage.2003.09.071)
- Gomez, R. L. & Gerken, L. 1999 Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition* **70**, 109–135. (doi:10.1016/S0010-0277(99)00003-7)
- Gomez, R. L. & Gerken, L. 2000 Infant artificial language learning and language acquisition. *Trends Cogn. Sci.* **4**, 178–186. (doi:10.1016/S1364-6613(00)01467-4)
- Goodsitt, J. V., Morgan, J. L. & Kuhl, P. K. 1993 Perceptual strategies in prelingual speech segmentation. *J. Child Lang.* **20**, 229–252.
- Goodz, N. S. 1989 Parental language mixing in bilingual families. *Inf. Mental Hlth. J.* **10**, 25–44. (doi:10.1002/1097-0355(198921)10:1<25::AID-IMHJ2280100104>3.0.CO;2-R)
- Gopnik, A. & Meltzoff, A. N. 1987 The development of categorization in the second year and its relation to other cognitive and linguistic developments. *Child Dev.* **58**, 1523–1531. (doi:10.2307/1130692)
- Gopnik, A. & Meltzoff, A. N. 1997 *Words, thoughts, and theories*. Cambridge, MA: MIT Press.
- Guenther, F. H., Nieto-Castanon, A., Ghosh, S. S. & Tourville, J. A. 2004 Representation of sound categories in auditory cortical maps. *J. Speech Lang. Hear. Res.* **47**, 46–57. (doi:10.1044/1092-4388(2004)005)
- Hauser, M. D., Newport, E. L. & Aslin, R. N. 2001 Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition* **78**, B53–B64. (doi:10.1016/S0010-0277(00)00132-3)
- Hauser, M. D., Chomsky, N. & Fitch, W. T. 2002a The faculty of language: what is it, who has it, and how did it evolve? *Science* **298**, 1569–1579. (doi:10.1126/science.298.5598.1569)
- Hauser, M. D., Weiss, D. & Marcus, G. 2002b Rule learning by cotton-top tamarins. *Cognition* **86**, B15–B22. (doi:10.1016/S0010-0277(02)00139-7)
- Hess, E. H. 1973 *Imprinting: early experience and the developmental psychobiology of attachment*. New York, NY: Von Nostrand Reinhold Company.
- Hillenbrand, J., Getty, L. A., Clark, M. J. & Wheeler, K. 1995 Acoustic characteristics of American English vowels. *J. Acoust. Soc. Am.* **97**, 3099–3111. (doi:10.1121/1.411872)
- Hollich, G., Hirsh-Pasek, K. & Golinkoff, R. 2000 Breaking the language barrier: an emergentist coalition model for the origins of word learning. *Monogr. Soc. Res. Child Dev.* **65**, 262.
- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M. & Lyons, T. 1991 Early vocabulary growth: relation to language input and gender. *Dev. Psychol.* **27**, 236–248. (doi:10.1037/0012-1649.27.2.236)
- Imada, T., Zhang, Y., Cheour, M., Taulu, S., Ahonen, A. & Kuhl, P. K. 2006 Infant speech perception activates Broca's area: a developmental magnetoencephalography study. *NeuroReport* **17**, 957–962. (doi:10.1097/01.wnr.0000223387.51704.89)
- Immelmann, K. 1969 Song development in the zebra finch and other estrildid finches. In *Bird vocalizations* (ed. R. Hinde), pp. 61–74. London, UK: Cambridge University Press.
- Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A. & Siebert, C. 2003 A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* **87**, B47–B57. (doi:10.1016/S0010-0277(02)00198-1)
- Iverson, P., Hazan, V. & Bannister, K. 2005 Phonetic training with acoustic cue manipulations: a comparison of methods for teaching English /r/-/l/ to Japanese adults. *J. Acoust. Soc. Am.* **118**, 3267–3278. (doi:10.1121/1.2062307)
- Jackson, P. L., Meltzoff, A. N. & Decety, J. 2006 Neural circuits involved in imitation and perspective taking. *NeuroImage* **31**, 429–439. (doi:10.1016/j.neuroimage.2005.11.026)
- 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. (doi:10.1016/0010-0285(89)90003-0)
- Jusczyk, P. W. 1997 *The discovery of spoken language*. Cambridge, MA: MIT Press.
- Jusczyk, P. W., Rosner, B. S., Cutting, J. E., Foard, C. F. & Smith, L. B. 1977 Categorical perception of nonspeech sounds by 2-month-old infants. *Percept. Psychophys.* **21**, 50–54.
- Jusczyk, P. W., Pisoni, D. B., Walley, A. & Murray, J. 1980 Discrimination of relative onset time of two-component tones by infants. *J. Acoust. Soc. Am.* **67**, 262–270. (doi:10.1121/1.383735)
- Jusczyk, P. W., Friederici, A. D., Wessels, J. M. I., Svenkerud, V. Y. & Jusczyk, A. M. 1993 Infants' sensitivity to the sound patterns of native language words. *J. Mem. Lang.* **32**, 402–420. (doi:10.1006/jmla.1993.1022)
- Jusczyk, P. W., Luce, P. A. & Charles-Luce, J. 1994 Infants' sensitivity to phonotactic patterns in the native language. *J. Mem. Lang.* **33**, 630–645. (doi:10.1006/jmla.1994.1030)
- Karmiloff-Smith, A. 1992 *Beyond modularity: a developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
- Kluender, K. R., Diehl, R. L. & Killeen, P. R. 1987 Japanese quail can learn phonetic categories. *Science* **237**, 1195–1197. (doi:10.1126/science.3629235)

- Knudsen, E. I. 2004 Sensitive periods in the development of the brain and behavior. *J. Cogn. Neurosci.* **16**, 1412–1425. (doi:10.1162/0898929042304796)
- Koyama, S., Akahane-Yamada, R., Gunji, A., Kubo, R., Roberts, T. P., Yabe, H. & Kakigi, R. 2003 Cortical evidence of the perceptual backward masking effect on /l/ and /r/ sounds from a following vowel in Japanese speakers. *NeuroImage* **18**, 962–974. (doi:10.1016/S1053-8119(03)00037-5)
- Kraus, N., McGee, T., Carrell, T., Zecker, S., Nicol, T. & Koch, D. 1996 Auditory neurophysiologic responses and discrimination deficits in children with learning problems. *Science* **273**, 971–973. (doi:10.1126/science.273.5277.971)
- Kuhl, P. K. 1980 Perceptual constancy for speech-sound categories in early infancy. In *Child phonology, vol. 2, Perception* (eds G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson), pp. 41–66. New York, NY: Academic Press.
- Kuhl, P. K. 1991a Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Percept. Psychophys.* **50**, 93–107.
- Kuhl, P. K. 1991b Perception, cognition, and the ontogenetic and phylogenetic emergence of human speech. In *Plasticity of development* (eds S. E. Brauth, W. S. Hall & R. J. Dooling), pp. 73–106. Cambridge, MA: MIT Press.
- Kuhl, P. K. 1992 Psychoacoustics and speech perception: internal standards, perceptual anchors, and prototypes. In *Developmental psychoacoustics* (eds L. A. Werner & E. W. Rubel), pp. 293–332. Washington, DC: American Psychological Association.
- Kuhl, P. K. 1993 Early linguistic experience and phonetic perception: implications for theories of developmental speech perception. *J. Phonet.* **21**, 125–139.
- Kuhl, P. K. 1994 Learning and representation in speech and language. *Curr. Opin. Neurobiol.* **4**, 812–822. (doi:10.1016/0959-4388(94)90128-7)
- Kuhl, P. K. 1998 Effects of language experience on speech perception. *J. Acoust. Soc. Am.* **103**, 29–31. (doi:10.1121/1.422159)
- Kuhl, P. K. 2000a A new view of language acquisition. *Proc. Natl Acad. Sci. USA* **97**, 11 850–11 857. (doi:10.1073/pnas.97.22.11850)
- Kuhl, P. K. 2000b Language, mind, and brain: experience alters perception. In *The new cognitive neurosciences* (ed. M. Gazzaniga) pp. 99–115, 2nd edn. Cambridge, MA: MIT Press.
- Kuhl, P. K. 2004 Early language acquisition: cracking the speech code. *Nat. Rev. Neurosci.* **5**, 831–843. (doi:10.1038/nrn1533)
- Kuhl, P. K. 2007 Is speech learning ‘gated’ by the social brain? *Dev. Sci.* **10**, 110–120. (doi:10.1111/j.1467-7687.2007.00572.x)
- Kuhl, P. K. & Meltzoff, A. N. 1982 The bimodal perception of speech in infancy. *Science* **218**, 1138–1141. (doi:10.1126/science.7146899)
- Kuhl, P. K. & Meltzoff, A. N. 1996 Infant vocalizations in response to speech: vocal imitation and developmental change. *J. Acoust. Soc. Am.* **100**, 2425–2438. (doi:10.1121/1.417951)
- Kuhl, P. K. & Miller, J. D. 1975 Speech perception by the chinchilla: voiced–voiceless distinction in alveolar plosive consonants. *Science* **190**, 69–72. (doi:10.1126/science.1166301)
- Kuhl, P. K. & Miller, J. D. 1978 Speech perception by the chinchilla: identification functions for synthetic VOT stimuli. *J. Acoust. Soc. Am.* **63**, 905–917. (doi:10.1121/1.381770)
- Kuhl, P. K. & Padden, D. M. 1982 Enhanced discriminability at the phonetic boundaries for the voicing feature in macaques. *Percept. Psychophys.* **32**, 542–550.
- Kuhl, P. K. & Padden, D. M. 1983 Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *J. Acoust. Soc. Am.* **73**, 1003–1010. (doi:10.1121/1.389148)
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N. & Lindblom, B. 1992 Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* **255**, 606–608. (doi:10.1126/science.1736364)
- Kuhl, P. K., Andruski, J., Christovich, I., Chistovich, L., Kozhevnikova, E., Ryskina, V., Stolyarova, E., Sungberg, U. & Lacerda, F. 1997 Cross-language analysis of phonetic units in language addressed to infants. *Science* **277**, 684–686. (doi:10.1126/science.277.5326.684)
- Kuhl, P. K., Tsao, F.-M. & Liu, H.-M. 2003 Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proc. Natl Acad. Sci. USA* **100**, 9096–9101. (doi:10.1073/pnas.1532872100)
- Kuhl, P. K., Coffey-Corina, S., Padden, D. & Dawson, G. 2005a Links between social and linguistic processing of speech in preschool children with autism: behavioral and electrophysiological measures. *Dev. Sci.* **8**, F1–F12. (doi:10.1111/j.1467-7687.2004.00384.x)
- Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T. & Pruitt, J. 2005b Early speech perception and later language development: implications for the “critical period”. *Lang. Learn. Dev.* **1**, 237–264.
- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S. & Iverson, P. 2006 Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Dev. Sci.* **9**, F13–F21. (doi:10.1111/j.1467-7687.2006.00468.x)
- Kuhl, P. K., Coffey-Corina, S. & Padden, D. In preparation. Brain and behavioral measures of learning in infants after exposure to a second language.
- Lacerda, F. In preparation. Acoustic analysis of Swedish /r/ and /l/.
- Lalonde, C. E. & Werker, J. F. 1995 Cognitive influences on cross-language speech perception in infancy. *Infant Behav. Dev.* **18**, 459–475. (doi:10.1016/0163-6383(95)90035-7)
- Lenneberg, E. 1967 *Biological foundations of language*. New York, NY: Wiley.
- Lieberman, A. M. & Mattingly, I. G. 1985 The motor theory of speech perception revised. *Cognition* **21**, 1–36. (doi:10.1016/0010-0277(85)90021-6)
- Lieberman, A. M., Cooper, F. S., Shankweiler, D. P. & Studdert-Kennedy, M. 1967 Perception of the speech code. *Psychol. Rev.* **74**, 431–461. (doi:10.1037/h0020279)
- Liu, H.-M., Kuhl, P. K. & Tsao, F.-M. 2003 An association between mothers’ speech clarity and infants’ speech discrimination skills. *Dev. Sci.* **6**, F1–F10. (doi:10.1111/1467-7687.00275)
- Liu, H.-M., Tsao, F.-M., & Kuhl, P. K. In press. Lexical tone exaggeration in Mandarin infant-directed speech. *Dev. Sci.*
- Lorenz, K. 1957 Companionship in bird life. In *Instinctive behavior: the development of a modern concept* (ed. C. H. Schiller), pp. 83–128. New York, NY: International Universities Press.
- Lotto, A. J., 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* (eds J. Slitka, S. Manuel & M. Matthies). Cambridge, MA: MIT Press.
- MacLeod, A. A. N. & Stoel-Gammon, C. 2005 Are bilinguals different? What VOT tells us about simultaneous bilinguals. *J. Multiling. Commun. Disord.* **3**, 118–127. (doi:10.1080/14769670500066313)
- Malsheen, B. 1980 Two hypotheses for phonetic clarification in the speech of mothers to children. In *Child phonology*,

- vol. 2, *Perception* (eds G. H. Yeni-Komshian J. F. Kavanaugh & C. A. Ferguson), pp. 173–184. New York, NY: Academic Press.
- Marcus, G. F., Vijayan, S., Bandi Rao, S. & Vishton, P. M. 1999 Rule learning by seven-month-old infants. *Science* **283**, 77–80. (doi:10.1126/science.283.5398.77)
- Martin, B. A., Shafer, V. L., Mara, L., Morr, M. L., Kreuzer, J. A. & Kurtzberg, D. 2003 Maturation of mismatch negativity: a scalp current density analysis. *Ear Hearing* **24**, 463–471. (doi:10.1097/01.AUD.0000100306.20188.0E)
- Mattys, S. L., Jusczyk, P. W., Luce, P. A. & Morgan, J. L. 1999 Phonotactic and prosodic effects on word segmentation in infants. *Cogn. Psychol.* **38**, 465–494. (doi:10.1006/cogp.1999.0721)
- Maye, J., Werker, J. F. & Gerken, L. 2002 Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* **82**, B101–B111. (doi:10.1016/S0010-0277(01)00157-3)
- Maye, J. C., Weiss, D. & Aslin, R. In press. Statistical phonetic learning in infants: facilitation and feature generalization. *Dev. Sci.*
- McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M. & McClelland, J. L. 2002 Success and failure in teaching the [r]–[l] contrast to Japanese adults: tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cogn. Affect. Behav. Neurosci.* **2**, 89–108.
- McMurray, B. & Aslin, R. N. 2005 Infants are sensitive to within-category variation in speech perception. *Cognition* **95**, B15–B26. (doi:10.1016/j.cognition.2004.07.005)
- Meltzoff, A. N. 2005 Imitation and other minds: the “Like me” hypothesis. In *Perspectives on imitation: from cognitive neuroscience to social science*, vol. 2 (eds S. Hurley & N. Chater), pp. 55–77. Cambridge, MA: MIT Press.
- Meltzoff, A. N. & Decety, J. 2003 What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. *Phil. Trans. R. Soc. B* **358**, 491–500. (doi:10.1098/rstb.2002.1261)
- Meltzoff, A. N. & Moore, M. K. 1997 Explaining facial imitation: a theoretical model. *Early Dev. Parenting* **6**, 179–192. (doi:10.1002/(SICI)1099-0917(199709/12)6:3/4<179::AID-EDP157>3.0.CO;2-R)
- Merzenich, M., Jenkins, W., Johnston, P. S., Schreiner, C., Miller, S. L. & Tallal, P. 1996 Temporal processing deficits of language-learning impaired children ameliorated by training. *Science* **271**, 77–81. (doi:10.1126/science.271.5245.77)
- Miller, J. L. & Eimas, P. D. 1996 Internal structure of voicing categories in early infancy. *Percept. Psychophys.* **58**, 1157–1167.
- Mills, D., Prat, C., Zangl, R., Stager, C., Neville, H. & Werker, J. 2004 Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14- and 20-month-olds. *J. Cogn. Neurosci.* **16**, 1452–1464. (doi:10.1162/0898929042304697)
- Mills, D., Conboy, B. T. & Paton, C. 2005a How learning new words shapes the organization of the infant brain. In *Symbol use and symbolic representation* (ed. L. Namy), pp. 123–153. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Mills, D. L., Plunkett, K., Prat, C. & Schafer, G. 2005b Watching the infant brain learn words: effects of vocabulary size and experience. *Cogn. Dev.* **20**, 19–31. (doi:10.1016/j.cogdev.2004.07.001)
- Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J. & Fujimura, O. 1975 An effect of linguistic experience: the discrimination of [r] and [l] by native speakers of Japanese and English. *Percept. Psychophys.* **18**, 331–340.
- Molfese, D. L. 2000 Predicting dyslexia at 8 years of age using neonatal brain responses. *Brain Lang.* **72**, 238–245. (doi:10.1006/brln.2000.2287)
- Molfese, D. L. & Molfese, V. 1997 Discrimination of language skills at five years of age using event-related potentials recorded at birth. *Dev. Neuropsychol.* **13**, 135–156.
- Moon, C., Cooper, R. P. & Fifer, W. P. 1993 2-day-olds prefer their native language. *Infant Behav. Dev.* **16**, 495–500. (doi:10.1016/0163-6383(93)80007-U)
- Moore, J. K. & Guan, Y. L. 2001 Cytoarchitectural and axonal maturation in human auditory cortex. *J. Assoc. Res. Otolaryngol.* **2**, 297–311. (doi:10.1007/s101620010052)
- Morr, M. L., Shafer, V. L., Kreuzer, J. A. & Kurtzberg, D. 2002 Maturation of mismatch negativity in typically developing infants and preschool children. *Ear Hear.* **23**, 118–136. (doi:10.1097/00003446-200204000-00005)
- Näätänen, R. et al. 1997 Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* **385**, 432–434. (doi:10.1038/385432a0)
- Newman, R., Ratner, N. B., Jusczyk, A. M., Jusczyk, P. W. & Dow, K. A. 2006 Infants’ early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. *Dev. Psychol.* **42**, 643–655. (doi:10.1037/0012-1649.42.4.643)
- Newport, E. L. & Aslin, R. N. 2004 Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cogn. Psychol.* **48**, 127–162. (doi:10.1016/S0010-0285(03)00128-2)
- Newport, E. L., Bavelier, D. & 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* (ed. E. Dupoux), pp. 481–502. Cambridge, MA: MIT Press.
- Newport, E. L., Hauser, M. D., Sepaepen, G. & Aslin, R. N. 2004 Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cogn. Psychol.* **49**, 85–117. (doi:10.1016/j.cogpsych.2003.12.002)
- Nittrouer, S. 2001 Challenging the notion of innate phonetic boundaries. *J. Acoust. Soc. Am.* **110**, 1598–1605. (doi:10.1121/1.1379078)
- Pang, E., Edmonds, G., Desjardins, R., Khan, S., Trainor, L. & Taylor, M. 1998 Mismatch negativity to speech stimuli in 8-month-old infants and adults. *Int. J. Psychophysiol.* **29**, 227–236. (doi:10.1016/S0167-8760(98)00018-X)
- Peña, M., Bonatti, L., Nespor, M. & Mehler, J. 2002 Signal-driven computations in speech processing. *Science* **298**, 604–607. (doi:10.1126/science.1072901)
- Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. & Fazio, F. 2003 The role of age of acquisition and language usage in early, high-proficient bilinguals: an fMRI study during verbal fluency. *Hum. Brain Mapp.* **19**, 170–182. (doi:10.1002/hbm.10110)
- Pettito, L. A., Holowka, S., Sergio, L. E., Levy, B. & Ostry, D. J. 2004 Baby hands that move to the rhythm of language: hearing babies acquiring sign language babble silently on the hands. *Cognition* **93**, 43–73. (doi:10.1016/j.cognition.2003.10.007)
- Phan, M. L., Pytte, C. L. & Vicario, D. S. 2006 Early auditory experience generates long-lasting memories that may subserve vocal learning in songbirds. *Proc. Natl Acad. Sci. USA* **103**, 1088–1093. (doi:10.1073/pnas.0510136103)
- Pinker, S. & Jackendoff, R. 2005 The faculty of language: what’s special about it? *Cognition* **95**, 201–236. (doi:10.1016/j.cognition.2004.08.004)
- Pisoni, D. B. & Lively, S. E. 1995 Variability and invariance in speech perception: a new look at some old problems in perceptual learning. In *Speech perception and linguistic experience: issues in cross-language research* (ed. W. Strange), pp. 429–455. Timonium, MD: York Press.
- Pisoni, D. B., Aslin, R. N., Perey, A. J. & Hennessy, B. L. 1982 Some effects of laboratory training on identification

- and discrimination of voicing contrasts in stop consonants. *J. Exp. Psychol. Hum. Percept. Perform.* **8**, 297–314. (doi:10.1037/0096-1523.8.2.297)
- Polka, L. & Bohn, O.-S. 1996 A cross-language comparison of vowel perception in English-learning and German-learning infants. *J. Acoust. Soc. Am.* **100**, 577–592. (doi:10.1121/1.415884)
- Polka, L. & Bohn, O.-S. 2003 Asymmetries in vowel perception. *Speech Commun.* **41**, 221–231. (doi:10.1016/S0167-6393(02)00105-X)
- Polka, L. & Werker, J. F. 1994 Developmental changes in perception of nonnative vowel contrasts. *J. Exp. Psychol. Hum. Percept. Perform.* **20**, 421–435. (doi:10.1037/0096-1523.20.2.421)
- Polka, L., Colantonio, C. & Sundara, M. 2001 A cross-language comparison of /d-/ð/ perception: evidence for a new developmental pattern. *J. Acoust. Soc. Am.* **109**, 2190–2201. (doi:10.1121/1.1362689)
- Raudenbush, S. W., Bryk, A. S. & Congdon, R. 2005 *HLM-6: hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International, Inc.
- Rivera-Gaxiola, M. & Romo, H. 2006 Infant head-start learners: brain and behavioral measures and family assessments. Paper presented at *From Synapse to Schoolroom: The Science of Learning. NSF Science of Learning Centers Satellite Symposium, Society for Neuroscience Annual Meeting, Atlanta, GA, 13 October*.
- Rivera-Gaxiola, M., Csibra, G., Johnson, M. H. & Karmiloff-Smith, A. 2000 Electrophysiological correlates of cross-linguistic speech perception in native English speakers. *Behav. Brain Res.* **111**, 11–23. (doi:10.1016/S0166-4328(00)00139-X)
- Rivera-Gaxiola, M., Klarman, L., Garcia-Sierra, A. & Kuhl, P. K. 2005a Neural patterns to speech and vocabulary growth in American infants. *NeuroReport* **16**, 495–498.
- Rivera-Gaxiola, M., Silva-Pereyra, J. & Kuhl, P. K. 2005b Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Dev. Sci.* **8**, 162–172. (doi:10.1111/j.1467-7687.2005.00403.x)
- Rivera-Gaxiola, M., Silva-Pereyra, J., Klarman, L., Garcia-Sierra, A., Lara-Ayala, L., Cadena-Salazar, C. & Kuhl, P. K. 2007 Principal component analyses and scalp distribution of the auditory P150-250 and N250-550 to speech contrasts in Mexican and American infants. *Dev. Neuropsychol.* **31**, 363–378.
- Rizzolatti, G., Fadiga, B., Gallese, V. & Fogassi, L. 1996 Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* **3**, 131–141. (doi:10.1016/0926-6410(95)00038-0)
- Rizzolatti, G., Fogassi, L. & Gallese, V. 2001 Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* **2**, 661–670. (doi:10.1038/35090060)
- Rosenthal, O., Fusi, S. & Hochstein, S. 2001 Forming classes by stimulus frequency: behavior and theory. *Proc. Natl Acad. Sci. USA* **98**, 4265–4270. (doi:10.1073/pnas.071525998)
- Saffran, J. R. 2002 Constraints on statistical language learning. *J. Mem. Lang.* **47**, 172–196. (doi:10.1006/jmla.2001.2839)
- Saffran, J. R. 2003 Statistical language learning: mechanisms and constraints. *Curr. Dir. Psychol. Sci.* **12**, 110–114. (doi:10.1111/1467-8721.01243)
- Saffran, J. R., Aslin, R. N. & Newport, E. L. 1996 Statistical learning by 8-month old infants. *Science* **274**, 1926–1928. (doi:10.1126/science.274.5294.1926)
- Sanders, L. D., Newport, E. L. & Neville, H. J. 2002 Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nat. Neurosci.* **5**, 700–703. (doi:10.1038/nn873)
- Seidenberg, M. S. & Zevin, J. D. 2006 Connectionist models in developmental cognitive neuroscience: critical periods and the paradox of success. In *Processes of change in brain and cognitive development - attention & performance XXI* (eds Y. Manakata & M. Johnson). Oxford, UK: Oxford University Press.
- Silva-Pereyra, J., Rivera-Gaxiola, M. & Kuhl, P. K. 2005 An event-related brain potential study of sentence comprehension in preschoolers: semantic and morphosyntactic processing. *Cogn. Brain Res.* **23**, 247–258. (doi:10.1016/j.cogbrainres.2004.10.015)
- Silva-Pereyra, J., Conboy, B. T., Klarman, L. & Kuhl, P. K. 2007 Grammatical processing without semantics? An event-related brain potential study of preschoolers using jabberwocky sentences. *J. Cogn. Neurosci.* **19**, 1–16. (doi:10.1162/jocn.2007.19.6.1050)
- Silven, M., Kouvo, A., Haapakoski, M., Lahteenmaki, V. & Kuhl, P. K. 2004 Language learning in infants from mono- and bilingual families. In *18th Biennial ISSBD Conference, Ghent, Belgium*.
- Skinner, B. F. 1957 *Verbal behavior*. New York, NY: Appleton-Century-Crofts, Inc.
- Streeter, L. A. 1976 Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature* **259**, 39–41. (doi:10.1038/259039a0)
- Sundara, M., Polka, L. & Genesee, F. 2006 Language-experience facilitates discrimination of /d- ð/ in monolingual and bilingual acquisition of English. *Cognition* **100**, 369–388. (doi:10.1016/j.cognition.2005.04.007)
- Sundberg, U. & Lacerda, F. 1999 Voice onset time in speech directed to infants and adults. *Phonetica* **56**, 186–199. (doi:10.1159/000028450)
- Swingle, D. & Aslin, R. N. 2002 Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychol. Sci.* **13**, 480–484. (doi:10.1111/1467-9280.00485)
- Tallal, P., Miller, S., Bedi, G., Byma, G., Wang, X., Nagarajan, S., Schreiner, C., Jenkins, W. & Merzenich, M. 1996 Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science* **271**, 81–84. (doi:10.1126/science.271.5245.81)
- Thal, D. 1991 Language and cognition in normal and late-talking toddlers. *Top. Lang. Disord.* **11**, 33–42.
- Tomasello, M. 2003 *Constructing a language*. Cambridge, MA: Harvard University Press.
- Tomasello, M. & Farrar, M. J. 1984 Cognitive bases of lexical development: object permanence and relational words. *J. Child Lang.* **11**, 477–493.
- Tomasello, M. & Farrar, M. J. 1986 Joint attention and early language. *Child Dev.* **57**, 1454–1463. (doi:10.2307/1130423)
- Trehub, S. E. 1976 The discrimination of foreign speech contrasts by infants and adults. *Child Dev.* **47**, 466–472. (doi:10.2307/1128803)
- Tsao, F.-M., Liu, H.-M. & 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. (doi:10.1111/j.1467-8624.2004.00726.x)
- Tsao, F.-M., Liu, H.-M. & Kuhl, P. K. 2006 Perception of native and non-native affricate-fricative contrasts: cross-language tests on adults and infants. *J. Acoust. Soc. Am.* **120**, 2285–2294. (doi:10.1121/1.2338290)
- Vallabha, G. K. & McClelland, J. L. 2007 Success and failure of new speech category learning in adulthood: consequences of learned Hebbian attractors in topographic maps. *Cogn. Affect. Behav. Neurosci.* **7**, 53–73.

- Vouloumanos, A. & Werker, J. F. 2004 Tuned to the signal: the privileged status of speech for young infants. *Dev. Sci.* **7**, 270–276. (doi:10.1111/j.1467-7687.2004.00345.x)
- Wang, Y., Sereno, J. A., Jongman, A. & Hirsch, J. 2003 fMRI evidence for cortical modification during learning of Mandarin lexical tone. *J. Cogn. Neurosci.* **15**, 1019–1027. (doi:10.1162/089892903770007407)
- Weber-Fox, C. M. & Neville, H. J. 1999 Functional neural subsystems are differentially affected by delays in second language immersion: ERP and behavioral evidence in bilinguals. In *Second language acquisition and the critical period hypothesis* (ed. D. Birdsong). Mahwah, NJ: Lawrence Erlbaum and Associates, Inc.
- Werker, J. F. 1995 Exploring developmental changes in cross-language speech perception. In *Invitation to cognitive science* (eds L. Gleitman & M. Liberman). Cambridge, MA: MIT Press.
- Werker, J. F. & Curtin, S. 2005 PRIMIR: a developmental Framework of infant speech processing. *Lang. Learn. Dev.* **1**, 197–234. (doi:10.1207/s15473341lld0102\_4)
- Werker, J. F. & Lalonde, C. E. 1988 Cross-language speech perception: initial capabilities and developmental change. *Dev. Psychol.* **24**, 672–683. (doi:10.1037/0012-1649.24.5.672)
- Werker, J. F. & Logan, J. S. 1985 Cross-language evidence for three factors in speech perception. *Percept. Psychophys.* **37**, 35–44.
- Werker, J. F. & Pegg, J. E. 1992 Infant speech perception and phonological acquisition. In *Phonological development: models, research, and implications* (eds C. A. Ferguson, L. Menn & C. Stoel-Gammon), pp. 285–311. Timonium, MD: York Press.
- Werker, J. F. & Tees, R. C. 1984a Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behav. Dev.* **7**, 49–63. (doi:10.1016/S0163-6383(84)80022-3)
- Werker, J. F. & Tees, R. C. 1984b Phonemic and phonetic factors in adult cross-language speech perception. *J. Acoust. Soc. Am.* **75**, 1866–1878. (doi:10.1121/1.390988)
- Werker, J. F. & Tees, R. C. 2005 Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Dev. Psychobiol.* **46**, 233–251. (doi:10.1002/dev.20060)
- Werker, J. F., Fennell, C. T., Corcoran, K. M. & Stager, C. L. 2002 Infants' ability to learn phonetically similar words: effects of age and vocabulary size. *Infancy* **3**, 1–30. (doi:10.1207/15250000252828226)
- Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L. & Amano, S. 2007 Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition* **103**, 147–162. (doi:10.1016/j.cognition.2006.03.006)
- West, M. & King, A. 1988 Female visual displays affect the development of male song in the cowbird. *Nature* **334**, 244–246. (doi:10.1038/334244a0)
- Yeni-Komshian, G. H., Flege, J. E. & Liu, S. 2000 Pronunciation proficiency in the first and second languages of Korean–English bilinguals. *Biling-Lang. Cogn.* **3**, 131–149. (doi:10.1017/S1366728900000225)
- Yoshida, K. A., Pons, F., Cady, J. C. & Werker, J. F. 2006 Distributional learning and attention in phonological development. Paper presented at *Int. Conf. on Infant Studies, Kyoto Japan, 19–23 June*.
- Yu, C., Ballard, D. H. & Aslin, R. N. 2005 The role of embodied intention in early lexical acquisition. *Cogn. Sci.* **29**, 961–1005.
- Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M. & Tohkura, Y. 2005 Effects of language experience: neural commitment to language-specific auditory patterns. *NeuroImage* **26**, 703–720. (doi:10.1016/j.neuroimage.2005.02.040)
- Zhang, Y., Kuhl, P. K., Imada, T., Iverson, P., Pruitt, J. & Stevens, E. B. Submitted. Neural signatures of phonetic learning in adulthood: a magnetoencephalography (MEG) study.