The Design of Animal Communication

edited by Marc D. Hauser and Mark Konishi

15 Speech, Language, and the Brain: Innate Preparation for Learning

Patricia K. Kuhl

Thus, the question to be addressed is not "do innate mechanisms exist?" but rather, "what is the nature of the innate mechanisms for learning, by what mechanisms do they operate, and what provisions do they make for the interaction between organisms and their environments?" In other words, what are the ways in which innate mechanisms impinge on the pervasive plasticity that behavior displays in the course of its development?
—Peter Marler, Innate learning preferences

Nature, Nurture, and a Historical Debate

Forty years ago, there was a historic confrontation between a strong nativist and a strong learning theorist. Chomsky's (1957) reply to Skinner's (1957) *Verbal Behavior* had just been published, reigniting the debate on the nature of language. On Chomsky's (1965, 1981) nativist view, universal rules encompassing the grammars and phonologies of all languages were taken to be innately specified. Language input served to "trigger" the appropriate subset of rules, and developmental change in language ability was viewed as biological growth akin to other bodily organs, rather than learning. On the Skinnerian view, language was explicitly learned. Language was thought to be brought about in the child through a process of feedback and external reinforcement (Skinner 1957).

Both views made assumptions about three critical parameters: (i) the biological preparation that infants bring to the task of language learning, (ii) the nature of language input, and (iii) the nature of developmental change. Chomsky asserted, through the "poverty of the stimulus" argument, that language input to the child is greatly underspecified. Critical elements are missing; hence the necessity for innately specified information. Skinner viewed speech as simply another operant behavior, shaped through parental feedback and reinforcement like all other behaviors.

In the decades that have passed since these positions were developed, the debate has been played out for language at the syntactic, semantic, and phonological levels. It has also been played out for communication in animal models. Peter Marler's work on the acquisition of birdsong provides the quintessential example. Beginning with his theoretical writings in the 1970s, it was clear that Marler was a pioneer in promoting the epigenetic view that innate predispositions prepare organisms for learning (Marler 1970). Marler's (1991) view that an "instinct to learn" was key to both language and birdsong has been seminal to later discussions on the topic by many theorists.

In this chapter, I will concentrate primarily on language, using the elementary components of sounds—the consonants and vowels that make up words—to structure an

argument about what is given by nature and gained by experience in the acquisition of language at the phonetic level. Studying the sound structure of language allows us to test the perception of language in infants just hours old and thereby to address the question of what language capacities are innate in infants. Then, by tracking the development of infants raised in various cultures listening to different languages, we can determine, again using tests of perception, when infants' use of language begins to diverge as a function of experience with a particular language. These methods provide a strong test of the historically opposing views, and the results of these tests deliver dramatic evidence of the interaction between biology and culture, leading to a new view. Parallels between speech and birdsong will be highlighted in describing this new approach (see also Doupe and Kuhl 1999).

Origins of Conceptual Distinctions and the Modern View

The discussion between linguists and psychologists regarding language is only one forum in which the nature-nurture issue has been debated. Begun by philosophers hundreds of years ago, the nativism-empiricism debate or the nativist-constructivist debate concerned the origins of knowledge and whether it stemmed from our native abilities or was empirically derived (Berkeley 1709; Locke 1690). The debate is of continuing interest across a wide variety of disciplines, fueled by arguments and data from ethology (Bateson 1991; Hauser 1996), neuroscience (Carew, Menzel, and Shatz 1998), language science (Kuhl 1994, 1998a; Pinker 1994), and developmental psychology (Bates and Elman 1996; Carey 1985; Gopnik and Meltzoff 1997; Karmiloff-Smith 1995).

These groups use different terms to distinguish complex behaviors that appear relatively immune from as opposed to wholly dependent on experience. The terminology reflects differences in emphasis between groups. In ethology, for example, the distinction has traditionally been drawn between *innate* or *instinctual* behaviors, considered to be genetically determined, and those *learned* as a function of exposure to the environment (Lorenz 1965; Thorpe 1959; Tinbergen 1951). The emphasis in early ethological writings was on explaining behaviors that existed at birth in the absence of experience (Lorenz's "innate release mechanisms").

In the early psychological literature on the mental development of the child (James 1890; Koffka 1924; Piaget 1954; Vygotsky 1962) and also in the neuroscience literature (Cajal 1906), the distinction has historically been drawn using the terms "development" and "learning." *Development* included changes in the organism over time that depend primarily on maturation or internal factors leading to the expression of

Conceptual relations between development and learning
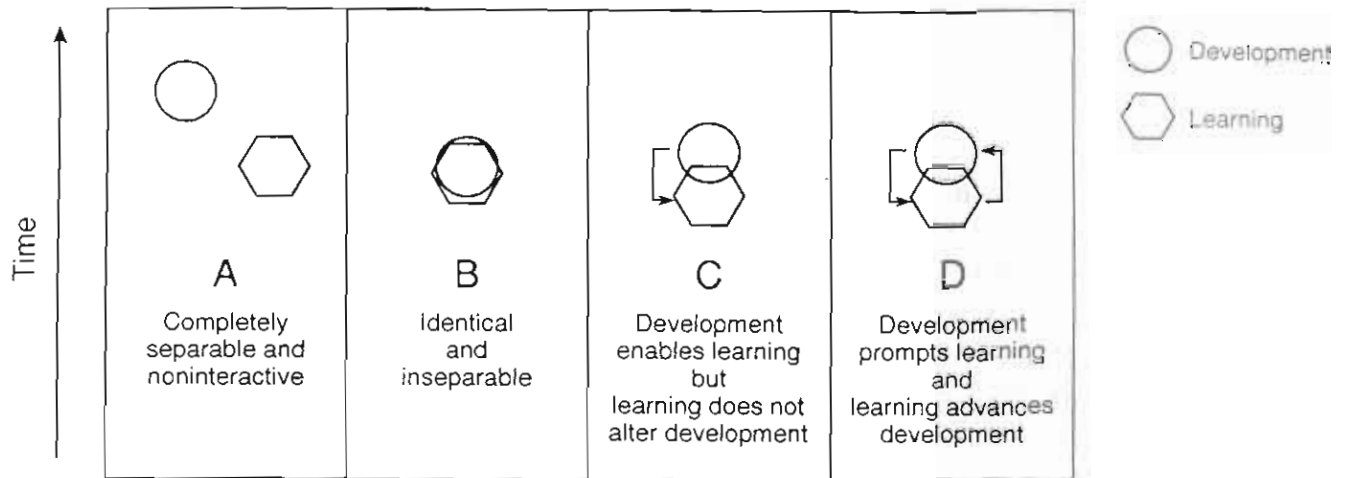
Figure 15.1
Conceptual relations between development and learning.

information specified in the genome. The term *development* and the term *innate* are thus similar, but not identical. *Development* (as opposed to *innate*) emphasized complex behaviors, thought to be under genetic control, that unfolded well after birth rather than those existing at birth (innate behavior). *Learning* encompassed processes that depended on explicit experience producing long-lasting changes in the organism.

Modern writers in all the aforementioned fields agree that behavior unfolds under the control of both a genetic blueprint and the environment, and that the debate now centers on the precise nature of the interplay between the two. Using the development-learning terminology, four alternatives can be conceptualized, as illustrated in figure 15.1. The first two are not interactionist accounts, whereas the last two can be described in this way.

Development and learning can be thought of as completely separable processes (figure 15.1a). Development follows a maturational course guided by a genetic blueprint, and learning neither follows from nor leads to changes in the preestablished course of development. Alternatively, they can be thought of as two processes so inseparable that they cannot be pulled apart even conceptually (figure 15.1b).

More commonly, development and learning are thought of as separate and distinguishable processes that interact in one way or another (figure 15.1c–d). Developmental psychologists, neuroscientists, and neurobiologists largely agree that the interactionist view is the correct one (Bonhoeffer and Shatz 1998; Carey 1985; Doupe 1998;

Doupe and Kuhl 1999; Fanselow and Rudy 1998; Gopnik and Meltzoff 1997; Karmiloff-Smith 1995; Kuhl 1994; Marler 1990, 1997). At issue, however, is exactly how the two systems interact, and particularly whether the interaction between development and learning is bidirectional.

Among the interactionist views, one model is that development enables learning, but that learning does not change the course of development, which unfolds more or less on its own timetable (figure 15.1c). Learning is seen as capitalizing on the achievements of development and as unable to occur unless a certain level of development has been achieved. The interaction is unidirectional, however. Development is not impacted by learning. In classical developmental psychology, this position is closest to the view of Piaget (1954). In modern neurobiology, the notion that there are "constraints" on learning, that development both prepares the organism and sets limits on learning, is consistent with this model (see Doupe 1998 and Marler 1974 for the case of birdsong). Greenough and Black's (1992) "experience-expectant" plasticity, wherein changes in neural development are thought to precede and prepare an organism to react to a reliably present environmental stimulus, provides a detailed example of this model. In each of these cases, development is conceived of as both enabling and limiting learning, but learning does not alter the course of development.

There is an alternative interactionist view. This model describes development and learning as mutually affecting one another (figure 15.1d). Development enables and even prompts learning, and learning in turn advances development. This view is closest to that developed by Vygotsky (1979). Vygotsky's theory, the "zone of proximal development" (ZPD), described development at two levels. One was the infant's actual developmental level, the level already achieved. The second was the level that was just within reach. The ZPD was the difference between the two. In Vygotsky's view, environmental stimulation slightly in advance of current development (in the ZPD) resulted in learning, and when this occurred, learning prompted development. Recent theories proposed by developmental psychologists to account for a wide variety of cognitive and linguistic tasks converge on the point that there is mutual interaction between development and learning (Carey 1985; Gopnik and Meltzoff 1997; Karmiloff-Smith 1991).

In linguistic theory, Chomsky's classic view, that the growth of language is largely determined by a maturational process, fits model C. Experience plays a role, but it is seen as triggering prespecified options, or as setting innately determined parameters (Chomsky 1981). The data reviewed here at the phonetic level of language come closer to the mutual interaction of model D. In the model of speech development I will describe, language input plays more than a triggering role in the process. Language input is mapped in a complex process that appears to code its subtle details.

Input thus goes beyond setting the parameters of prespecified options. Moreover, early mapping of the perceptual regularities of language input is argued to allow infants to recognize words and phrases, thus advancing development.

In summary, there is a great deal of support for interactionist views (models C and D) over noninteractionist views (models A and B). Although the relations between learning and development may differ across species and systems, there is an emerging consensus across diverse disciplines including neurobiology, psychology, linguistics, and neuroscience that development and learning are not independent entities. Both birdsong and speech fall clearly on the interactionist side. However, the form of the interaction remains a question, with a cutting-edge issue being whether (and how) learning can alter development. The model I will propose here on the basis of recent research on speech development goes some distance toward addressing this issue.

Explanations for Developmental Change in Speech

One of the puzzles in language development is to explain the orderly transitions that all infants go through. Infants the world over achieve certain milestones in linguistic development at roughly the same time, regardless of the language they are exposed to. Moreover, developmental change can also include cases in which infants' early skills exceed their later ones. Explaining these transitions is one of the major goals of developmental linguistic theory.

One of these transitions occurs in speech perception. At birth, infants discern differences between all the phonetic units used in the world's languages (Eimas, Miller, and Jusczyk 1987). All infants show these universal skills, regardless of the language environment in which they are being raised. Data on nonhuman animals' perception of speech suggest that the ability to partition the basic building blocks of speech is deeply rooted in our evolutionary history (Kuhl 1991a).

When do infants from different cultures begin to diverge in their perceptual abilities? Infants' initial language-universal perceptual abilities are highly constrained just one year later. By the end of the first year, infants fail to discriminate foreign-language contrasts they once discriminated (Werker and Tees 1984), resembling the adult pattern. Adults often find it difficult to perceive differences between sounds not used to distinguish words in their native language. Adult native speakers of Japanese, for example, have great difficulty discriminating American English /r/ and /l/ (Best 1993; Strange 1995), and American English listeners have great difficulty hearing the difference between Spanish /b/ and /p/ (Abramson and Lisker 1970).

Infants' abilities change over a 3-month period. A recent study we have just completed in Japan shows, for example, that at 7 months of age Japanese infants respond

to the /r-l/ distinction and are as accurate in perceiving it as American 7-month-old infants. By 10 months. Japanese infants no longer demonstrate this ability, even though American infants at that same age have become even better at discriminating the two sounds (Kuhl, Kiritani, et al. 1997).

A similar transition occurs in speech production. Regardless of culture, all infants show a universal progression in the development of speech which encompasses five distinct phases: *cooing* (1–4 months), in which infants produce sounds that resemble vowels, *canonical babbling* (5–10 months), during which infants produce strings of consonant-vowel syllables, such as "babababa" or "mamamama," *first words* (10–15 months), wherein infants use a consistent phonetic form to refer to an object, *two-word utterances* (18–24 months), in which two words are combined in a meaningful way, and *meaningful speech* (15 months and beyond), in which infants produce both babbling and meaningful speech to produce long intonated utterances (Ferguson, Menn, and Stoel-Gammon 1992). Interestingly, deaf infants exposed to a natural sign language, such as American Sign Language (ASL), are purported to follow the same progression using a visual-manual mode of communication (Petitto and Marentette 1991).

Although infants across cultures begin life producing a universal set of utterances that cannot be distinguished, their utterances soon begin to diverge, reflecting the influence of the ambient language they are listening to. By the end of the first year of life, the utterances of infants reared in different countries begin to be separable; infants show distinct patterns of speech production, both in the prosodic (intonational patterns) and phonetic aspects of language, that are unique to the culture in which they are being raised (de Boysson-Bardies 1993). In adulthood, these distinctive speech motor patterns that we initially learned contribute to our "accents" when attempting to speak another language (Flege 1988).

The transitions in speech perception and production from a pattern that is initially universal across languages to one that is highly specific to one particular language present one of the most intriguing problems in language acquisition. What causes the transition? We know that it is not simply maturational change. In the absence of natural language input, as in the case of socially isolated children (Curtiss 1977) or deaf children not exposed to manual sign language and unable to hear oral language, full-blown linguistic skills do not develop. Linguistic input and interaction, provided early in life, appear to be necessary.

The thesis developed here at the phonetic level is that linguistic experience produces a special kind of developmental change. Language input alters the brain's processing of the signal, resulting in the creation of complex mental maps. The mapping "warps" underlying dimensions, altering perception in a way that high-

lights distinctive categories. This mapping is not like traditional psychological learning. In the psychological literature, learning involves explicit teaching. In the Skinnerian tradition (1957), for example, this kind of learning depended on the presence of external reinforcement (specific contingencies that rewarded certain responses), feedback about the correctness of the response, and step-by-step shaping of the response. The kind of learning reflected in language is just the opposite. Although it does depend on external information from the environment (language input), it does not require explicit teaching or reinforcement contingencies. Given exposure to language in a normal and socially interactive environment, language learning occurs, and the knowledge gained about a specific language is long-lasting and difficult to undo.

Language Experience Alters Perception

The thesis developed here for the phonetic level of language is that ambient language experience produces a "mapping" that alters perception. A research finding that helps explain how this occurs is called the "perceptual magnet effect." It is observed when tokens perceived as exceptionally good representatives of a phonetic category ("prototypes") are used in tests of speech perception (Kuhl 1991b). The notion that categories have "prototypes" stems from cognitive psychology. Findings in that field show that the members of common categories (like the category *bird* or *dog*) are not equal. An ostrich is not as representative of the category *bird* as is a robin; a Terrier is not as representative of the category *dog* as is a Collie. These "prototypes" or best instances of categories are easier to remember, result in shorter reaction times when identified, and are often preferred in tests that tap our favorite instances of categories (Rosch 1977). Because of this literature, we were motivated to test the concept that phonetic categories had "prototypes" or best instances, in the early 1980s.

Our results demonstrated that phonetic prototypes did exist (Grieser and Kuhl 1989; Kuhl 1991b), that they differed in speakers of different languages (Kuhl 1992; Näätänen et al. 1997; Willerman and Kuhl 1996), and that phonetic prototypes function like perceptual magnets for other sounds in the category (Kuhl 1991b). When listeners hear a phonetic prototype and attempt to discriminate it from sounds that surround it in acoustic space, the prototype displays an attractor effect on the surrounding sounds (figure 15.2). It perceptually pulls other members of the category toward it, making it difficult to hear differences between the prototype and surrounding stimuli. Poor instances from the category (nonprototypes) do not function in this way. A variety of experimental tasks produce this result with both consonants and vowels (Iverson and Kuhl 1995, 1996; Sussman and Lauckner-Morano 1995).
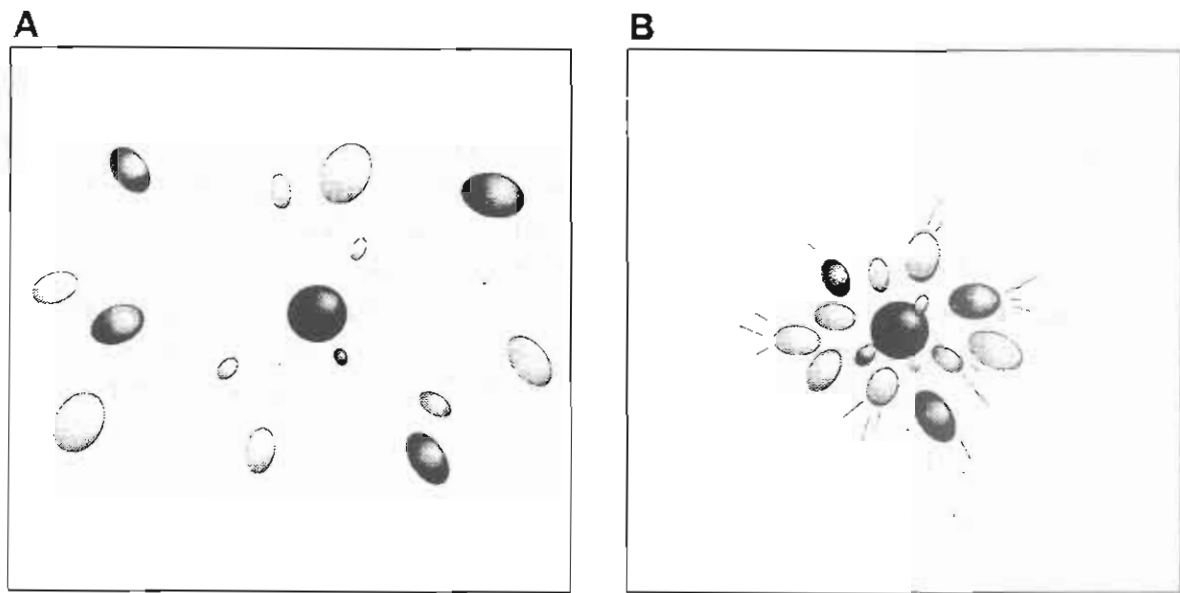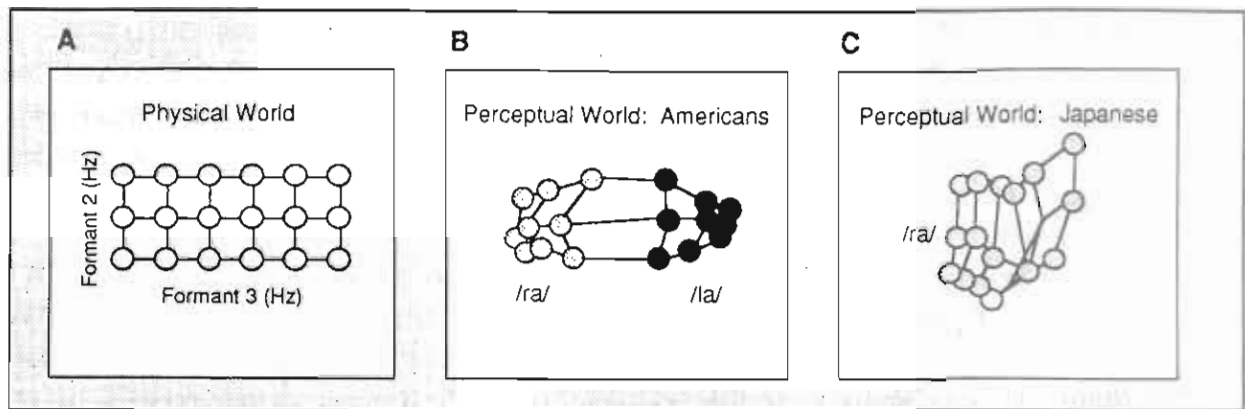
Figure 15.2
The perceptual magnet effect. When a variety of sounds in a category surround the category prototype (A), they are perceptually drawn toward the prototype (B). The prototype appears to function like a magnet for other stimuli in the category.

Other studies confirm listeners' skills in identifying phonetic prototypes and show that they are language-specific (Kuhl 1992; Miller 1994; Willerman and Kuhl 1996).

Developmental tests revealed that the perceptual magnet effect was exhibited by six-month-old infants for the sounds of their native language (Kuhl 1991b). In later studies, cross-language experiments demonstrated that the magnet effect is the product of linguistic experience (Kuhl et al. 1992). In the cross-language experiment, infants in the United States and Sweden were tested. The infants from both countries were tested with two vowel prototypes, an American English vowel prototype, /i/ (as in "peep"), and a Swedish vowel prototype, /y/ (as in "fye"). The results demonstrated that the perceptual magnet effect in six-month-old infants was influenced by exposure to a particular language. American infants demonstrated the magnet effect only for the American English /i/; they treated the Swedish /y/ like a nonprototype. Swedish infants showed the opposite pattern, demonstrating the magnet effect for the Swedish /y/ and treating the American English /i/ as a nonprototype. This is the youngest age at which language experience has been shown to affect phonetic perception.

The perceptual magnet effect thus occurs prior to word learning. What this means is that in the absence of formal language understanding or use—before infants utter

Figure 15.3
Physical (acoustic) versus perceptual distance. Consonant tokens of /r/ and /l/ were generated to be equally distant from one another in acoustic space (A). However, American listeners perceive perceptual space as "shrunk" near the best instances of /r/ (gray dots) and /l/ (black dots) and "stretched" at the boundary between the two (B). Japanese listeners' perceptual world differs dramatically; neither magnet effects nor a boundary between the two categories are seen (C).

or understand their first words—infants' perceptual systems strongly conform to the characteristics of the ambient language. We previously believed that word learning caused infants to recognize that phonetic changes that they could hear, such as the change that Japanese infants perceived between /r/ and /l/, did not change the meaning of a word in their language. This discovery was thought to cause the change in phonetic perception. We now know that just the opposite is true. Language input sculpts the brain to create a perceptual system that highlights the contrasts used in the language, while deemphasizing those that are not, and this happens prior to word learning. The change in phonetic perception thus assists word learning, rather than the reverse.

Further tests on adults suggested that the magnet effect distorted perception to highlight sound contrasts in the native language. Studies on the perception of the phonetic units /r/ and /l/ as in the words "rake" and "lake," illustrate this point. The /r-l/ distinction is one notoriously difficult for Japanese speakers and our studies sought to determine how adults from different cultures perceived these two sounds. To conduct the study, we used computer-synthesized syllables beginning with /r/ and /l/, spacing them at equal physical intervals in a two-dimensional acoustic grid (Iverson and Kuhl 1996) (figure 15.3a). American listeners identified each syllable as beginning with either /r/ or /l/, rated its category goodness, and estimated the perceived similarity for all possible pairs of stimuli using a scale from "1" (very dissimilar) to "7" (very similar). Similarity ratings were scaled using multidimensional

scaling (MDS) techniques. The results revealed that perception distorts physical space. The physical (acoustic) differences between pairs of stimuli were equal; however, perceived distance was "warped" (figure 15.3b). The perceptual space around the best /r/ and the best /l/ was greatly reduced, as predicted by the perceptual magnet effect, while the space near the boundary between the two categories was expanded.

This experiment has now been done using Japanese monolingual listeners (Kuhl, Yamada, Tohkura, Iverson, and Stevens, submitted) and the results show a strong contrast in the way the /r-l/ stimuli are perceived by American and Japanese adults (figure 15.3c). Japanese adults hear almost all the sounds as /r/; there is no /l/ in Japanese. More striking is the complete absence of magnet and boundary effects in the Japanese MDS solution. The results suggest that linguistic experience induces the formation of perceptual maps specifying the perceived distances between stimuli. These maps increase internal category cohesion while maximizing the distinction between categories. The critical point for theory is that neither group perceives the physical reality, the actual physical differences between the sounds. For each language group, experience has altered perception to create a language-specific map of auditory similarities and differences, one that highlights the sound contrasts of the speaker's native language. These mental maps for speech are the front-end of the language mechanism. In this sense, they point infants in the direction needed to focus on the aspects of the acoustic signal that will separate categories in their own native language. They provide a kind of attentional network that may function as a highly tuned filter for language. Such a network would promote semantic and syntactic analysis.

The theoretical position developed here is that the mental maps for speech are developed early in infancy, prior to the development of word acquisition. The mapping of phonetic information is seen as enabling infants to recognize word candidates. For example, our work shows that Japanese infants fail to discriminate American English /r/ from /l/ at 10 months of age, though they did so perfectly well at 7 months of age (Kuhl et al. submitted). This is argued to assist Japanese infants in word recognition. The collapsing of /r/ and /l/ into a single category makes it possible for Japanese infants to perceive their parents' productions of /r/-like and /l/-like sounds as one entity at 10 months, when the process of word acquisition begins. If they did not do so, it would presumably make it more difficult to relate sound patterns to objects and events.

The view that phonetic mapping enhances the recognition of higher-order units is supported by data showing that slightly later in development infants use information about phonetic units to recognize wordlike forms. Work by Jusczyk and his col-

leagues shows that by 9 months of age, infants prefer word patterns that are typical of the native language, which requires recognition of native-language phonetic units (Jusczyk et al. 1993). Infants have also been shown to be capable of learning the statistical probabilities of sound combinations contained in artificial words (Saffran, Aslin, and Newport 1996). Infants' mapping at the phonetic level is thus seen as assisting infants in "chunking" the sound stream into higher-order units, suggesting that "learning" promotes "development." Infants' discovery of statistical regularities in language input suggests new ways of conceiving of learning. During the first year of life, infants come to recognize the recurring properties of their native language and mentally store those properties in some form. This occurs in the absence of any formal instruction or reinforcement of the infant's behavior. In this sense, the learning that transpires is outside the realm of the historic versions of learning described by psychologists.

A Theory of Speech Development

Given these findings, how do we reconceptualize infants' innate predispositions as preparing them for experience? One view can be summarized as a three-step model of speech development, called the *native language magnet* (NLM) model (Kuhl 1994). NLM describes infants' initial state as well as changes brought about by experience with language (figure 15.4). The model demonstrates how infants' developing native-language speech representations might alter both speech perception and production. The same principles apply to both vowel and consonant perception; the example developed here is for vowels.

Phase 1 describes infants' initial abilities. At birth, infants distinguish all the phonetic distinctions of all languages of the world. This is illustrated by a hypothetical F1/F2 coordinate vowel space partitioned into categories (figure 15.4a). These divisions separate perceptually the vowels of all languages. According to NLM, infants' abilities at this stage do not depend on specific language experience. The boundaries initially structure perception in a phonetically relevant way. However, they are not due to a "language module." This notion is buttressed by the fact that nonhuman animals also hear speech distinctions, displaying abilities that were once thought to be exclusively human (Kuhl 1991a; see also, e.g., Dooling, Best, and Brown 1995; Kluender, Diehl, and Killeen 1987 for demonstrations of speech perception abilities in nonhuman mammals).

Phase 2 describes the vowel space at six months of age for infants reared in three very different language environments, Swedish, English, and Japanese (figure 15.4b).
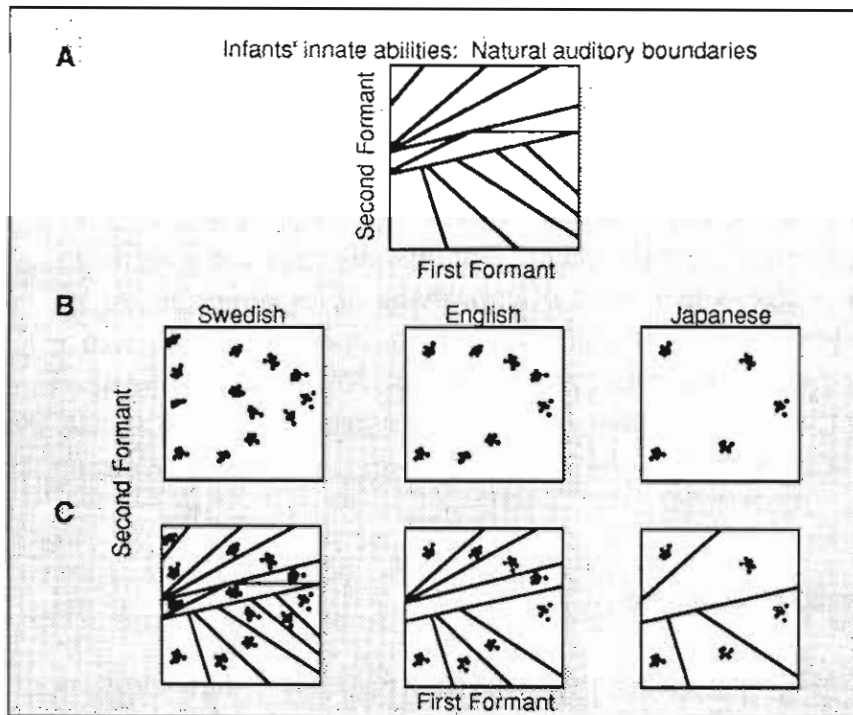
Figure 15.4
The native language magnet (NLM) model. (A) At birth, infants perceptually partition the acoustic space underlying phonetic distinctions in a language-universal way. They are capable of discriminating all phonetically relevant differences in the world's languages. (B) By 6 months of age, infants reared in different linguistic environments show an effect of language experience. Infants store incoming vowel information in memory in some form. The resulting representations (shown by the dots) are language-specific, and reflect the distributional properties of vowels in the three different languages. (C) After language-specific magnet effects appear, some of the natural boundaries that existed at birth "disappear." Infants now fail to discriminate foreign-language contrasts they once discriminated.

By six months of age, infants show more than the ability to perceptually separate all phonetic categories, as shown in phase 1. By six months, our informal calculations indicate that infants have heard many vowels. The distributional properties of vowels heard by infants being raised in Sweden, America, and Japan differ. According to NLM, infants mentally store this information in some way. As shown in figure 15.4b, infants' stored representations differ, reflecting the distributional differences between their languages. In each case, linguistic experience has produced stored representations that reflect the vowel system of the ambient language. Language-specific magnet effects, produced by the stored representations, are exhibited by infants at this stage.

An interesting question about phase 2 magnet effects is: How much language input does it take to develop these effects, and is all language heard by the child

(including that from a radio or television) effective in producing this special kind of learning? This, in fact, was the question President Clinton asked me at the White House Conference on Early Learning and the Brain, in April of 1997. My answer was that at present we have little data suggesting how much language input it takes to show these effects, and we do not know whether language from a disembodied source (TV, radio) would be sufficient to produce it. By six months of age, the earliest age at which we have evidence of magnet effects, our estimates suggest that infants have heard thousands of instances of vowels in *en face* communication with their parents (Kuhl 1994), but we do not know what amount is necessary. These are exciting questions for future research, and we have studies underway at present that are providing some interesting information on these issues.

Phase 3 shows how magnet effects recursively alter the initial state of speech perception, and affect the processing of foreign-language stimuli. Magnet effects cause certain perceptual distinctions to be minimized (those near the magnet attractors) while others are maximized (those near the boundaries between two magnets). The consequence is that some of the boundaries that initially divided the space "disappear" as the perceptual space is reconfigured to incorporate a language's particular magnet placement (figure 15.4c). Magnet effects functionally erase certain boundaries—those relevant to foreign but not native languages.

In phase 3, a perceptual space once characterized by basic "auditory cuts"—boundaries that divide all speech categories and ones demonstrated in nonhuman animals—has been replaced by a dramatically warped space, dominated by magnet effects, that completely restructure the space. It is at this phase that infants fail to discriminate foreign-language contrasts that were once discriminable. The mapping of incoming speech has altered which stimulus differences infants respond to, producing a language-specific listener for the first time.

A natural question arising from these data is what happens to infants exposed to two different languages throughout phase 3. We are only beginning to study this issue, but the theory predicts that infants will develop magnet effects for the sound categories of both languages. Interestingly, preliminary data from studies underway suggest that development of two sets of magnet effects is particularly likely when the two languages are spoken by different speakers (mother speaks one language, father speaks another). Presumably having each language spoken by a different speaker allows infants to separate perceptually the maps for the two languages.

Infants at six months of age have no awareness of the fact that sound units are used contrastively in language to name things. Yet the infant's perceptual system organizes itself to reflect language-specific phonetic categories. At the next stage in linguistic development, when infants acquire word meanings by relating sounds to

objects and events in the world, the language-specific mapping that has already occurred should greatly assist this process.

NLM theory offers an explanation for the developmental change observed in speech perception. A developing magnet pulls sounds that were once discriminable toward it, making them less discriminable. Magnet effects should therefore developmentally precede changes in infants' perception of foreign-language contrasts; preliminary data indicate that they do (Werker and Polka 1993). The magnet effect also helps account for the results of studies on the perception of sounds from a foreign language by adults (Best 1993; Flege 1993). For example, NLM theory may help explain Japanese listeners' difficulty with American /r/ and /l/. The magnet effect for the Japanese /r/ category prototype (which is neither American /r/ nor /l/) will attract both /r/ and /l/, making the two sounds difficult for native-speaking Japanese people to discriminate (Kuhl et al. submitted). NLM theory argues that early experience establishes a complex perceptual network through which language passes. On this view, one's primary language and the map that results from early experience will determine how other languages are perceived.

Reinterpreting "Critical Periods"

Evidence for the interaction between genetic programming and environmental stimulation is nowhere more predominant than in the literature on critical periods in learning (Marler 1970; Thorpe 1961). Critical periods are no longer viewed as strictly timed developmental processes with rigid cut-off periods that restrict learning to a specific time frame. Recent studies showing that learning can be stretched by a variety of factors has caused a shift in the terminology used to refer to this period. It is now understood that during "sensitive periods" exposure to specific kinds of information may be more effective than at other times, but that a variety of factors can alter the period of learning. Knudsen's work, for example, on the sound-localization system in the barn owl shows that the sensitive period for learning the auditory-visual map in the optic tectum can be altered by a variety of factors that either shorten or extend the learning period; the learning period closes much earlier, for instance, if experience occurs in a more natural environment (Knudsen and Brainerd 1995; Knudsen and Knudsen 1990).

The idea that sensitive periods define "windows of opportunity" for learning, during which environmental stimulation is highly effective in producing developmental change, remains well supported in both the human and the animal literature. Whether language or birdsong is the focus of inquiry, the ability to learn is not

equivalent over time (Doupe and Kuhl 1999). The question is: What causes changes in the ability to learn throughout the individual's lifespan?

The sensitive period denotes a process of learning that is constrained primarily by time, or other factors (hormones, etc.)—factors outside the learning process itself (see Doupe and Kuhl 1999). There is an alternative possibility suggested by studies on speech. Later learning may be limited by the fact that learning itself alters the brain, and learning may produce a kind of "interference" effect that impacts later learning. To illustrate, if NLM's claim that learning involves the creation of mental maps for speech is true, this would mean that learning "commits" neural structure in some way. According to the model, ongoing processing of speech is affected by this neural structure, and future learning is affected as well. The mechanisms governing an organism's general ability to learn may not have changed. Rather, initial learning may result in a structure that reflects environmental input, and once committed, the learned structure may interfere with the processing of information that does not conform to the learned pattern. On this account, initial learning can alter future learning independent of a strictly timed period.

On this *interference* account, plasticity would be governed from a statistical standpoint. When additional input does not cause the overall statistical distribution to change substantially, the organism becomes less sensitive to input. Hypothetically, for instance, the infants' representation of the vowel /a/ might not change when the one-millionth token of the vowel /a/ is heard. Plasticity might thus be independent of time, but be dependent on the amount and variability provided by experience. At some time in the lifetime of an organism, one could conceive of a point where new input no longer alters the underlying distribution, and this could, at least in principle, reduce the system's "plasticity."

The interference view may account for some aspects of second language learning. When acquiring a second language, certain phonetic distinctions are notoriously difficult to master both in speech perception and production. Take the case of the /r-l/ distinction for native speakers of Japanese. Hearing and producing the distinction is very difficult for native speakers of Japanese (Goto 1971; Miyawaki et al. 1975; Yamada and Tohkura 1992). According to NLM, this is because exposure to Japanese early in life altered the Japanese infant's perceptual system, resulting in magnet effects for the Japanese phoneme /r/, but not for American English /r/ or American English /l/. Once in place, the magnet effects appropriate for Japanese would not make it easy to process American English. Both American English /r/ and /l/ would be assimilated to Japanese /r/ (see also Best 1993). Teaching Japanese adults to distinguish /r/ and /l/ may require the use of stimuli that avoid the interference effect.

A second language learned later in life may require separation between the two systems to avoid interference. Data gathered using fMRI techniques indicate that adult bilinguals who learned both languages early in life activate overlapping regions of the brain when processing the two languages, whereas those who learned the second language later in life activate two distinct regions of the brain for the two languages (Kim, Relkin, Lee, and Hirsch 1997). This is consistent with the idea that the brain's processing of a primary language can interfere with the second language. This problem is avoided if both are learned early in development.

The general thesis is that acquiring new phonetic categories as adults is difficult because the brain's mental maps for speech, formed on the basis of the primary language, are incompatible with those required for the new language; hence interference results. Early in life, interference effects are minimal and new categories can be acquired because input continues to revise the statistical distribution. As mentioned earlier, limited evidence suggests that infants exposed to two languages do much better if each parent speaks one of the two languages, rather both parents speaking both languages. This may be the case because it is easier to map two different sets of phonetic categories (one for each of the two languages) if there is some way to keep them perceptually separate. Males and females produce speech in different frequency ranges, and this could make it easier to maintain separation.

These two factors—a maturationally defined temporal window and initial learning that makes later learning more difficult—could both be operating to produce constraints on learning a second language later in life. If a maturational process induces "readiness" at a particular time, input that misses this timing could reduce learning. At the same time, an "interference" factor might provide an independent mechanism that contributes to the difficulty in readily learning a second language in adulthood.

Vocal Learning

As Peter Marler has so clearly described, vocal learning—an organism's dependence on auditory input and feedback to acquire a vocal repertoire—is not common among mammals (see Snowdon and Hausberger 1997 for exceptions), but it is exhibited strikingly in songbirds. In certain songbirds, and in humans, young members of the species not only learn the perceptual properties of their conspecific communicative signals but become proficient producers of those signals. A great deal of research on birds (Marler 1990, 1997) and infants (Stoel-Gammon and Otomo 1986; Oller, Wieman, Doyle, and Ross 1976) has shown that input is essential to the development of vocalizations. Deaf infants, for example, do not babble normally (Oller and Mac-

Neilage 1983), nor do deafened birds (Konishi 1965; Nottebohm 1967). In the case of humans, the learned motor patterns that underlie speech become difficult to alter. Speakers who learn a second language later in life, for example, produce it with an "accent" typical of their primary language (Flege 1993). Most speakers of a second language would like to speak like a native speaker, without an accent, but this is difficult to do.

When do we adopt the indelible speech patterns that will mark us as native speakers of a particular language for our entire lives? Developmental studies suggest that by one year of age, language-specific patterns of speech production appear in infants' spontaneous utterances (de Boysson-Bardies 1993; Vihman and de Boysson-Bardies 1994). However, the fundamental capacity to reproduce the sound patterns one hears is in place much earlier. In a recent study, Kuhl and Meltzoff (1996) recorded infant utterances at 12, 16, and 20 weeks of age while the infants watched and listened to a video recording of a woman producing a vowel, either /a/, /i/, or /u/. Infants watched the video for five minutes on each of three consecutive days. Infants' utterances were analyzed both perceptually (phonetic transcription) and instrumentally (computerized spectrographic analysis).

The results showed that there was developmental change in infants' vowel productions between 12 and 20 weeks of age. The areas of vowel space occupied by infants' /a/, /i/, and /u/ vowels become progressively more tightly clustered at each age, and by 20 weeks, a "vowel triangle" typical of that produced in every language of the world had emerged in infants' own region of the vowel space (figure 15.5).
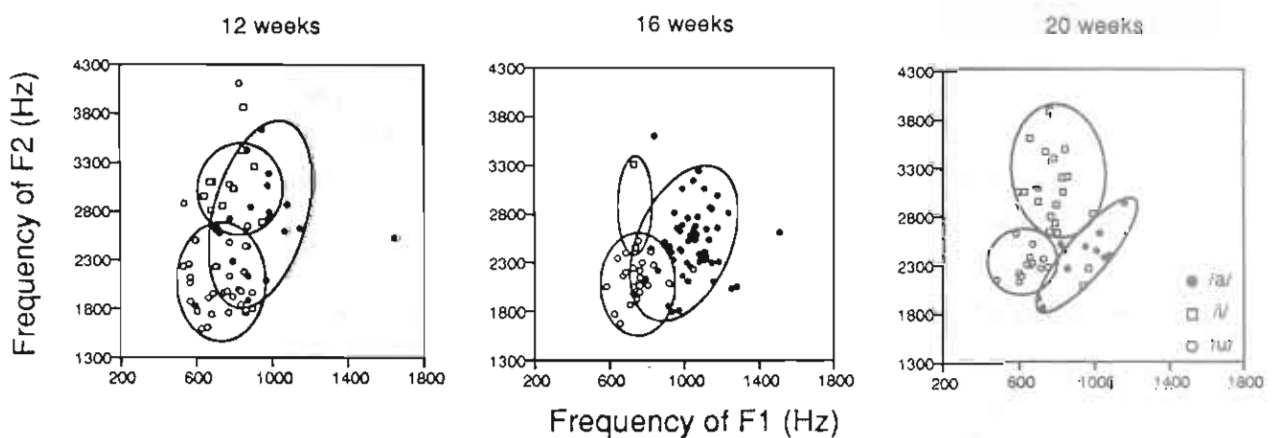
Figure 15.5
The location of /a/, /i/, and /u/ vowels produced by 12-, 16-, and 20-week-old infants. Infants' vowel productions show progressively tighter clustering in vowel space over the 8-week period and reflect differences between the three vowel categories seen in adults' productions.
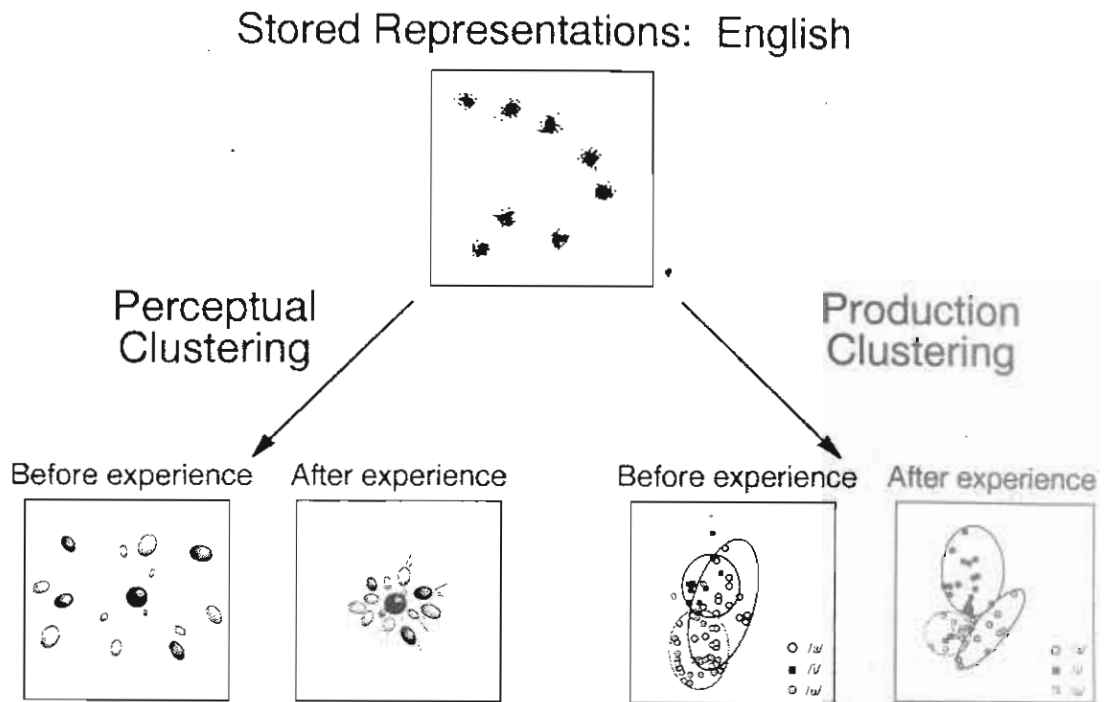
This suggested the possibility that infants were listening to language and attempting to vocally imitate the sound patterns they heard (Kuhl and Meltzoff 1996).

Direct evidence that infants were vocally imitating was also obtained in the study. By 20 weeks, infants were shown to reproduce the vowels they heard. Infants exposed to /a/ were more likely to produce /a/ than when exposed to either /i/ or /u/; similarly, infants exposed to either /i/ or /u/ were more likely to produce the vowel in that condition than when listening to either of the two alternate vowels. The total amount of exposure to a specific vowel in the laboratory was only 15 minutes, yet this was sufficient to influence infants' productions. If 15 minutes of laboratory exposure to a vowel is sufficient to influence infants' vocalizations, then listening to ambient language for weeks would be expected to provide a powerful influence on infants' production of speech. These data suggest that infants' stored representations of speech not only alter infant perception but alter production as well, serving as auditory patterns that guide motor production. Stored representations are thus viewed as the common cause for both the tighter clustering observed in infant vowel production and the tighter clustering observed in infant vowel perception (figure 15.6).

This pattern of learning and self-organization, in which perceptual patterns stored in memory serve as guides for production, is strikingly similar for birdsong (Doupe 1998; Doupe and Kuhl 1999), in visual-motor learning in which nonspeech oral movements like tongue protrusion and mouth opening are imitated (Meltzoff and Moore 1977, 1994), and in language involving both sign (Petitto and Marentette 1991) and speech (Kuhl and Meltzoff 1996). In each of these cases, perceptual experience establishes a representation that guides sensory-motor learning. In the case of infants and speech, perception affects production in the earliest stages of language learning, which reinforces the idea that the speech motor patterns of a specific language are formed very early in life. Once learned, motor patterns may also further development by altering the probability that infants will acquire words that contain items they are capable of producing (see, e.g., Vihman 1993).

The Role of Vision in Speech Perception: Polymodal Speech Representation

The link between perception and production can be seen in another experimental situation in speech, and there is some suggestion that this is mirrored in birdsong learning. Speech perception in adults is strongly affected by the sight of a talker's mouth movements during speech, indicating that our representational codes for speech contain both auditory and visual information. One of the most compelling examples of the polymodal nature of speech representations is auditory-visual illu-

Stored Representations: English

Perceptual Clustering

Production Clustering

Before experience   After experience   Before experience   After experience

Figure 15.6
Stored representations of native-language speech affect both speech perception, producing the perceptual clustering evidenced by the magnet effect, as well as speech production, producing the increased clustering seen in infants' vocalizations over time.

sions that result when discrepant information is sent to two separate modalities. One such illusion occurs when auditory information for /b/ is combined with visual information for /g/ (McGurk and MacDonald 1976; Green, Kuhl, Meltzoff, and Stevens 1991; Kuhl, Tsuzaki, Tohkura, and Meltzoff 1994; Massaro 1987). Perceivers report the phenomenal impression of an intermediate articulation (/da/ or /tha/) despite the fact that this information was not delivered by or to either sense modality. This is a robust phenomenon and is readily obtained even when the information from the two modalities comes from different speakers, such as when a male voice is combined with a female face (Green et al. 1991). In this case, there is no doubt that the auditory and visual signals do not belong together. Yet the illusion is still unavoidable—our perceptual systems combine the multimodal information (auditory and visual) to give a unified percept.

Young infants are also affected by visual information. Infants just 18–20 weeks old recognize auditory-visual correspondences for speech, akin to what we as adults

do when we lip-read; in these studies, infants looked longer at a face pronouncing a vowel that matched the vowel sound they heard, rather than at a mismatched face (Kuhl and Meltzoff 1982). Young infants demonstrate knowledge about both the auditory and visual information contained in speech, which supports the notion that infants' stored speech representations contain information of both kinds. Additional demonstrations of auditory-visual speech perception in infants suggest that there is a left-hemisphere involvement in the process (MacKain, Studdert-Kennedy, Spieker, and Stern 1983), and more recent data by Rosenblum, Schmuckler, and Johnson (1997) and Walton and Bower (1993) suggest that the ability to match auditory and visual speech is present in newborns.

Visual information thus plays a strong role in speech perception. Studies show that when listeners watch the face of the talker, perception of the message is greatly enhanced, in effect contributing the equivalent of a 20 dB boost (quite substantial) in the signal. This supports the view that speech in humans is polymodally represented. In birds, syllables from alien bird species' songs can be learned in the presence of the sight of a conspecific bird (Petrinovich and Baptista 1987). These visual effects suggest how strongly the visual-motor system is linked to perception in both birdsong and speech (see also Liberman 1993).

Nature of Language Input to the Child

Peter Marler's early studies (e.g., 1970) on birdsong learning highlighted the role of input by manipulating it. These studies showed the dramatic effects of input on song learning, as well as the birds' selectivity for certain kinds of signals. In the realm of language, attention is being increasingly devoted to understanding exactly how much and what kind of language input infants hear.

Estimates indicate that a typical listening day for a two year old includes 20,000–40,000 words (Chapman et al. 1992). Research has shown that the speech addressed to infants (often called "motherese" or "parentese") is unique: it has a characteristic prosodic structure that includes a higher pitch, a slower tempo, and exaggerated intonation contours, and it is syntactically and semantically simplified. Studies show that this speaking style is nearly universal in the speech of caretakers around the world, and that infants prefer it over other complex acoustic signals (Fernald 1985; Grieser and Kuhl 1989; Fernald and Kuhl 1987).

In recent studies, we have uncovered another modification made by parents when addressing infants that may be important to infant learning. We examined natural-language input at the phonetic level to infants in the United States, Russia, and
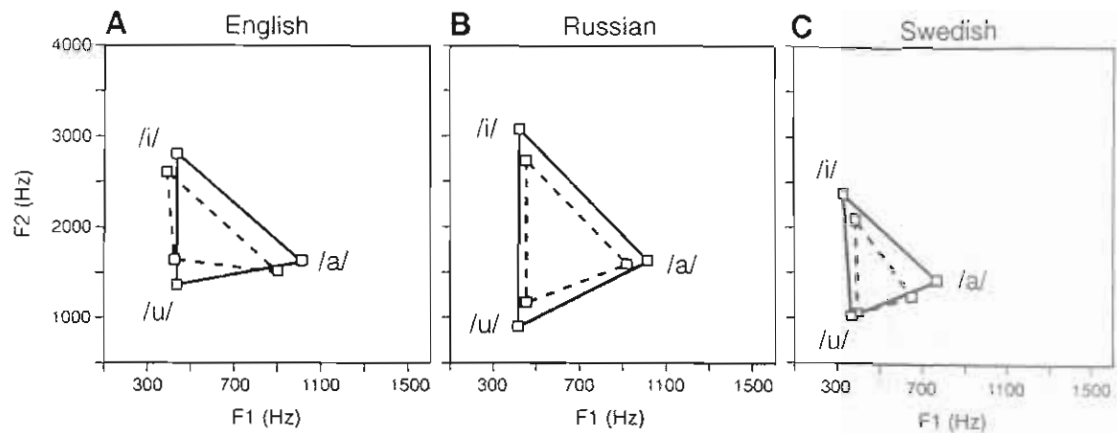
Figure 15.7
The vowel triangle of maternal speech directed toward infants (solid line) across three diverse languages shows a "stretching" relative to the adult-directed vowel triangle (dashed line), an effect that both makes vowels more discriminable and highlights the abstract features that infants must use to produce speech themselves.

Sweden (Kuhl, Andruski, et al. 1997). The study shows that across three very diverse languages, infant-directed speech exhibited a universal alteration of phonetic units when compared with adult-directed speech. Parents addressing their infants produced acoustically more extreme tokens of vowel sounds, resulting in a "stretching" of the acoustic space encompassing the vowel triangle (figure 15.7). A stretched vowel triangle not only makes speech more discriminable for infants, it also highlights critical spectral parameters that allow speech to be produced by the child. The results suggest that at the phonetic level of language, linguistic input to infants provides exceptionally well-specified information about the units that form the building blocks for words.

A stretched vowel space is not necessary from the standpoint of the infant's capacity to distinguish vowels. The formant frequency changes from adult-directed to infant-directed speech were substantial and would clearly be registered by the infant auditory system. Previous data on infants' capacities to discern subtle differences between vowels indicate that infants are capable of hearing differences a great deal smaller than those produced by mothers in the study (Kuhl 1991a).

If not required for infant discrimination, what function does a stretched vowel space serve? We hypothesized that stretching the vowel triangle could benefit infants in three ways. First, it increases the distance between vowels, making them more distinct from one another. In recent studies, language-delayed children showed substantial

improvements in measures of speech and language after treatment in a program in which they listened to speech altered by computer to exaggerate phonetic differences (Merzenich et al. 1996; Tallal et al. 1996). Normally developing infants may benefit similarly from the enhanced acoustic differences provided in infant-directed speech.

Second, to achieve the stretching, mothers produce vowels that go beyond those produced in typical adult conversation. From both an acoustic and articulatory perspective, these vowels are "hyperarticulated" (Lindblom 1990). Hyperarticulated vowels are perceived by adults as "better instances" of vowel categories (Iverson and Kuhl 1995; Johnson, Flemming, and Wright 1993), and laboratory tests show that when listening to good instances of phonetic categories, infants show greater phonetic categorization ability. The present study shows that hyperarticulated vowels are a part of infants' linguistic experience and raises the possibility that they play an important role in the development of infants' vowel categories.

Third, expanding the vowel triangle allows mothers to produce a greater variety of instances representing each vowel category without creating acoustic overlap between vowel categories. Greater variety may cause infants to attend to non-frequency-specific spectral dimensions that characterize a vowel category, rather than to any one particular set of frequencies that the mother uses to produce a vowel, and this may assist learning (see Lively, Logan, and Pisoni 1993). Converting the formant values to spectral features shows that infant-directed speech maximizes the featural contrast between vowels (Kuhl, Andruski et al. 1997). This is especially critical for infants because they cannot duplicate the absolute frequencies of adult speech—their vocal tracts are too small (Kent and Murray 1982). To speak, infants must reproduce the appropriate spectral dimensions in their own frequency range (Kuhl and Meltzoff 1996). Our recent study indicates that maternal language input emphasizes these dimensions. It is unknown whether adult birds tutor their infants using a special kind of vocal signal, the equivalent of "songbird parentese."

Brain Correlates

Since the classic reports of Broca (1861) and Wernicke (1874) on patients with language deficits typical of aphasia, we have known that the two hemispheres are not equal in the extent to which they subserve language, and this had in the past been reported for birds as well (though see Doupe and Kuhl 1999 few recent data). In the 1960s, behavioral studies on language processing in normal adults contributed additional evidence of the left-hemisphere specialization for language (Kimura 1961). Using a method known as the "dichotic listening task," researchers sent two differ-

ent speech signals simultaneously to the right and left ears and asked subjects to identify the stimuli they heard. The results demonstrated that people are significantly more accurate at identifying speech sent to the right ear than to the left ear, and that music showed a right-ear advantage; both were argued to be attributable to the increased strength of contralateral over ipsilateral fibers connecting the auditory input to the brain, and the fact that the left hemisphere (LH) of the brain was specialized for language, whereas the right hemisphere (RH) was specialized for music (Kimura 1961; Milner, Taylor, and Sperry 1968). The LH specialization was shown not only for spoken language but for individuals words, the consonants and vowel sounds contained in words, and even the phonetic features that formed the building blocks for phonemes (Studdert-Kennedy and Shankweiler 1970). These studies provided powerful evidence that the left hemisphere controlled language processing in adults, for units as small as the phonetic feature.

With the advent of modern neuroimaging techniques such as PET, positron emission tomography (Peterson, Fox, Snyder, and Raichle 1990; Posner, Peterson, Fox, and Raichle 1988), fMRI, functional magnetic resonance imaging, (Neville et al. 1998), ERP, event-related potentials (Neville, Mills, and Lawson 1992; Osterhout and Holcomb 1992), and MEG, magnetoencephalography (Näätänen et al. 1997), studies of the brain's organization of language have appeared in increasing numbers. It is early in this field, but there are some conclusions that take us beyond what early studies have shown regarding language and the brain. The new studies suggest, for example, that there is not one unified "language area" in the brain in which linguistic signals are processed. Different brain systems subserve different aspects of language processing, and the language-processing areas of the brain include many more regions than the classic Broca's and Wernicke's areas (see Binder 1994 for review).

The imaging studies support the dissociation for processing speech and nonspeech signals found in earlier behavioral studies. Zatorre, Evans, Meyer, and Gjedde (1992) used PET scans to examine phonetic as opposed to pitch processing. Subjects in the study had to judge the final consonant of the syllable in the phonetic task and the pitch of the syllable (high or low) in the pitch task. The results showed that phonetic processing engaged the LH whereas pitch processing engaged the RH. This dissociation between the phonetic and nonphonetic processing of auditory dimensions is mirrored in studies using vowels at different pitches and MEG, even though the lateralization of vowels was less strong than that shown for consonants, as might be expected from behavioral studies (Poeppel et al. 1997).

Studies of deaf individuals who communicate using sign language provide an interesting test for the hypothesis that the LH handles language processing. Because visual-spatial information is typically thought to involve right-hemisphere analysis,

studies of sign-language processing by deaf speakers provided a method for examining whether language involves left-hemisphere analysis in speakers whose language is visual-spatial, a signal normally associated with right-hemisphere processing.

Studies of deaf individuals who have sustained cerebral damage and exhibit aphasia indicate that in such individuals the left hemisphere rather than right hemisphere is impaired (Bellugi, Poizner, and Klima 1989; Klima, Bellugi, and Poizner 1988). Studies using event-related potential methods show that in deaf individuals the left-hemisphere regions normally used for spoken language processing are active during processing of signed language signals (Neville, Mills, and Lawson 1992).

Given these results, an important question from the standpoint of development is when the left-hemisphere becomes dominant in the processing of linguistic information. Lenneberg (1967) hypothesized that the two hemispheres are equipotential for language until approximately two years of age, at which time left-hemisphere dominance begins to develop and continues until puberty. However, this theory was contradicted by data suggesting that children with LH versus RH damage at an early age displayed different deficits in language abilities; early LH damage affected language abilities more than early RH damage (Witelson 1987).

Behavioral studies established the right-ear advantage (REA) for speech for verbal stimuli and the left-ear advantage (LEA) for musical and environmental sounds using the dichotic listening task in children as young as 2.5 years of age (Bever 1971; Kimura 1963). But what of infants? Glanville, Best, and Levenson (1977) reported an REA for speech contrasts and a LEA for musical sounds using a cardiac orienting procedure in 3-month-olds. There have been only two studies providing data on infants' differential responses to speech and music at ages younger than two months. In one study, infants' discriminative capacities for speech and music were examined in 2-, 3-, and 4-month-old infants using a cardiac-orienting procedure. The results demonstrated a REA for speech discrimination in 3-month and 4-month-old infants, but not in 2-month-old infants. In addition, a LEA was shown for musical-note discrimination at all three ages. Finally, tests on 2-month-olds conducted by Vargha-Khadem and Corballis (1979) showed infants discriminated speech contrasts equally with both hemispheres. In other words, the results suggest that speech may not be lateralized in 2-month-old infants, but becomes so by 4 months of age.

This issue of the onset of laterality and the extent to which it depends on linguistic experience may be resolved with future research using techniques that can be used throughout the lifespan. One such technique is the mismatched negativity (MMN), a component of the auditorily evoked event-related potential (ERP). The MMN is automatically elicited by a discriminable change in a repetitive sound pattern and its

generation appears to represent the detection of a change in the neural sensory-memory representation established by the repeated stimulus (see Näätänen 1990, 1992 for review). MMNs to a change in a speech stimulus have been well documented in adult listeners (see, e.g., Kraus et al. 1992). The MMN has been established in children (Kraus et al. 1993) and most recently in infants (Cheour-Luhtanen et al. 1995; Kuhl 1998b; Pang et al. 1998). In adult ERP studies, a left-hemisphere asymmetry is evident (Pang et al. 1998), but it has not been observed in newborns (Cheour-Luhtanen et al. 1995).

In summary, there is no strong evidence at present that the bias toward left-hemisphere processing for language is present at birth. The data suggest that it may take experience with linguistically patterned information to produce the left-hemisphere specialization. Thus there is support for a specialization for language in infancy, but one that develops, rather than one that exists at birth. Moreover, the input that is eventually lateralized to the left hemisphere can be either speech or sign, indicating that it is the linguistic or communicative significance of the signals, rather than their specific form, that accounts for the specialization.

Conclusions

As first suggested by Peter Marler (1970), birdsong and speech share a number of striking features: (i) evidence of innate perceptual predispositions that aid the acquisition of vocal communication; (ii) dramatic effects during sensitive periods to experience with conspecific communicative signals both in perception and production; (iii) vocal learning's dependence on auditory input; (iv) visual effects that enhance processing indicating perceptual-motor links and polymodal mapping of information; and (v) early brain correlates supporting the control of communication. Theories will be furthered by interactions between these two fields. In humans, theoretical progress is being made with detailed results of behavioral experiments on development. In birds, information regarding the neural substrates are providing detailed information, thus making the two fields highly complimentary.

To summarize these points: research has shown that in the first year of life infants learn a great deal about the perceptual characteristics of their native language, and this subsequently alters their perception and production of speech. According to the native language magnet model, perceptual learning early in life results in the formation of stored representations that capture native-language regularities. The theory emphasizes the role of linguistic input. Input does not act like a trigger for innately stored information. Rather, it is mapped in such a way as to "warp" the underlying

acoustic space. Stored representations act like perceptual magnets for similar patterns of sound, resulting in maps that specify perceived distances between sounds and create categories. The map shrinks perceptual distances near a category's most typical instances and stretches perceptual distances between categories. Perceptual maps differ in adults who speak different languages and are polymodally mapped, containing auditory, visual, and motor information. The magnet effects and the mental maps they produce help explain how native-language speech develops, as well as our relative inability as adults to readily acquire a foreign language.

Acknowledgment

References

Abramson, A. S. and Lisker, L. (1970) Discriminability along the voicing continuum: Cross-language tests. In *Proceedings of the Sixth International Congress of Phonetic Sciences, Prague 1967*, pp. 569–573. Prague: Academia.

Bates, E. and Elman, J. (1996) Learning rediscovered. *Science* 274:1849–1850.

Bateson, P. (1991) Are there principles of behavioral development? In *The Development and Integration of Behaviour*, ed. by P. Bateson, pp. 19–39. Cambridge: Cambridge University Press.

Bellugi, U., Poizner, H., and Klima, E. S. (1989) Language, modality, and the brain. *Trends Neurosci.* 12:380–388.

Berkeley, G. (1709) *An Essay Toward a New Theory of Vision*. Dublin: Pepyat.

Best, C. T. (1993) Language-specific changes in non-native speech perception: A window on early phonological development. In *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, ed. by B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, and J. Morton, pp. 289–304. Dordrecht, Netherlands: Kluwer Academic Publishers.

Bever, T. G. (1971) The nature of cerebral dominance in speech behavior of the child and adult. In *Language Acquisition: Models and Methods*, ed. by R. Huxley and E. Ingram. London: Academic Press.

Binder, J. R. and Rao, S. M. (1994) Human brain mapping with functional magnetic resonance imaging. In *Localization and Neuroimaging in Neuropsychology*, ed. by A. Kertesz, pp. 185–212. San Diego: Academic Press.

Bonhoeffer, T. and Shatz, C. J. (1998) Neurotrophins and visual system plasticity. In *Mechanistic Relationships between Development and Learning*, ed. by T. J. Carew, R. Menzel, and C. J. Shatz, pp. 93–112. New York: Wiley.

Broca, P. (1861) Nouvelle observation d'aphemie produite par une lesion de la motie posterieure des deuxieme et troiseme circonvolutions frontales. *Bulletin de la Societe d'Anatomique* 6:398–407.

Cajal, R. S. (1906 [1967]) The structure and connections of neurons. In *Nobel Lectures in Physiology or Medicine 1901–1921*, pp. 220–253. New York: Elsevier.

Carew, T. J., Menzel, R., and Shatz, C. J. (1998) *Mechanistic Relationships between Development and Learning*. New York: Wiley.

Carey, S. (1985) *Conceptual Change in Childhood.* Cambridge, MA: MIT Press.

Chapman, R. S., Streim, N. W., Crais, E. R., Salmon, D., Strand, E. A., and Negri, N. A. (1992) Child talk: Assumptions of a developmental process model for early language learning. In *Processes in Language Acquisition and Disorders,* ed. by R. S. Chapman, pp. 3–19. St. Louis, MO: Mosby Year Book.

Cheour-Luhtanen, M., Alho Kujala, K. T., Sainio Reinikainen, K. K., Renlund, M., Aaltonen, O., Eerola, O., and Näätänen, R. (1995) Mismatch negativity indicates vowel discrimination in newborns. *Hearing Research* 82:53–58.

Chomsky, N. (1957) A review of B. F. Skinner's *Verbal Behavior. Language* 35:26–58.

Chomsky, N. (1965) *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.

Chomsky, N. (1981) *Rules and Representations.* New York: Columbia University Press.

Curtiss, S. (1977) *Genie: A Psycholinguistic Study of a Modern Day "Wild Child."* New York: Academic.

de Boysson-Bardies, B. (1993) Ontogeny of language-specific syllabic productions. In *Developmental Neurocognition: Speech and Face Processing in the First Year of Life,* ed. by B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, and J. Morton, pp 353–363. Dordrecht, Netherlands: Kluwer Academic Publishers.

Dooling, R. J., Best, C. T., and Brown, S. D. (1995) Discrimination of synthetic full-formant and sine-wave /ra-la/ continua by budgerigars (*Melopsittacus undulatus*) and zebra finches (*Taeniopygia guttata*). *J. Acoust. Soc. Am.* 97:1839–1846.

Doupe, A. (1998) Development and learning in the birdsong system: Are there shared mechanisms? In *Mechanistic Relationships between Development and Learning,* ed. by T. J. Carew, R. Menzel, and C. J. Shatz, pp. 29–52. New York: Wiley.

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

Eimas, P. D., Miller, J. L., and Jusczyk, P. W. (1987) On infant speech perception and the acquisition of language. In *Categorical Perception: The Groundwork of Cognition,* ed. by S. Harnad, pp. 161–195. New York: Cambridge University Press.

Fanselow, M. S. and Rudy, J. W. (1998) Convergence of experimental and developmental approaches to animal learning and memory processes. In *Mechanistic Relationships between Development and Learning,* ed. by T. J. Carew, R. Menzel, and C. J. Shatz, pp. 15–28. New York: Wiley.

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

Fernald, A. (1985) Four-month-old infants prefer to listen to motherese. *Inf. Behav. Develop.* 8:181–195.

Fernald, A. and Kuhl, P. (1987) Acoustic determinants of infant preference for Motherese speech. *Infant Behavior and Development* 10:279–293.

Flege, J. E. (1988) Factors affecting degree of perceived foreign accent in English sentences. *J. Acoust. Soc. Am.* 84:70–79.

Flege, J. E. (1993) Production and perception of a novel, second-language phonetic contrast. *J. Acoust. Soc. Am.* 93:1589–1608.

Glanville, B. B., Best, C. T., and Levenson, R. (1977) A cardiac measure of cerebral asymmetries in infant auditory perception. *Develop. Psych.* 13:54–59.

Gopnik, A. and Meltzoff, A. N. (1997) *Words, Thoughts, and Theories.* Cambridge, MA: MIT Press.

Goto, H. (1971) Auditory perception by normal Japanese adults of the sounds "l" and "r." *Neuropsychologia* 9:317–323.

Green, K. P., Kuhl, P. K., Meltzoff, A. N., and Stevens, E. B. (1991) Integrating speech information across talkers, gender, and sensory modality: Female faces and male voices in the McGurk effect. *Perception & Psychophysics* 50:524–536.

Greenough, W. T. and Black, J. E. (1992) Induction of brain structure by experience: Substrates for cognitive development. In *The Minnesota Symposia on Child Psychology*, Vol. 24: *Developmental Behavioral Neuroscience*, ed. by M. Gunnar and C. Nelson, pp. 155–200. Hillsdale, NJ: Erlbaum.

Grieser, D. and Kuhl, P. K. (1989) Categorization of speech by infants: Support for speech-sound prototypes. *Develop. Psych.* 25:577–588.

Hauser, M. D. (1996) *The Evolution of Communication*. Cambridge, MA: MIT Press.

Iverson, P. and Kuhl, P. K. (1995) Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *J. of the Acoust. Soc. Am.* 97:553–562.

Iverson, P. and Kuhl, P. K. (1996) Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. *J. of the Acoust. Soc. of Am.* 99:1130–1140.

James, W. (1890 [1958]) *Principles of Psychology*. New York: Holt.

Johnson, K., Flemming, E., and Wright, R. (1993) The hyperspace effect: Phonetic targets are hyperarticulated. *Language* 69:505–528.

Jusczyk, P. W., Friederici, A. D., Wessels, J. M. I., Svenkerud, V. Y., and Jusczyk, A. M. (1993) Infants' sensitivity to the sound patterns of native language words. *J. of Mem. and Lang.* 32:402–420.

Karmiloff-Smith, A. (1991) Beyond modularity: Innate constraints and developmental change. In *The Epigenesis of Mind: Essays on Biology and Cognition*, ed. by S. Carey, and R. Gelman, pp. 171–197. Hillsdale, NJ: Erlbaum.

Karmiloff-Smith, A. (1995) Annotation: The extraordinary cognitive journey from foetus through infancy. *J. Child Psych. & Psychiat. & All. Disciplines* 36:1293–1313.

Kent, R. D. and Murray, A. D. (1982) Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *J. of the Acoust. Soc. of Am.* 72:353–365.

Kim, K. H. S., Relkin, N. R., Lee, K. M., and Hirsch, J. (1997) Distinct cortical areas associated with native and second languages. *Nature* 388:172–174.

Kimura, D. (1961) Cerebral dominance and the perception of verbal stimuli. *Can J. Psych.* 15:166–171.

Kimura, D. (1963) Speech lateralization in young children as determined by an auditory test. *J. of Comp. and Physiol. Psych.* 56:899–901.

Klima, E. S., Bellugi, U., and Poizner, H. (1988) Grammar and space in sign aphasiology. *Aphasiol.* 2:319–327.

Kluender, K. R., Diehl, R. L., and Killeen, P. R. (1987) Japanese quail can learn phonetic categories. *Science* 237:1195–1197.

Knudsen, E. I. and Brainerd, M. S. (1995) Creating a unified representation of visual and auditory space in the brain. *Ann. Rev. Neurosci.* 18:19–43.

Knudsen, E. I. and Knudsen, P. F. (1990) Sensitive and critical periods for visual calibration of sound localization by barn owls. *J. Neurosci.* 10:222–232.

Koffka, K. (1924) *The Growth of the Mind*. New York: Harcourt Brace.

Konishi, M. (1965) Effects of deafening on song development in American robins and black-headed grosbeaks. *Z. Tierpsychol.* 22:584–599.

Kraus, N., McGee, T., Sharma, A., Carrell, T., and Nicol, T. (1992) Mismatch negativity event-related potential elicited by speech stimuli. *Ear and Hearing* 13:158–164.

Kraus, N., McGee, T., Micco, A., Sharma, A., Carrell, T., and Nicol, T. (1993) Mismatch negativity in school-age children to speech stimuli that are just perceptibly different. *Electroenceph. Clin. Neurophys.* 88:123–130.

Kuhl, P. K. (1991a) Perception, cognition, and the ontogenetic and phylogenetic emergence of human speech. In *Plasticity of Development*, ed. by S. E. Brauth, W. S. Hall, and R. J. Dooling, pp. 73–106. Cambridge, MA: MIT Press.

Kuhl, P. K. (1991b) Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics* 50:93–107.

Kuhl, P. K. (1992) Infants' perception and representation of speech: Development of a new theory. In *Proceedings of the International Conference on Spoken Language Processing*, ed. by J. J. Ohala, T. M. Nearey, B. L. Derwing, M. M. Hodge, and G. E. Wiebe, pp. 449–456. Edmonton, Alberta: University of Alberta.

Kuhl, P. K. (1994) Learning and representation in speech and language. *Current Opinion in Neurobiology* 4:812–822.

Kuhl, P. K. (1998a) The development of language. In C. von Euler, I. Lundberg, and R. Llinás (eds.), *Basic mechanisms in cognition and language*, pp. 175–195. New York: Elsevier.

Kuhl, P. K. (1998b) Effects of language experience on speech perception. Paper presented at the 135th meeting of the Acoustical Society of America, Seattle, WA.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., and Lacerda, F. (1997) Cross-language analysis of phonetic units in language addressed to infants. *Science* 277:684–686.

Kuhl, P. K., Kiritani, S., Deguchi, T., Hayashi, A., Stevens, E. B., Dugger, C. D., and Iverson, P. (1997) Effects of language experience on speech perception: American and Japanese infants' perception of /ra/ and /la/. *J. Acoust. Soc. Am.* 102:3135.

Kuhl, P. K. and Meltzoff, A. N. (1982) The bimodal perception of speech in infancy. *Science* 218:1138–1141.

Kuhl, P. K. and Meltzoff, A. N. (1996) Infant vocalizations in response to speech: Vocal imitation and developmental change. *J. Acoust. Soc. Am.* 100:2425–2438.

Kuhl, P. K., Tsuzaki, M., Tohkura, Y., and Meltzoff, A. N. (1994) Human processing of auditory-visual information in speech perception: Potential for multimodal human-machine interfaces. In *Proceedings of the International Conference on Spoken Language Processing*, pp. 539–542. Tokyo: Acoustical Society of Japan.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (1992) Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255:606–608.

Lenneberg, E. H. (1967) *Biological Foundations of Language.* New York: Wiley.

Liberman, A. M. (1993) In the perception of speech, time is not what it seems. In *Temporal Information Processing in the Nervous System.* Annals of the New York Academy of Sciences, ed. by P. Tallal, A. M. Galaburda, R. R. Llinás, and C. von Euler, vol. 682, pp. 264–271. New York: The New York Academy of Sciences.

Lindblom, B. (1990) Explaining phonetic variation: A sketch of the H&H theory. In *Speech Production and Speech Modeling*, ed. by W. J. Hardcastle, and A. Marchal, pp. 403–439. Dordrecht, Netherlands: Kluwer Academic Publishers.

Lively, S. E., Logan, J. S., Pisoni, D. B. (1993) Training Japanese listeners to identify English /r/ and /l/. II. The role of phonetic environment and talker variability in learning new perceptual categories. *J. Acoust. Soc. Am.* 94:1242–1255.

Locke, J. (1690) *An Essay Concerning Human Understanding.* London: Bassett.

Lorenz, K. (1965) *Evolution and Modification of Behavior.* Chicago: Chicago University Press.

MacKain, K., Studdert-Kennedy, M., Spieker, S., and Stern, D. (1983) Infant intermodal speech perception is a left-hemisphere function. *Science* 219:1347–1349.

Marler, P. (1970). A comparative approach to vocal learning: Song development in white-crowned sparrows. *J. Comp. and Physiol. Psych.* 71:1–25.

Marler, P. (1974) Constraints on learning: Development of bird song. In *Ethology and Psychiatry*, ed. by W. F. Norman, pp. 69–83. Toronto: University of Toronto Press.

Marler, P. (1990) Innate learning preferences: Signals for communication. *Dev. Psychobiol.* 23:557–568.

Marler, P. (1991) The instinct to learn. In *The Epigenesis of Mind: Essays on Biology and Cognition*, ed. by S. Carey, and R. Gelman, pp. 37–66. Hillsdale, NJ: Erlbaum.

Marler, P. (1977) Three models of song learning: Evidence from behavior. *J. of Neurobiol.* 33:501–506.

Massaro, D. W. (1987) *Speech Perception by Ear and Eye: A Paradigm for Psychological Inquiry*. Hillsdale, NJ: Erlbaum.

McGurk, H. and MacDonald, J. (1976) Hearing lips and seeing voices. *Nature* 264:746–748.

Meltzoff, A. N. and Moore, M. K. (1977) Imitation of facial and manual gestures by human neonates. *Science* 198:75–78.

Meltzoff, A. N. and Moore, M. K. (1994) Imitation, memory, and the representation of persons. *Infant Behav. and Dev.* 17:83–99.

Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., and Tallal, P. (1996) Temporal processing deficits of language-learning impaired children ameliorated by training. *Science* 271:77–81.

Miller, J. L. (1994) On the internal structure of phonetic categories: A progress report. *Cognition* 50:271–285.

Milner, B., Taylor, L., and Sperry, R. W. (1968) Lateralized suppression of dichotically presented digits after commissural section in man. *Science* 161:184–186.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., and 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.

Näätänen, R. (1990) The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behav. Brain Sci.* 13:201–288.

Näätänen, R. (1992) The mismatch negativity (MMN). In *Attention and Brain Function*, pp. 137–210. New Jersey: Erlbaum Assoc.

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

Neville, H. J., Bavelier, D., Corina, D., Rauschecker, J., Karni, A., Lalwani, A., et al. (1998) Cerebral organization for language in deaf and hearing subjects: Biological constraints and effects of experience. *Proceedings of the Nat. Acad. Sci.* 95:922–929.

Neville, H. J., Mills, D. L., and Lawson, D. S. (1992) Fractionating language: Different neural subsystems with different sensitive periods. *Cerebral Cortex* 2:244–258.

Nottebohm, F. (1967) The role of sensory feedback in the development of avian vocalizations. *Proceedings of the 14th International Ornithological Congress*, pp. 265–280.

Oller, D. K. and MacNeilage, P. F. (1983) Development of speech production: Perspectives from natural and perturbed speech. In *The Production of Speech*, ed. by P. F. MacNeilage, pp. 91–108. New York: Springer-Verlag.

Oller, D. K., Wieman, L. A., Doyle, W. J., and Ross, C. (1976) Infant babbling and speech. *J. Child Lang.* 3:1–11.

Osterhout, L. and Holcomb, P. L. (1992) Event-related brain potentials elicited by syntactic anomaly. *J. Mem. Lang.* 31:785–806.

Pang, E. W., Edmonds, G. E., Desjardins, R., Khan, S. C., Trainor, L. J., and Taylor, M. J. (1998) Mismatch negativity to speech stimuli in 8-month-old infants and adults. *Int. J. Psychophysiol.* 29:227–236.

Petersen, S. E., Fox, P. T., Snyder, A. Z., and Raichle, M. E. (1990) Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. *Science* 249:1041–1044.

Petitto, L. A. and Marentette, P. F. (1991) Babbling in the manual mode: Evidence for the ontogeny of language. *Science* 251:1493–1496.

Petrinovich, L. and Baptista, L. F. (1987) Song development in the white-crowned sparrow: Modification of learned song. *Anim. Behav.* 35:961–974.

Piaget, J. (1954) *The Construction of Reality in the Child.* New York: Basic Books.

Pinker, S. (1994) *The Language Instinct.* New York: Morrow.

Poeppel, D., Phillips, C., Yellin, E., Rowley, H. A., Roberts, T. P., and Marantz, A. (1997) Processing of vowels in supratemporal auditory cortex. *Neurosci. Letters* 221:145–148.

Posner, M. I., Petersen, S. E., Fox, P. T., and Raichle, M. E. (1988) Localization of cognitive operations in the human brain. *Science* 240:1627–1631.

Rosch, E. H. (1977) Human categorization. In *Studies in Cross-cultural Psychology,* vol. 1, ed by N. Warren, pp. 1–49. San Francisco: Academic Press.

Rosenblum, L. D., Schmuckler, M. A., and Johnson, J. A. (1997) The McGurk effect in infants. *Percept. Psychophys.* 59:347–357.

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

Skinner, B. F. (1957) *Verbal Behavior.* New York: Appleton-Century-Crofts.

Snowdon, C. T. and Hausberger, M. (eds.) (1997) *Social Influences on Vocal Development.* Cambridge: Cambridge University Press.

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

Strange, W. (1995) *Speech Perception and Linguistic Experience: Issues in Cross-language Research.* Timonium, MD: York.

Studdert-Kennedy, M. and Shankweiler, D. (1970) Hemispheric specialization for speech perception. *J. Acoust. Soc. Am.* 48:579–594.

Sussman, J. E. and Lauckner-Morano, V. J. (1995) Further tests of the "perceptual magnet effect" in the perception of [i]: Identification and change/no-change discrimination. *J. Acoust. Soc. Am.* 97:539–552.

Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, G., Nagarajan, S. S., Schreiner, C., Jenkins, W. M., and Merzenich, M. M. (1996) Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science* 271:81–84.

Thorpe, W. H. (1959) Talking birds and the mode of action of the vocal apparatus of birds. *Proceedings Zoo. Soc. London* 132:441–455.

Thorpe, W. H. (1961) *Bird Song: The Biology of Vocal Communication and Expression in Birds.* New York: Cambridge University Press.

Tinbergen, N. (1951) *The Study of Instinct.* New York: Clarendon.

Vargha-Khadem, F. and Corballis, M. C. (1979) Cerebral asymmetry in infants. *Brain & Lang.* 8:1–9.

Vihman, M. M. (1993) Vocal motor schemes, variation and the production-perception link. *J. Phonetics* 21:163–169.

Vihman, M. M. and de Boysson-Bardies, B. (1994) The nature and origins of ambient language influence on infant vocal production and early words. *Phonetica* 51:159–169.

Vygotsky, L. S. (1962) *Thought and Language.* Cambridge, MA: MIT Press.

Vygotsky, L. S. (1979) Interaction between learning and development. In *Mind in Society: The Development of Higher Psychological Processes,* ed. by M. Cole, V. John-Steiner, S. Scribner, and E. Souberman, pp. 79–91. Cambridge, MA: Harvard University Press.

Walton, G. E. and Bower, T. G. R. (1993) Amodal representation of speech in infants. *Inf. Behav. Develop.* 16:233–243.

Werker, J. F. and Tees, R. C. (1984) Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Inf. Behav. Develop.* 7:49–63.

Werker, J. F. and Polka, L. (1993) The ontogeny and developmental significance of language-specific phonetic perception. In *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, ed. by B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, and J. Morton, pp. 275–288. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Wernicke, C. (1874) *Der aphasische symptomenkomplex.* Breslau: Cohn and Weigert.

Willerman, R. and Kuhl, P. K. (1996) Cross-language speech perception: Swedish, English, and Spanish speakers' perception of front rounded vowels. *Proc. 1996 Int. Conf. Spoken Lang. Proc.* 1:442–445.

Witelson, S. F. (1987) Neurobiological aspects of language in children. *Child Develop.* 58:653–688.

Yamada, R. A. and Tohkura, Y. (1992) The effects of experimental variables on the perception of American English /r/ and /l/ by Japanese listeners. *Percept. Psychophys.* 52:376–392.

Zatorre, R. J., Evans, A. C., Meyer, E., and Gjedde, A. (1992) Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256:846–849.