Colloquium

A new view of language acquisition

Patricia K. Kuhl*

Department of Speech and Hearing Sciences and Center for Mind, Brain, and Learning, University of Washington, Box 357920, Seattle, WA 98195

At the forefront of debates on language are new data demonstrating infants' early acquisition of information about their native language. The data show that infants perceptually ''map'' critical aspects of ambient language in the first year of life before they can speak. Statistical properties of speech are picked up through exposure to ambient language. Moreover, linguistic experience alters infants' perception of speech, warping perception in the service of language. Infants' strategies are unexpected and unpredicted by historical views. A new theoretical position has emerged, and six postulates of this position are described.

The last half of the 20th century has produced a revolution in our understanding of language and its acquisition. Studies of infants across languages and cultures have provided valuable information about the initial state of the mechanisms underlying language, and more recently, have revealed infants' unexpected learning strategies. The learning strategies—demonstrating pattern perception, as well as statistical (probabilistic and distributional) computational skills—are not predicted by historical theories. The results lead to a new view of language acquisition, one that accounts for both the initial state of linguistic knowledge in infants and infants' extraordinary ability to learn simply by listening to ambient language. The new view reinterprets the critical period for language and helps explain certain paradoxes—why infants, for example, with their immature cognitive systems, far surpass adults in acquiring a new language. The goal of this paper is to illustrate the recent work and offer six principles that shape the new perspective.

Historical Theoretical Positions

In the last half of the 20th century, debate on the origins of language was ignited by a highly publicized exchange between a strong nativist and a strong learning theorist. In 1957, the behavioral psychologist B F. Skinner proposed a learning view in his book *Verbal Behavior*, arguing that language, like all animal behavior, was an ''operant'' that developed in children as a function of external reinforcement and shaping (1). By Skinner's account, infants learn language as a rat learns to press a bar—through the monitoring and management of reward contingencies.

Noam Chomsky, in a review of *Verbal Behavior*, took a very different theoretical position (2, 3). Chomsky argued that traditional reinforcement learning had little to do with humans' abilities to acquire language. He posited a ''language faculty'' that included innately specified constraints on the possible forms human language could take. Chomsky argued that infants' innate constraints for language included specification of a universal grammar and universal phonetics. Language was one of the primary examples of what Fodor called a module—domain-specific, informationally encapsulated, and innate (4).

The two approaches took strikingly different positions on all of the critical components of a theory of language acquisition: (*i*) the initial state of knowledge, (*ii*) the mechanisms responsible for developmental change, and (*iii*) the role played by ambient language input. On Skinner's view, no innate information was necessary, developmental change was brought about through
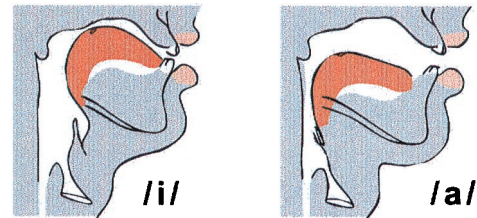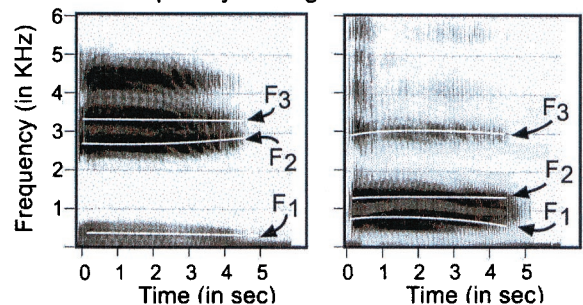

Fig. 1. Vocal tract positions (*Upper*) and spectrographic displays (*Lower*) for the vowels /i/ as in ''heat'' and /a/ as in ''hot.'' Formant frequencies, regions of the frequency spectrum in which the concentration of energy is high, are marked for each vowel.

reward contingencies, and language input did not cause language to emerge. On Chomsky's view, infants' innate knowledge of language was a core tenet, development constituted "growth" or maturation of the language module, and language input triggered (or set the parameters for) a particular pattern from among those innately provided.

A great deal has been learned since the debate ensued, caused largely by experiments conducted on infants. Infants' perception of the phonetic units of speech, which requires tracking the formant frequencies (Fig. 1) (5), and their detection of words from cues in running speech (Fig. 2) (6) support a different view. The emerging view argues that the kind of learning taking place in early language acquisition cannot be accounted for by Skinnerian reinforcement. On the other hand, the idea that language acquisition involves a selectionist process wherein language input operates on innately specified options also is not supported. The emerging view suggests that infants engage in a new kind of learning in which language input is mapped in detail by the infant brain. Six principles reflecting this view are offered.

Only a millio-naire can have a house with round-windowed sun- lighted rooms

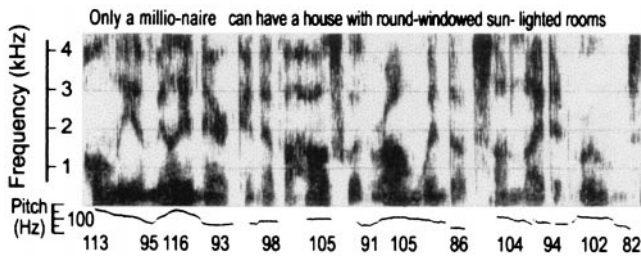113 95 116 93   98 105 91 105   86 104 94 102 82

Fig. 2. Spectrographic display of running speech showing the formant frequencies and the pitch (fundamental frequency) of the voice over time. Increases in pitch indicate primary stress in the utterance. [Reproduced with permission from ref. 6 (Copyright 1999, Allyn & Bacon).]

Fig. 3. Humans' and animals' identification functions for two series ranging from a voiced to a voiceless phonetic unit (/d/to/t/and/g/to/k/) showing that animals' perceptual boundaries correspond to humans' phonetic boundaries. VOT, voice onset time. [Reproduced with permission from ref. 19 (Copyright 1978, Acoustical Society of America).]

Initial Perception Parses Speech Correctly and Is Universal, but Not Domain Specific or Species Specific

Any theory of language acquisition has to specify how infants parse the auditory world to make the critical units of language available. This is a formidable problem as indicated by the difficulty computers have in segmenting speech (7–9). Early experiments on infants confirmed their abilities to parse speech correctly at the phonetic level and revealed that their abilities are universal across languages. Interestingly, however, the data also demonstrated that the kind of partitioning seen for speech is not limited to humans or limited to speech.

The evidence derived from tests of categorical perception (10). When adult listeners were tested on a continuum that ranges from one syllable (such as "bat") to another ("pat"), perception appeared absolute. Adults discriminated phonetic units that crossed the "phonetic boundary" between categories but not stimuli that fell within a category. The phenomenon was language-specific; Japanese adults, for example, failed to show a peak in discrimination at the phonetic boundary of an American English/ra-la/series (as in "rake" vs. "lake") (11).

Categorical perception provided an opportunity to test whether infants could parse the basic units of language, and discrimination tests confirmed that they did. Infants discriminated only between stimuli from different phonetic categories (12–14). Moreover, unlike adults, infants demonstrated the effect for the phonetic units of all languages (15, 16). Eimas hypothesized that infants' abilities reflected innate "phonetic feature detectors" that evolved for speech and theorized that infants are biologically endowed with neural mechanisms that respond to the phonetic contrasts used by the world's languages (17).

Experimental tests on nonhuman animals altered this conclusion (18, 19). Animals also exhibited categorical perception; they demonstrated perceptual "boundaries" at locations where humans perceive a shift from one phonetic category to another (18, 19) (Fig. 3). In tests of discrimination, monkeys showed peaks in sensitivity that coincided with the phonetic boundaries used by languages (20–22) (Fig. 4). The results were subsequently replicated in a number of species (23, 24). Recently, additional tests on infants and monkeys revealed similarities in their perception of the prosodic cues of speech as well (25).

Two conclusions were drawn from the initial comparative work (26). First, infants' parsing of the phonetic units at birth was a discriminative capacity that could be accounted for by a general auditory processing mechanism, rather than one that evolved specifically for speech. Differentiating the units of speech did not imply *a priori* knowledge of the phonetic units themselves, merely the capacity to detect differences between them, which was constrained in an interesting way (18, 19, 25, 27). Second, in the evolution of language, acoustic differences detected by the auditory perceptual processing mechanism strongly influenced
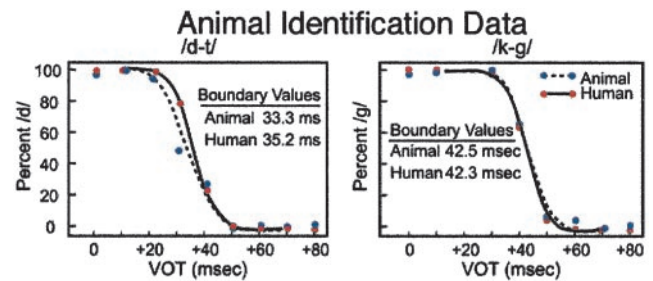
the selection of phonetic units used in language. On this view, particular auditory features were exploited in the evolution of the sound system used in language (19, 26, 27). This ran counter to two prevailing principles at the time: (*i*) the view that phonetic units were prespecified in infants, and (*ii*) the view that language evolved in humans without continuity with lower species.

Categorical perception also was demonstrated with nonspeech stimuli that mimicked speech features without being perceived as speech, in both adults (28, 29) and infants (30). This finding supported the view that domain-general mechanisms were responsible for infants' initial partitioning of the phonetic units of language.

Development Is Not Based on Selection

Eimas' early model of speech perception was selectionist in nature. An innate neural specification of all possible phonetic units allowed selection of a subset of those units to be triggered by language input (17). The notion was that linguistic experience produced either maintenance or loss. Detectors stimulated by ambient language were maintained, whereas those not stimulated by language input atrophied.

Developmental studies were initially seen as providing support for the selectionist view. Werker and her colleagues demonstrated that, by 12 months of age, infants no longer discriminate non-native phonetic contrasts, even though they did so at 6 months of age (31). The finding was interpreted as support for a selectionist theory; there was a "loss" of a subset of phonetic units initially specified.

Modifications regarding the extent to which listeners "lost" the ability to discriminate non-native phonetic units were quick to follow (32). Adult performance on non-native contrasts could be increased by a number of factors: (*i*) the use of techniques that minimize the effects of memory (33, 34), (*ii*) extensive training
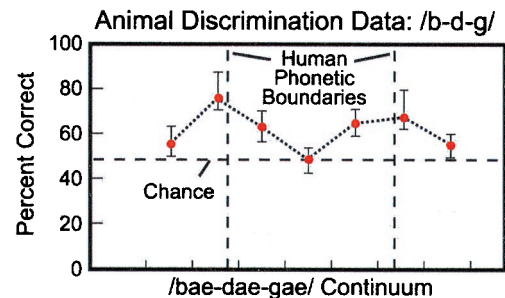

Fig. 4. Monkey discrimination performance for pairs of stimuli drawn from a continuum of speech sounds ranging from/b/to/d/to/g/, showing that sensitivity increases near the locations of humans' phonetic boundaries. [Reproduced with permission from ref. 22 (Copyright 1978, Acoustical Society of America.]

(35, 36), and (*iii*) the use of contrasts, such as Zulu clicks, that are not related to native-language categories (37, 38). These data, indicating that there is not an immutable loss of phonetic abilities for non-native units (32), did not refute the selectionist position. The fact that listeners do not completely lose the ability to discriminate non-native contrasts does not alter the basic tenet of the selectionist view, which is that the role of language experience is to maintain or decrease the activity of innately specified neural detectors. To refute the selectionist position, studies must demonstrate that infants listening to ambient language are engaged in some other kind of learning process, a process that is not fundamentally subtractive in nature. New studies on learning provide that demonstration.

Infants' Learning Strategies "Map" Language Input

Learning theory as a mechanism for language acquisition had been dismissed by early theorists because of the failure of existing learning models, such as Skinner's, to explain the facts of language development (2). At present, however, learning models figure prominently in debates on language (39–42). What has changed? The discoveries of the last two decades, demonstrating that by simply listening to language infants acquire sophisticated information about its properties, have created new views of learning.

Three important examples of a new kind of learning have emerged. First, infants detect patterns in language input. Second, infants exploit the statistical properties of the input, enabling them to detect and use distributional and probabilistic information contained in ambient language to identify higher-order units. Third, infant perception is altered—literally warped—by experience to enhance language perception. No speaker of any language perceives acoustic reality; in each case, perception is altered in the service of language.

Infants Abstract Patterns. A major requirement of language processing is the detection of similarities, or patterns, in language input, a stumbling block for computer speech recognition (7). Infants demonstrate excellent skills at pattern recognition for speech. A number of studies have shown that 6-month-old infants, trained to produce a head-turn response when a sound from one category is presented (such as the vowel /a/ in "pop"), and to inhibit that response when an instance from another vowel category is presented (/i/ in "peep"), demonstrate the ability to perceptually sort novel instances into categories (43).

For example, infants perceptually sort vowels that vary across talkers and intonation contours (44, 45), as well as syllables that vary in their initial consonant (those beginning with /m/ as opposed to /n/, or those beginning with /s/ versus /ʃ/) across variations in talkers and vowel contexts (46, 47). Moreover, infants perceptually sort syllables based on a phonetic feature shared by their initial consonants, such as a set of nasal consonants, /m/, /n/, and /ŋ/, as opposed to a set of stop consonants, /b/, /d/, and /g/ (46). Recent tests show that 9-month-old infants are particularly attentive to the initial portions of syllables (48).

Infants' detection of patterns is not limited to phonetic units. More global prosodic patterns contained in language also are detected. At birth, infants have been shown to prefer the language spoken by their mothers during pregnancy, as opposed to another language (49–51). This skill requires infant learning of the stress and intonation pattern characteristic of the language (the pitch information shown in Fig. 2), information that is reliably transmitted through bone conduction to the womb (52). Additional evidence that the learning of speech patterns commences *in utero* stems from studies showing infant preference for their mother's voice over another female at birth (53) and their preference for stories read by the mother during the last 10 weeks of pregnancy (54).

Between 6 and 9 months, infants exploit prosodic patterns related to the stress or emphasis typical of words in their native language. In English, a strong/weak pattern of stress, with emphasis on the initial syllable ("*ba*by," "*mom*my," "*ta*ble") is typical, whereas a weak/strong pattern predominates in other languages. American infants tested at 6 months show no listening preference for words with the strong/weak as opposed to the weak/strong pattern, but by 9 months they exhibit a strong preference for the pattern typical of their native language (55). Infants also use prosodic cues to detect major constituent boundaries, such as clauses. At 4 months of age, infants listen equally long to Polish and English speech samples that have pauses inserted at clause boundaries as opposed to within clauses, but by 6 months, infants listen preferentially to pauses inserted at the clause boundaries appropriate only to their native language (41, 56).

By 9 months of age, infants detect patterns related to the orderings of phonemes that are legal for their language. In English, for example, the combination *zw or vl* is not legal; in Dutch, they are permissible. By 9 months of age, but not at 6 months of age, American infants listen longer to English words, whereas Dutch infants show a listening preference for Dutch words (57). At this age, infants do not recognize the words themselves, but recognize the perceptual patterns typical of words in their language. They develop a "perceptual sleeve" in which words fit; a description of word candidates assists them in identifying potential words in running speech.

Infants Exploit Statistical Properties of Language Input. Running speech presents a problem for infants because, unlike written speech, there are no breaks between words. New research shows that infants detect and exploit the statistical properties of the language they hear to find word candidates in running speech before they know the meanings of words. Goodsitt, Morgan, and Kuhl (58) demonstrated this in 7-month-old infants by using artificial words.

Goodsitt *et al.* examined infants' abilities to maintain the discrimination of two isolated syllables, /de/ and /ti/, when these target syllables were later embedded in three-syllable strings. The three-syllable strings contained the target syllable and a bisyllable composed of the syllables /ko/ and/ga/. The arrangement of /ko/ and /ga/ was manipulated to change the degree to which they could be perceived as a likely word candidate. Three conditions were tested. In *a*, /koga/ was an invariantly ordered "word," appearing either after the target syllables, /dekoga/ and /tikoga/, or before it, /kogade/ and /kogati/. In this condition, the transitional probability between the /ko/ and /ga/ was always 1.0. If infants detect /koga/ as a unit, it should assist infants in detecting and discriminating /de/ from /ti/. In *b*, the two syllables could either appear in variable order, either /koga/ or /gako/, reducing the transitional probabilities to 0.3 and preventing infants from perceiving /koga/ as a word. In *c*, one of the context syllables was repeated (e.g., /koko/). In this case, /koko/ could be perceived as a unit, but the basis of the perception would not be high transitional probabilities; the transitional probabilities between syllables in *c* remain low (0.3).

The results confirmed the hypothesis that 7-month-old infants exploit transitional probabilities. Infants discriminated the target syllables in condition *a* significantly more accurately than in either *b* or *c*, the latter of which showed equally poor discrimination. These strategies also have been shown to be effective for adults presented with artificial nonspeech analogs created by computer (42, 59).

In further work, Saffran, Aslin, and Newport (42) directly assessed 8-month-old infants' abilities to learn pseudowords based on transitional probabilities. Infants were exposed to 2-min strings of synthetic speech composed of four different

pseudowords that followed one another equally often. There were no breaks, pauses, stress differences, or intonation contours to aid infants in recovering these "words" from the strings of syllables. During the test phase, infants listened to two of the original pseudowords and two new words formed by combining parts of two of the original words. The results demonstrated that infants' listened longer to the new words, demonstrating that they are capable of using statistical regularities to detect words (60).

Additional examples of the computation and use of probability statistics have been uncovered. Nine-month-old infants detect the probability of occurrence of legal sequences that occur in English (61). Certain combinations of two consonants are more likely to occur within words whereas others occur at the juncture between words. The combination "ft" is more common within words whereas the combination "vt" is more common between words. Nine-month-olds were tested with consonant (C) and vowel (V) strings of the form CVCCVC. These items contained embedded CCs that were either frequent or infrequent in English. Infants listened significantly longer to the lists containing frequent within-word CCs.

The results reveal that an old principle of Gestalt psychology, referred to as "common fate" (58), plays a role in speech perception. Phonemes that are typically linked, and thus share a common fate, are perceived as units by infants. It is interesting to note that early object perception also may rely on this principle. Physical entities whose properties cohere in space, and move together, are perceived as individuated objects (62). Whether the constraints underlying infants' detection of common fate information for physical objects and speech are identical or different is important to theory and remains to be examined.

Language Experience Warps Perception. Language experience not only produces a change in infants' discriminative abilities and listening preferences, it results in a "mapping" that alters perception. A research finding that helps explain this is called the perceptual magnet effect. The magnet effect is observed when tokens perceived as exceptionally good representatives of a phonetic category ("prototypes") are used in tests of speech perception (63–66). Many behavioral (63–69) and brain (70–73) studies indicate that native-language phonetic prototypes evoke special responses when compared with nonprototypes.

When tested with a phonetic prototype as opposed to a nonprototype from the same category, infants show greater ability to generalize to other category members (63, 64). The prototype appears to function as a "magnet" for other stimuli in the category, in a way similar to that shown for prototypes of other cognitive categories (74, 75). Moreover, the perceptual magnet effect depends on exposure to a specific language (65). Six-month-old infants being raised in the United States and Sweden were tested with two vowel prototypes, an American English /i/ vowel prototype and a Swedish /y/ vowel prototype, using the exact same stimuli (Fig. 5A), techniques, and testers in the two countries. American infants demonstrated the magnet effect only for the American English /i/, treating 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 (Fig. 5B). The results show that by 6 months of age, perception is altered by language experience.

Categorical perception and the perceptual magnet effect make different predictions about the perception and organization underlying speech categories and appear to arise from different mechanisms (76). Interestingly, comparative tests show that, unlike categorical perception, animals do not exhibit the perceptual magnet effect (64).

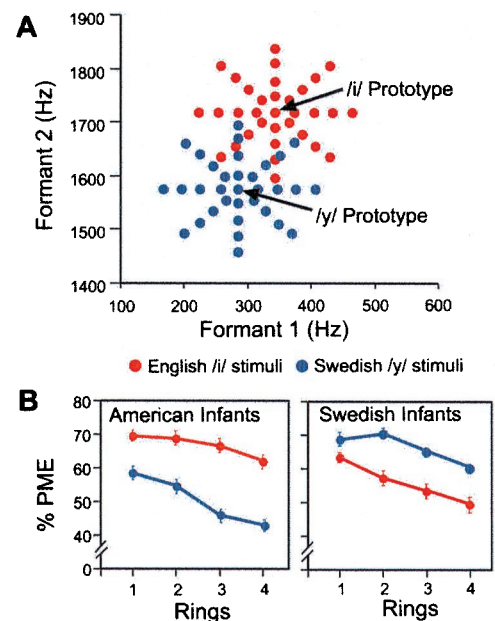In adults, the distortion of perception caused by language

Fig. 5. (A) Formant frequencies of vowels surrounding an American/i/ prototype (red) and a Swedish/y/prototype (blue). (B) Results of tests on American and Swedish infants indicating an effect of linguistic experience. Infants showed greater generalization when tested with the native-language prototype. PME, Perceptual magnet effect. [Reproduced with permission from ref. 65 (Copyright 1992, American Association for the Advancement of Science).]

experience is well illustrated by a study on the perception of American English /r/ and /l/ in American and Japanese listeners. The /r-l/ distinction is difficult for Japanese speakers to perceive and produce; it is not used in the Japanese language (77, 78). In the study, Iverson and Kuhl (79) used computer-synthesized syllables beginning with /r/ and /l/, spacing them at equal physical intervals in a two-dimensional acoustic grid (Fig. 6A). American listeners identified each syllable as /ra/ or /la/, rated its category goodness, and estimated the perceived similarity for all possible pairs of syllables. Similarity ratings were scaled by using multidimensional scaling techniques. The results provide a map of the perceived distances between stimuli—short distances for strong similarity and long distances for weak similarity. In the American map (Fig. 6B), magnet effects (seen as a shrinking of perceptual space) occur in the region of each category's best instances. Boundary effects (seen as a stretching of perceptual space) occur at the division between the two categories.

The experiment has recently been completed with Japanese monolingual listeners, and the results show a striking contrast in the way the /r-l/ stimuli are perceived by American and Japanese speakers. The map revealed by multidimensional scaling analysis is totally different—no magnet effects or boundary effects appear. Japanese listeners hear one category of sounds, not two, and attend to different dimensions of the same stimuli. The results suggest that linguistic experience produces mental maps for speech that differ substantially for speakers of different languages (40, 69, 79).

The important point regarding development is that the initial perceptual biases shown by infants in tests of categorical perception (12–16), as well as asymmetries in perception seen in infancy (80, 81), produce a contouring of the perceptual space that is universal. This universal contouring soon gives way to a language-specific mapping that distorts perception, completely revising the perceptual space underlying speech processing (65).
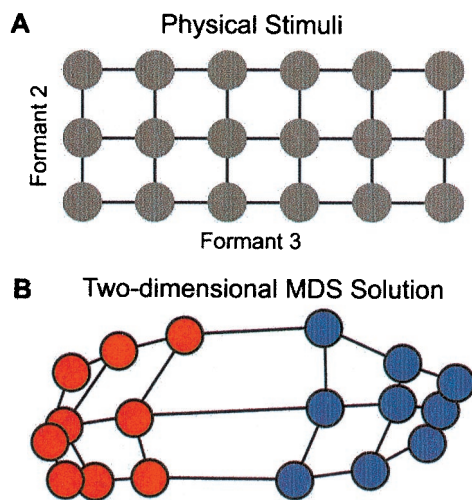
A Physical Stimuli

Formant 2 / Formant 3

B Two-dimensional MDS Solution

Fig. 6. (A) Physical distance between /ra-la/ syllables in a grid created by varying formants 2 and 3 in equal steps. (B) Perceptual distance between syllables for American listeners showing a warping of acoustic space. MDS, Multidimensional scaling. [Reproduced with permission from ref. 79 (Copyright 1996, Acoustical Society of America).]

A model reflecting this developmental sequence from universal perception to language-specific perception, called the Native Language Magnet model, proposes that infants' mapping of ambient language warps the acoustic dimensions underlying speech, producing a complex network, or filter, through which language is perceived (39, 40, 82). The language-specific filter alters the dimensions of speech we attend to, stretching and shrinking acoustic space to highlight the differences between language categories. Once formed, language-specific filters make learning a second language much more difficult because the mapping appropriate for one's primary language is completely different from that required by other languages. Studies of adult bilinguals, who were exposed to their second language after the age of 6, demonstrate magnet effects only for the first language, illustrating the potent effects of early linguistic experience (66). According to the Native Language Magnet theory, infants' transition in speech perception between 6 and 12 months reflects the formation of a language-specific filter.

In summary, the studies on speech learning, demonstrating that infants detect patterns, extract statistical information, and have perceptual systems that can be altered by experience, cannot be explained by recourse to Skinnerian reinforcement learning. This is a different kind of learning, one ubiquitous during early development. Its study will be valuable beyond what it tells us about language learning.

Are the new learning strategies observed for speech domain-specific and/or species-specific? Research on cognitive development confirms the fact that categorization (83), statistical learning (84), and prototype effects (85) are not unique to speech. Further tests need to be done to determine the constraints operating on these abilities in infants by using linguistic and nonlinguistic events. What about animal tests? Thus far, the data suggest differences between animals and humans on these kinds of learning. For instance, monkeys do not exhibit the perceptual magnet effect (64). Animals do show some degree of internal structure for speech categories after extensive training (24), but it is unlikely the perceptual magnet effect would be spontaneously produced in an animal after 6 months' experience listening to language, as seen in human infants. Similarly, animals are sensitive to transitional probabilities (86), but it is unlikely that an animal would spontaneously exhibit statistical learning after

simply listening to language, as human infants have been shown to do. These issues can be resolved with empirical tests.

Vocal Learning Unifies Perception and Production

Infants not only learn the perceptual characteristics of their language, they become native speakers, which requires imitation of the patterns of speech they hear others produce. Vocal learning critically depends on hearing the vocalizations of others and hearing oneself produce sound. This is true both for humans, who do not learn spoken language (or even babble normally) if they are deaf (87), and also for song birds (88). Production plays a role in normal language development; infants tracheostomized at the time at which they normally would babble show abnormal patterns of development that persist (89). These cases illustrate the strong dependency between perception and production and suggest why speech motor patterns learned early in life become difficult to alter later. Speakers who learn a second language after puberty produce it with an "accent" typical of their primary language, even after long-term instruction (90).

Imitation forges this early link between perception and production. By 1 year of age infants' spontaneous utterances reflect their imitation of ambient language patterns (91, 92). Laboratory studies indicate that the fundamental capacity to imitate sound patterns is in place even earlier. In a recent study, Kuhl and Meltzoff (93) 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/ for 5 min on each of 3 successive days. The results demonstrate developmental change between 12 and 20 weeks—by 20 weeks, there is clear separation between the three vowel categories for infants (Fig. 7). At this age, infants clearly imitate the model, and their vowels have appropriate formant frequency values in relation to one another, even though infants' vowels occur in a much higher frequency range (93).

Early theories of speech perception held that speech was perceived with reference to production (10). The developmental data suggest a different conclusion—early in life, perceptual representations of speech are stored in memory. Subsequently, these representations guide the development of motor speech. The two systems are thus tightly coupled early on, but the coupling is seen as a coregistration of auditory and motor information, a polymodal mapping, rather than one in which the representation is specified in motor terms. Perceptual experience that guides sensory-motor learning also is seen in infants' imitation of nonspeech oral movements (94, 95) and in sign language (96). The perception-action links observed for speech thus may rely on domain-general capabilities.

In related studies, infants also show an ability to link mouth movements they see to auditory signals they hear. Studies on 18- to 20-week-old infants show that they look longer at a face pronouncing a vowel that matches one they hear as opposed to a mismatched face (97). Infants' polymodal speech representations are thus likely to contain information regarding visual, as well as auditory instantiations of speech (ref. 98, see also refs. 99–101).

"Motherese" Is Instructive

Historically, language input was seen as a trigger for selecting among innately specified options. New data suggest that language addressed to infants plays a much more important role. The universal speaking style used by caretakers around the world when they address infants, often called "motherese" or "parentese" (102), has been shown to be preferred over adult-directed speech by infants given a choice (103, 104). Moreover, the exaggerated stress and increased pitch typical of infant-directed speech assists infants in discriminating phonetic units (105).

Infant-directed speech also is altered at the phonetic level and these alterations are argued to help infants learn. In a recent
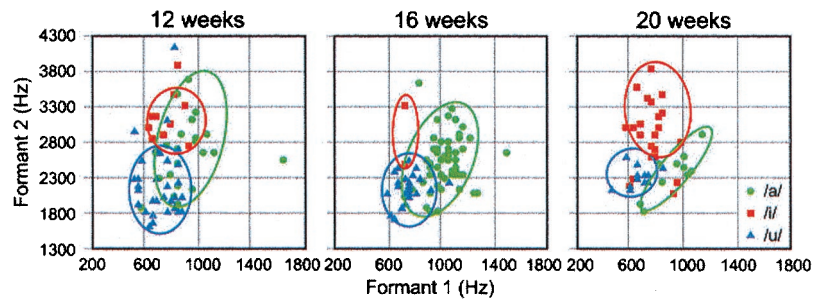
Fig. 7. Infants' vowels recorded as they imitate an adult show developmental change between 12 and 20 weeks of age. [Reproduced with permission from ref. 93 (Copyright 1996, Acoustical Society of America).]

study, women were recorded while speaking to their 2-month-old infants and to another adult in the United States, Russia, and Sweden (106). Mothers used the vowels /i/, /a/, and /u/, in both settings, and their speech was analyzed spectrographically. The results demonstrated that the phonetic units of infant-directed speech are acoustically exaggerated. The results show a stretching of the acoustic space encompassing speech (Fig. 8). Exaggerating speech not only makes it more discriminable for infants, it highlights critical parameters used in the native language. This may aid infants' discovery of the dimensions of sound used in their native language. Mothers addressing infants also increase the variety of exemplars they use, behaving in a way that makes mothers resemble many different talkers, a feature shown to assist category learning in second-language learners (107). In recent studies, language-delayed children show substantial improvements in measures of speech and language after listening to speech altered by computer to exaggerate phonetic differences (108, 109).

Mothers addressing infants make other adjustments that appear to aid learning. When introducing new words, parents repeat the word often in sterotyped frames ("Where's the ____," "See the ____," "That's a ____" (110), which would highlight the items in sentence-final position. They also present new words in a great variety of contexts, which would highlight the internal transitional probabilities of the new words against the backdrop of a variety of contexts (58). These new data suggest that the modifications made by adults unconsciously when they speak to infants plays a role in helping infants map native-language input. This represents a change in theoretical perspective with regard to the role of motherese in language acquisition.
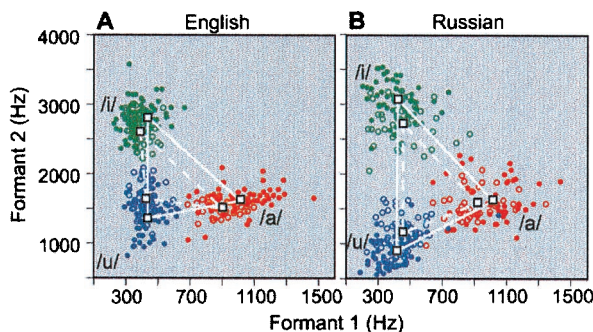

Fig. 8. Formant frequencies for vowels produced by American (*A*) and Russian (*B*) mothers as they talked to their infants (solid symbols) and to another adult (open symbols), indicating that vowels in infant-directed speech are acoustically exaggerated. [Reproduced with permission from ref. 106 (Copyright 1997, American Association for the Advancement of Science).]

The Critical Period for Language Learning Depends on Experience, Not Just Time

There is no doubt that children learn language more naturally and efficiently than adults, a paradox given adults' superior cognitive skills. The question is: Why?

Language acquisition often is cited as an example of a "critical period" in development, a learning process that is constrained by time, or factors such as hormones, that are outside the learning process itself. The studies on speech suggest an alternative (40, 82). The work on speech suggests that later learning may be constrained by the initial mapping that has taken place. For instance, if learning involves the creation of mental maps for speech, as suggested by the Native Language Magnet model (65, 82), it likely "commits" neural structure in some way. Measurements of brain activity, for example, confirm left-hemisphere effects for native-language sounds in the mismatched negativity (MMN), an event-related potential elicited by a change in a repetitive sound pattern (72). In infants, the MMN is observed to changes in both native and nonnative contrasts at 6 months of age. At 12 months of age, the MMN exists only for native language contrasts (73). Neural commitment to a 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.

Support for the neural commitment view comes from two sources, second language learning, and training studies. When acquiring a second language, certain phonetic distinctions are notoriously difficult to master both in speech perception and production, as shown, for example, by the difficulty of the /r-l/ distinction for native speakers of Japanese, even after training (11, 78, 111, 112). The hypothesis is that, for Japanese people, learning to process English requires the development of a new map, one more appropriate for English. New training studies suggest that exaggerating the dimensions of foreign language contrasts (36), as well as providing listeners with multiple instances spoken by many talkers (113), are effective training methods. These studies show that feedback and reinforcement are not necessary in this process; listeners simply need the right kind of listening experience (36, 113). Interestingly, the features shown to assist second-language learners—exaggerated acoustic cues, multiple instances by many talkers, and mass listening experience—are features that motherese provides infants.

Early in life, interference effects are minimal and two different mappings can be acquired, as is the case for infants learning two languages. Anecdotal evidence suggests that infants exposed to two languages do much better if each parent speaks one of the two languages, rather than 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 they can be perceptually separated. A second

language learned later in life (after puberty) may require another form of separation between the two systems to avoid interference. Data gathered by using functional MRI techniques indicate that adult bilinguals who acquire both languages early in life activate overlapping regions of the brain when processing the two languages, whereas those who learn the second language later in life activate two distinct regions of the brain for the two languages (114). This is consistent with the idea that the brain's processing of a primary language can interfere with the second language. The problem is avoided if both are learned early in development.

Conclusions

The framework that emerges from this research is very different from that held historically. Infants are neither the tabula rasas that Skinner described nor the innate grammarians that Chomsky envisioned. Infants have inherent perceptual biases that segment phonetic units without providing innate descriptions of them. They use inherent learning strategies that were not expected, ones thought to be too complex and difficult for infants to use. Adults addressing infants unconsciously modify speech in ways that assist the brain mapping of language. In combination, these factors provide a powerful discovery procedure for language. Six tenets of a new view of language acquisition are offered: (*i*) infants' initially parse the basic units of speech allowing them to acquire higher-order units created by their combinations; (*ii*) the developmental process is not a selectionist one in which innately specified options are selected on the basis of experience; (*iii*) rather, a perceptual learning process, unrelated to Skinnerian learning, commences with exposure to language, during which infants detect patterns, exploit statistical properties, and are perceptually altered by that experience; (*iv*) vocal imitation links speech perception and production early, and auditory, visual, and motor information are coregistered for speech categories; (*v*) adults addressing infants unconsciously alter their speech to match infants' learning strategies, and this is instrumental in supporting infants' initial mapping of speech; and (*vi*) the critical period for language is influenced not only by time, but by the neural commitment that results from experience.

Taken together, these principles suggest that what is innate regarding language is not a universal grammar and phonetics, but innate biases and strategies that place constraints on perception and learning. They allow infants to recover from language input the rules by which people in their community communicate. Language is thus innately discoverable, but not innate in the way that selectionist models suggested. The learning strategies used by infants may themselves have influenced the nature of language, in much the same way that general auditory processing influenced the selection of phonetic units for language during its evolution. The continued study of language development by infants promises to reveal the precise nature of the relationship between language and mind.

1. Skinner, B. F. (1957) *Verbal Behavior* (Appleton-Century-Crofts, New York).
2. Chomsky, N. (1957) *Language* **35,** 26–58.
3. Wexler, K. & Culicover, P. W. (1980) *Formal Principles of Language Acquisition* (MIT Press, Cambridge, MA).
4. Fodor, J. A. (1983) *The Modularity of Mind: An Essay on Faculty Psychology* (MIT Press, Cambridge, MA).
5. Stevens, K. N. (1998) *Acoustic Phonetics* (MIT Press, Cambridge, MA).
6. Pickett, J. M. (1999) *The Acoustics of Speech Communication* (Allyn and Bacon, Boston).
7. Gross, N., Judge, P. C., Port, O. & Wildstrom, S. H. (1998) *BusinessWeek* February **23,** 60–72.
8. Waibel, A. (1986) in *Pattern Recognition by Humans and Machines,* eds. Schwab, E. C. & Nusbaum, H. C. (Academic, New York), pp. 159–186.
9. Bernstein, J. & Franco, H. (1996) in *Principles of Experimental Phonetics,* ed. Lass, N. J. (Mosby, St. Louis), pp. 408–434.
10. Liberman, A. M., Cooper, F. S., Shankweiler, D. P. & Studdert-Kennedy, M. (1967) *Psychol. Rev.* **74,** 431–461.
11. Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J. & Fujimura, O. (1975) *Percept. Psychophys.* **18,** 331–340.
12. Eimas, P. D., Siqueland, E. R., Jusczyk, P. & Vigorito, J. (1971) *Science* **171,** 303–306.
13. Eimas, P. D. (1974) *Percept. Psychophys.* **16,** 513–521.
14. Eimas, P. D. (1975) *Percept. Psychophys.* **18,** 341–347.
15. Lasky, R. E., Syrdal-Lasky, A. & Klein, R. E. (1975) *J. Exp. Child Psych.* **20,** 215–225.
16. Streeter, L. A. (1976) *Nature (London)* **259,** 39–41.
17. Eimas, P. D. (1975) in *Infant Perception: Vol. 2. From Sensation to Cognition,* eds. Cohen, L. B. & Salapatek, P. (Academic, New York), pp. 193–231.
18. Kuhl, P. K. & Miller, J. D. (1975) *Science* **190,** 69–72.
19. Kuhl, P. K. & Miller, J. D. (1978) *J. Acoust. Soc. Am.* **63,** 905–917.
20. Kuhl, P. K. (1981) *J. Acoust. Soc. Am.* **70,** 340–349.
21. Kuhl, P. K. & Padden, D. M. (1982) *Percept. Psychophys.* **32,** 542–550.
22. Kuhl, P. K. & Padden, D. M. (1983) *J. Acoust. Soc. Am.* **73,** 1003–1010.
23. Dooling, R. J., Best, C. T. & Brown, S. D. (1995) *J. Acoust. Soc. Am.* **97,** 1839–1846.
24. Kluender, K. R., Diehl, R. L. & Killeen, P. R. (1987) *Science* **237,** 1195–1197.
25. Ramus, F., Hauser, M. D., Miller, C., Morris, D. & Mehler, J. (2000) *Science* **288,** 349–351.
26. Kuhl, P. K. (1991) in *Plasticity of Development,* eds. Brauth, S. E., Hall, W. S. & Dooling, R. J. (MIT Press, Cambridge, MA), pp. 73–106.
27. Kuhl, P. K. (1988) *Hum. Evol.* **3,** 19–43.
28. Miller, J. D., Wier, C. C., Pastore, R. E., Kelly, W. J. & Dooling, R. J. (1976) *J. Acoust. Soc. Am.* **60,** 410–417.
29. Pisoni, D. B. (1977) *J. Acoust. Soc. Am.* **61,** 1352–1361.
30. Jusczyk, P. W., Rosner, B. S., Cutting, J. E., Foard, C. F. & Smith, L. B. (1977) *Percept. Psychophys.* **21,** 50–54.
31. Werker, J. F. & Tees, R. C. (1984) *Inf. Behav. Dev.* **7,** 49–63.
32. Werker, J. F. (1995) in *An Invitation to Cognitive Science: Language,* eds. Gleitman, L. R. & Liberman, M. (MIT Press, Cambridge, MA), pp. 87–107.
33. Werker, J. F. & Logan, J. S. (1985) *Percept. Psychophys.* **37,** 35–44.
34. Carney, A. E., Widin, G. P. & Viemeister, N. F. (1977) *J. Acoust. Soc. Am.* **62,** 961–970.
35. Logan, J. S., Lively, S. E. & Pisoni, D. B. (1991) *J. Acoust. Soc. Am.* **89,** 874–886.
36. McClelland, J. L., Thomas, A., McCandliss, B. D. & Fiez, J. A. (1999) in *Brain, Behavioral, and Cognitive Disorders: The Neurocomputational Perspective,* eds. Reggia, J., Ruppin, E. & Glanzman, D. (Elsevier, Oxford), pp. 75–80.
37. Best, C. T., McRoberts, G. W. & Sithole, N. M. (1988) *J. Exp. Psych. Hum. Percept. Perform.* **14,** 345–360.
38. Best, C. T. (1995) in *Advances in Infancy Research,* eds. Rovee-Collier, C. & Lipsitt, L. P. (Ablex, Norwood, NJ), pp. 217–304.
39. Kuhl, P. K. (1994) *Curr. Opin. Neurobiol.* **4,** 812–822.
40. Kuhl, P. K. (2000) in *The New Cognitive Neurosciences,* ed. Gazzaniga, M. S. (MIT Press, Cambridge, MA), 2nd Ed., pp. 99–115.
41. Jusczyk, P. W. (1997) *The Discovery of Spoken Language* (MIT Press, Cambridge, MA).
42. Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996) *Science* **274,** 1926–1928.
43. Kuhl, P. K. (1985) in *Neonate Cognition: Beyond the Blooming Buzzing Confusion,* eds. Mehler, J. & Fox, R. (Erlbaum, Hillsdale, NJ), pp. 231–262.
44. Kuhl, P. K. (1979) *J. Acoust. Soc. Am.* **66,** 1668–1679.
45. Kuhl, P. K. (1983) *Inf. Behav. Dev.* **6,** 263–285.
46. Hillenbrand, J. (1983) *J. Speech Hear. Res.* **26,** 268–282.
47. Kuhl, P. K. (1980) in *Child Phonology: Vol. 2. Perception,* eds. Yeni-Komshian, G. H., Kavanagh, J. F. & Ferguson, C. A. (Academic, New York), pp. 41–66.
48. Jusczyk, P. W. (1999) *Trends Cognit. Sci.* **3,** 323–328.
49. Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J. & Amiel-Tison, C. (1988) *Cognition* **29,** 143–178.
50. Moon, C., Cooper, R. P. & Fifer, W. P. (1993) *Inf. Behav. Dev.* **16,** 495–500.
51. Nazzi, T., Bertoncini, J. & Mehler, J. (1998) *J. Exp. Psychol. Hum. Percept. Perform.* **24,** 756–766.
52. Lecanuet, J. P. & Granier-Deferre, C. (1993) in *Developmental Neurocognition: Speech and Face Processing in the First Year of Life,* eds. de Boysson-Bardies, B., de Schonen, S., Jusczyk, P., McNeilage, P. & Morton, J. (Kluwer, Dordrecht, The Netherlands).
53. DeCasper, A. J. & Fifer, W. P. (1980) *Science* **208,** 1174–1176.
54. DeCasper, A. J. & Spence, M. J. (1986) *Inf. Behav. Dev.* **9,** 133–150.

55. Jusczyk, P. W., Cutler, A. & Redanz, N. J. (1993) *Child Dev.* **64,** 675–687.
56. Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Cassidy, K. W., Druss, B. & Kennedy, L. (1987) *Cognition* **26,** 269–286.
57. Jusczyk, P. W., Friederici, A. D., Wessels, J. M. I., Svenkerud, V. Y. & Jusczyk, A. M. (1993) *J. Mem. Lang.* **32,** 402–420.
58. Goodsitt, J. V., Morgan, J. L. & Kuhl, P. K. (1993) *J. Child Lang.* **20,** 229–252.
59. Wolff, J. G. (1977) *Br. J. Psych.* **68,** 97–106.
60. Aslin, R. N., Saffran, J. R. & Newport, E. L. (1998) *Psychol. Sci.* **9,** 321–324.
61. Mattys, S. L., Jusczyk, P. W., Luce, P. A. & Morgan, J. L. (1999) *Cog. Psych.* **38,** 465–494.
62. Spelke, E. (1994) *Cognition* **50,** 431–445.
63. Grieser, D. & Kuhl, P. K. (1989) *Dev. Psych.* **25,** 577–588.
64. Kuhl, P. K. (1991) *Percept. Psychophys.* **50,** 93–107.
65. Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N. & Lindblom, B. (1992) *Science* **255,** 606–608.
66. Bosch, L., Costa, A. & Sebastian-Galles, N. (2000) *Eur. J. Cognit. Psychol.* **12,** 189–221.
67. Samuel, A. G. (1982) *Percept. Psychophys.* **31,** 307–314.
68. Miller, J. L. (1994) *Cognition* **50,** 271–285.
69. Iverson, P. & Kuhl, P. K. (1995) *J. Acoust. Soc. Am.* **97,** 553–562.
70. Aaltonen, O., Eerola, O., Hellström, A., Uusipaikka, E. & Lang, A. H. (1997) *J. Acoust. Soc. Am.* **101,** 1090–1105.
71. Sharma, A. & Dorman, M. F. (1998) *J. Acoust. Soc. Am.* **104,** 511–517.
72. Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., *et al.* (1997) *Nature (London)* **385,** 432–434.
73. Cheour-Luhtanen, M., Alho, K., Kujala, T., Sainio, K., Reinikainen, K., Renlund, M., Aaltonen, O., Eerola, O. & Näätänen, R. (1995) *Hear. Res.* **82,** 53–58.
74. Medin, D. L. & Barsalou, L. W. (1987) in *Categorical Perception: The Groundwork of Cognition,* ed. Harnad, S. (Cambridge Univ. Press, New York), pp. 455–490.
75. Mervis, C. B. & Rosch, E. (1981) *Annu. Rev. Psychol.* **32,** 89–115.
76. Iverson, P. & Kuhl, P. K. (2000) *Percept. Psychophys.* **62,** 874–886.
77. Strange, W. & Dittmann, S. (1984) *Percept. Psychophys.* **36,** 131–145.
78. Goto, H. (1971) *Neuropsychologia* **9,** 317–323.
79. Iverson, P. & Kuhl, P. K. (1996) *J. Acoust. Soc. Am.* **99,** 1130–1140.
80. Polka, L. & Bohn, O. S. (1996) *J. Acoust. Soc. Am.* **100,** 577–592.
81. Miller, J. L. & Eimas, P. D. (1996) *Percept. Psychophys.* **58,** 1157–1167.
82. Kuhl, P. K. (1998) in *Mechanistic Relationships Between Development and Learning,* eds. Carew, T. J., Menzel, R. & Shatz, C. J. (Wiley, New York), pp. 53–73.
83. Younger, B. A. & Cohen, L. B. (1985) in *The Psychology of Learning and Motivation,* ed. Bower, G. H. (Academic, San Diego), Vol. 19, pp. 211–247.
84. Saffran, J. R., Johnson, E. K., Aslin, R. N. & Newport, E. L. (1999) *Cognition* **70,** 27–52.
85. Quinn, P. C. & Eimas, P. D. (1998) *J. Exp. Child Psychol.* **69,** 151–174.
86. Gallistel, C. R. (1990) *The Organization of Learning* (MIT Press, Cambridge, MA).
87. Oller, D. K. & MacNeilage, P. F. (1983) in *The Production of Speech,* ed. MacNeilage, P. F. (Springer, New York), pp. 91–108.
88. Doupe, A. & Kuhl, P. K. (1999) *Annu. Rev. Neurosci.* **22,** 567–631.
89. Locke, J. L. & Pearson, D. M. (1990) *J. Child Lang.* **17,** 1–16.
90. Flege, J. E. (1993) *J. Acoust. Soc. Am.* **93,** 1589–1608.
91. de Boysson-Bardies, B. (1993) in *Developmental Neurocognition: Speech and Face Processing in the First Year of Life,* eds. de Boysson-Bardies, B., de Schonen, S., Jusczyk, P., McNeilage, P. & Morton, J. (Kluwer, Dordrecht, The Netherlands), pp. 353–363.
92. Vihman, M. M. & de Boysson-Bardies, B. (1994) *Phonetica* **51,** 159–169.
93. Kuhl, P. K. & Meltzoff, A. N. (1996) *J. Acoust. Soc. Am.* **100,** 2425–2438.
94. Meltzoff, A. N. & Moore, M. K. (1977) *Early Dev. Parent.* **6,** 179–192.
95. Meltzoff, A. N. & Moore, M. K. (1994) *Inf. Behav. Dev.* **17,** 83–99.
96. Petitto, L. A. & Marentette, P. F. (1991) *Science* **251,** 1493–1496.
97. Kuhl, P. K. & Meltzoff, A. N. (1982) *Science* **218,** 1138–1141.
98. Kuhl, P. K. & Meltzoff, A. N. (1997) in *The Inheritance and Innateness of Grammars,* ed. Gopnik, M. (Oxford Univ. Press, New York), pp. 7–44.
99. MacKain, K., Studdert-Kennedy, M., Spieker, S. & Stern, D. (1983) *Science* **219,** 1347–1349.
100. Rosenblum, L. D., Schmuckler, M. A. & Johnson, J. A. (1997) *Percept. Psychophys.* **59,** 347–357.
101. Walton, G. E. & Bower, T. G. R. (1993) *Inf. Behav. Dev.* **16,** 233–243.
102. Ferguson, C. A. (1964) *Am. Anthropol.* **66,** 103–114.
103. Fernald, A. (1985) *Inf. Behav. Dev.* **8,** 181–195.
104. Fernald, A. & Kuhl, P. (1987) *Inf. Behav. Dev.* **10,** 279–293.
105. Karzon, R. G. (1985) *J. Exp. Child. Psych.* **39,** 326–342.
106. Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U. & Lacerda, F. (1997) *Science* **277,** 684–686.
107. Lively, S. E., Logan, J. S. & Pisoni, D. B. (1993) *J. Acoust. Soc. Am.* **94,** 1242–1255.
108. Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L. & Tallal, P. (1996) *Science* **271,** 77–81.
109. Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., Schreiner, C., Jenkins, W. M. & Merzenich, M. M. (1996) *Science* **271,** 81–84.
110. Peters, A. M. (1983) *The Units of Language Acquisition* (Cambridge Univ. Press, Cambridge).
111. Flege, J. E., Takagi, N. & Mann, V. (1995) *Lang. Speech* **38,** 25–55.
112. Yamada, R. A. & Tohkura, Y. (1992) *Percept. Psychophys.* **52,** 376–392.
113. Pisoni, D. B. (1992) in *Speech Perception, Production and Linguistic Structure,* eds. Tohkura, Y., Vatikiotis-Bateson, E. & Sagisaka, Y. (Ohmsha, Tokyo), pp. 143–151.
114. Kim, K. H. S., Relkin, N. R., Lee, K. M. & Hirsch, J. (1997) *Nature (London)* **388,** 172–174.