PUBLISHED BY THE NEW YORK ACADEMY OF SCIENCES

Cover: "A World in a Grain of Sand" (after William Blake). A thin slice of rock from the earth's mantle, exposed in the Twin Sisters body, Cascade Mountains, Washington. Each grain of olivine (peridot when gem quality) is denoted by a consistent color, resulting from using a polarizing light filter above and below the thin section. Courtesy of Basil Tikoff, Geology and Geophysics, University of Wisconsin, Madison, Wisconsin.

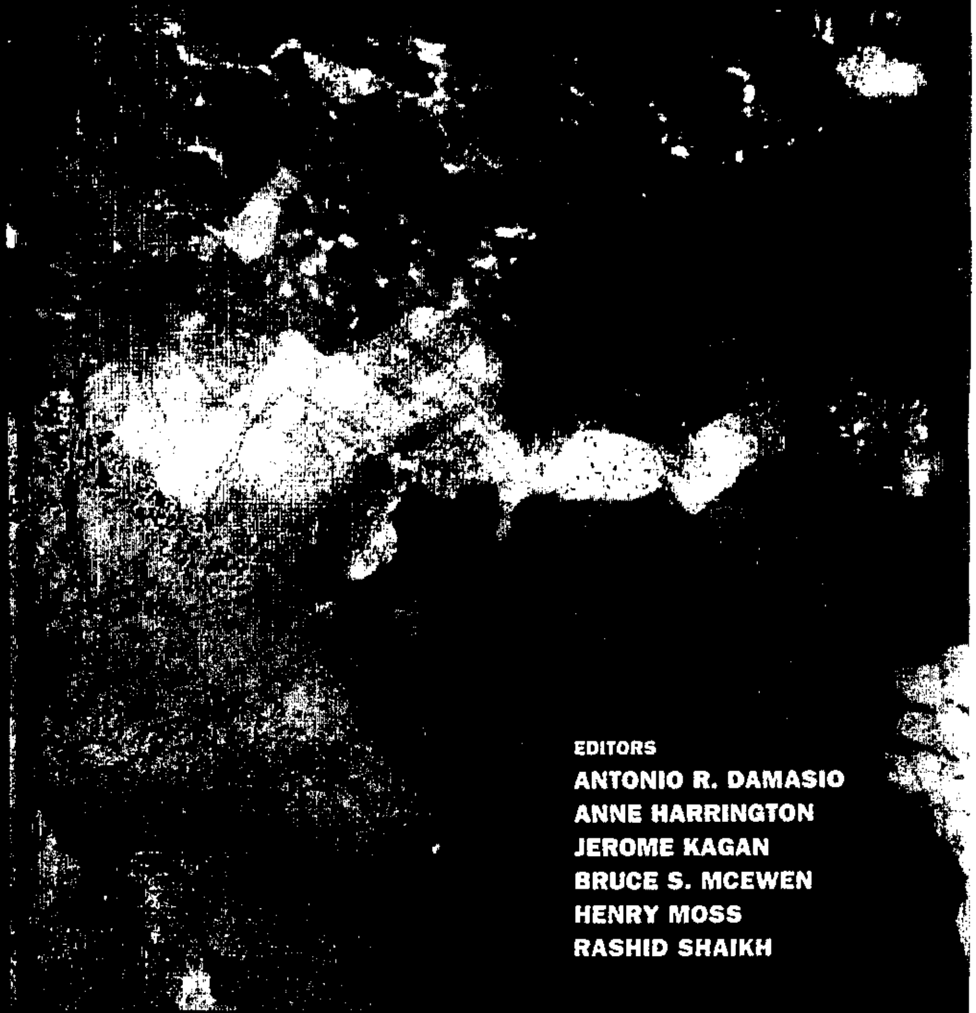
Cover design: Marguerite Z. Bunyan

ISBN 1-57331-311-4

UNITY OF KNOWLEDGE
THE CONVERGENCE OF NATURAL AND HUMAN SCIENCE
2001 ANN. N.Y. ACAD. SCI. VOL. 935

UNITY OF KNOWLEDGE

THE CONVERGENCE OF NATURAL AND HUMAN SCIENCE



EDITORS
ANTONIO R. DAMASIO
ANNE HARRINGTON
JEROME KAGAN
BRUCE S. MCEWEN
HENRY MOSS
RASHID SHAIKH

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES VOLUME 935

Language/Culture/Mind/Brain

Progress at the Margins between Disciplines

PATRICIA K. KUHL, FENG-MING TSAO, HUEI-MEI LIU, YANG ZHANG, AND BART DE BOER

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

ABSTRACT: At the forefront of research on language are new data demonstrating infants' strategies in the early acquisition of language. The data show that infants perceptually "map" critical aspects of ambient language in the first year of life before they can speak. Statistical and abstract properties of speech are picked up through exposure to ambient language. Moreover, linguistic experience alters infants' perception of speech, warping perception in a way that enhances native-language speech processing. Infants' strategies are unexpected and unpredicted by historical views. At the same time, research in three additional disciplines is contributing to our understanding of language and its acquisition by children. Cultural anthropologists are demonstrating the universality of adult speech behavior when addressing infants and children across cultures, and this is creating a new view of the role adult speakers play in bringing about language in the child. Neuroscientists, using the techniques of modern brain imaging, are revealing the temporal and structural aspects of language processing by the brain and suggesting new views of the critical period for language. Computer scientists, modeling the computational aspects of childrens' language acquisition, are meeting success using biologically inspired neural networks. Although a consilient view cannot yet be offered, the cross-disciplinary interaction now seen among scientists pursuing one of humans' greatest achievements, language, is quite promising.

KEYWORDS: Language development; Brain plasticity; Speech perception; Linguistic experience; Motherese; Informatics; Artificial intelligence; Brain imaging

INTRODUCTION

A large-scale scientific experiment is currently under way at universities around the world. Centers housing interdisciplinary research are being formed, supporting the widespread belief that collaborations among scientists with different basic training and approaches will promote cutting-edge discoveries. At the University of Washington, we have formed such a center, the Center for Mind, Brain, and Learning (CMBL, pronounced "symbol"), to foster cross-disciplinary collaboration on the study of human development.

Address for correspondence: Patricia K. Kuhl, Ph.D., Co-director, Center for Mind, Brain, and Learning, University of Washington, Box 357988, Seattle, WA 98105-6246. Voice: 206-685-1921; fax: 206-543-5771.
pkkuhl@u.washington.edu

CMBL scientists come not only from the fields of developmental psychology and the speech and hearing sciences, disciplines with a focus on children, but also include neuroscientists, computer scientists, geneticists, linguists, evolutionary and molecular biologists, anthropologists, engineers, informatics specialists, musicians, and educational researchers. Our goal is to understand how biology and culture cooperate to produce the young child's remarkable linguistic, cognitive, and social skills and the implications of those findings for society.

Language is an example of the kind of problem tackled by CMBL scientists. The last half of the 20th century has produced a revolution in our understanding of language and its acquisition. At the forefront of this revolution on language are new facts about its development in young infants. Behavioral studies of infants across cultures and languages, first undertaken in the early 1970s, have provided valuable information about the initial state of the mechanisms underlying speech perception, and more recently have revealed infants' use of unexpected learning strategies in the mastery of language during the first year of life. The learning strategies—demonstrating infants' statistical, probabilistic, distributional, and other computational skills—are unexpected and unpredicted by historical theories. The results are creating a new view of language and its relationship to the mind.

These findings on the psychology of language are enhanced by work in three other disciplines: cultural anthropology, computer science, and neuroscience. The goal of this chapter is to describe the research on language across the four disciplines and to discuss the extent to which a coherent view is emerging. Language is a defining characteristic of our species, a topic that has attracted a wide variety of scientists. It is a good test of the notion that consilience may emerge from cross-disciplinary collaboration.

The four disciplines take very different approaches to language. Psychologists study the babies themselves. Cultural anthropologists examine the ways in which adults across cultures play a role in bringing about language in the child. Computer scientists program computers to respond to language, and in the process try to model what babies do when they acquire language. Finally, neuroscientists examine the brain using new technology: positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and magnetoencephalography (MEG) in adult studies and event-related potentials (ERPs) in both infants and adults. The brain studies are expected to reveal not only the systems involved in language processing, but also how the brain codes language information, a topic with implications for critical periods in biology. Other branches of neuroscience, neurology and molecular biology, contribute through the study of individuals with language impairments, both inherent and acquired, with the expectancy that research will lead to models of treatment and intervention.

THE PSYCHOLOGY OF LANGUAGE

The last decade has produced surprising new data on language and its acquisition. Studies of infants across languages and cultures have produced a description of the innate state of the mechanisms underlying language, and more recently, have revealed infants' unexpected learning strategies once they are exposed to language. The learning strategies—demonstrating pattern perception, statistical learning, and a "warping" of perception brought about by exposure to a specific language—are not predicted by historical theories. The results have led 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.

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 careful parental 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.

The two approaches took strikingly different positions on all the critical components of a theory of language acquisition: (a) the initial state of knowledge, (b) the cause of developmental change, and (c) the role played by ambient language heard by the child. In Skinner's view, no innate information was necessary, developmental change was cause by reward contingencies, and language input could not, by itself, cause language to emerge. In Chomsky's view, infants possessed innate knowledge of language, development constituted growth of the "language module," and language input triggered a particular pattern from among those innately provided.

Tests on the Building Blocks of Speech: Phonetic Units

The original theorists had not studied babies, and when experimental tests were initiated, they focused on phonetic perception, perception of the consonants and vowels that make up words. Like grammar, phonetics demonstrates the dual patterning unique to language. A finite set of primitives (phonetic units, made up of subphonetic features) can be combined to create an infinite set of words or word-like strings that are legal combinations in a particular language (just as a finite set of words can be combined to create an infinite number of sentences).

Languages contain between 25 and 40 phonetic units, and their phonetic inventories differ dramatically. For example, French uses many "front-rounded" vowels, such as /y/ (produce American English "ee" and then hold that tongue position while rounding your lips to produce an "oo" vowel). English uses predominantly unrounded vowels, and Swedish uses both. Language perception requires learning which phonetic units are used in the language, and mastering the rules for their combination to form words.

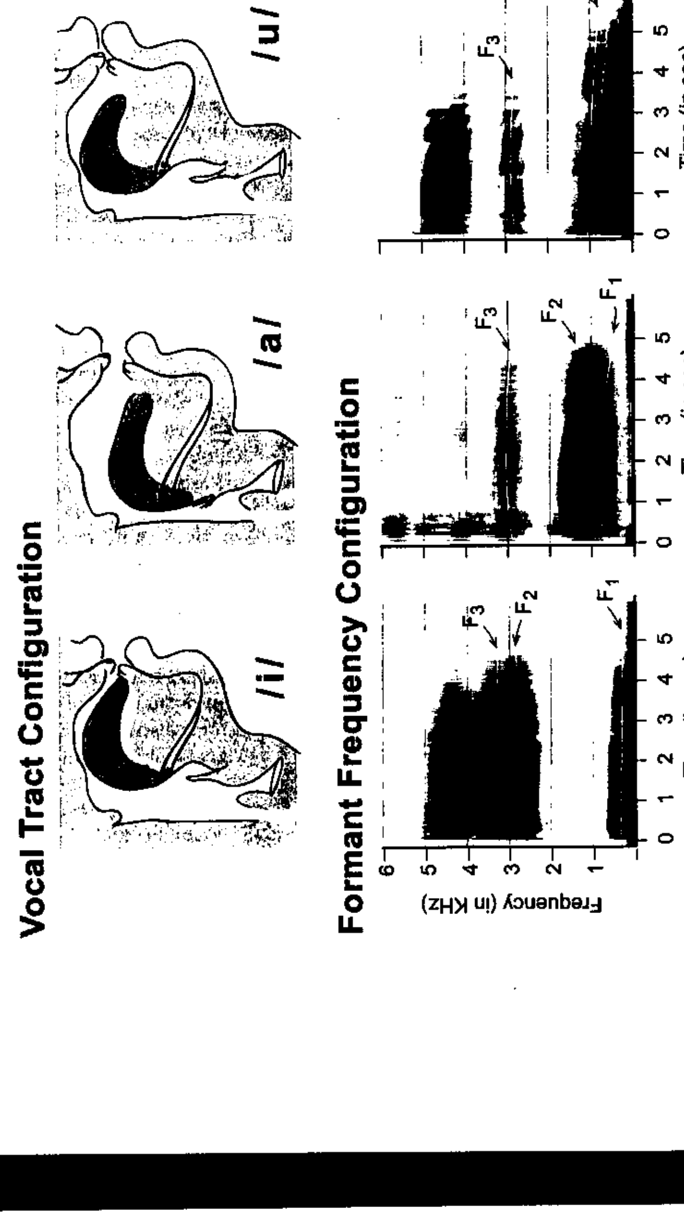


FIGURE 1. Vocal tract positions (*top*) and spectrographic displays (*bottom*) for the vowels /i/ as in "heat," /a/ as in "hot," and /u/ as in "hoot." Formant frequencies, regions in the frequency spectrum in which the concentration of energy is high, are marked for each formant.

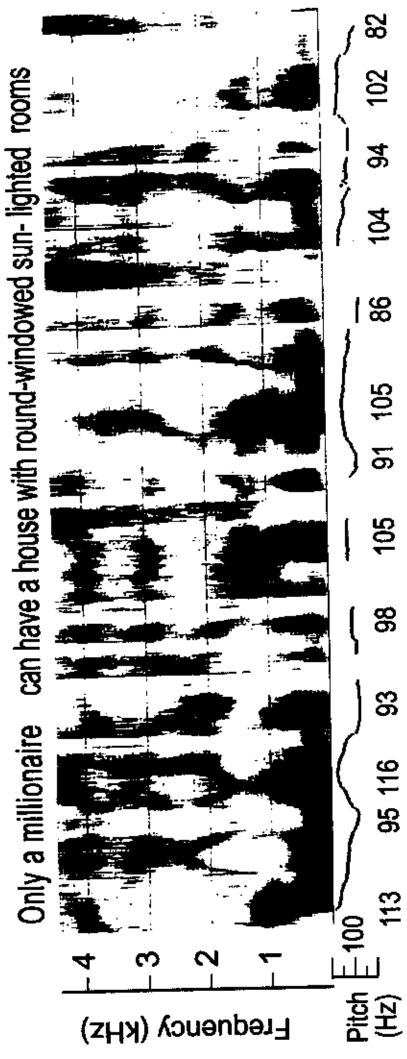


FIGURE 2. Display of running speech showing the formant frequencies and the fundamental frequency (pitch) of the voice over time. Increases in pitch mark primary stress in the utterance. (Reproduced with permission from Allyn and Bacon.[173])

Experiments have shown that the critical information listeners use to perceive phonetic units are contained in the formant frequencies of speech, regions of the frequency spectrum in which the concentration of energy is high. Formant frequencies can be easily seen in three vowels that are universal in the world's languages, /i/ as in "heat," /a/ as in "hot," and /u/ as in "hoot" (FIG. 1). Formant differences are much more difficult to discern, however, when close vowels such those in the words "cot," "caught," and "cat" are analyzed. Each talker's vocal tract differs in size and shape, so when close vowels like these are measured, the formant frequencies overlap considerably, making it impossible to specify an exact formant pattern that will be perceived as a particular vowel.[4] Moreover, phonetic units vary acoustically when placed in different phonetic contexts[5] or when spoken at different rates of speech,[6] factors that vary constantly in spoken speech. Infants have to overcome these problems to distinguish words in speech.

In ongoing speech, additional complications occur. Unlike written language, running speech has no breaks or pauses between words (FIG. 2). The formant frequencies found in isolated vowels (FIG. 1) are never in steady state when talkers actually speak. Prosodic cues come into play when syllables and words are strung together to form sentences. Across syllables, relative changes in pitch, loudness, and duration provide information about the overall stress pattern of words, patterns that are unique to a given language. In English, for example, words with a *strong/weak* pattern, with emphasis on the first syllable rather than the second, are typical (e.g., baseball, popcorn, apple, grandpa). Many other languages use a *fixed-stress* pattern which has either stress on the first syllable of every word (e.g., Finnish) or on the penultimate syllable (e.g., Polish).[7] Pitch changes over time identify the intonation contour of a sentence, which in English differentiates between a question and a statement. Intonation also marks syntactic junctures, such as the ends of major phrases and clauses, and these landmarks aid speech processing (e.g., see Wingfield, Lombardi, and Sokol[8]). In other languages, pitch and intonation are used differently; in Mandarin Chinese, for example, pitch changes alone can signal a change in word meaning.[9]

Initial Findings and Theories

Any theory of language acquisition has to specify how infants parse the auditory world to make the critical units of language available. The first experiments on infants asked whether infants could discern differences between the phonetic units used in the world's languages. Experiments in the '70s and '80s demonstrated convincingly across laboratories that infants were adept at speech discrimination. Virtually every phonetic contrast tested was shown to be discriminated by infants (see Jusczyk[10] for summary). Interestingly, however, the data also demonstrated that the partitioning of the phonetic units of speech is not limited to humans nor, when nonspeech is used with human listeners, limited to speech.

The evidence derived from tests of categorical perception.[5] 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 employed 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] and, in tests of discrimination, showed peaks in sensitivity that coincided with the phonetic boundaries employed by languages.[20–22] 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 original comparative work.[26] First, infants' parsing of the phonetic units at birth appears to be a discriminative capacity that can be accounted for by a general auditory processing mechanism, rather than one that evolved specifically for language. Infants' abilities to differentiate 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.[18,19,27] (See also Ramus *et al.*[25]) Second, in the evolution of language, acoustic differences detected by the auditory perceptual processing mechanism were seen as providing a strong influence on the selection of phonetic units used in language. In this view, particular auditory dimensions and features were exploited in the evolution of the sounds used in languages.[18,26,27] This ran counter to two prevailing views at the time: (a) the view that phonetic units themselves were innately prespecified in infants, and (b) the view that language evolved in humans without continuity with lower species.

Categorical perception was also 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, rather than language-specific ones, were responsible for infants' initial partitioning of the phonetic units of language.

Is Developmental Change Based on Selection?

Early models of speech perception were selectionist in nature. Theorists argued that an innate neural specification of all possible phonetic units allowed selection of a subset of those units to be triggered by language input.[17,31] The notion was that linguistic experience produced either maintenance or loss. Detectors stimulated by ambient language were maintained, while those not stimulated by language input atrophied.

Developmental studies on infants initially supported this view. Werker and her colleagues demonstrated that, by 12 months of age, infants no longer discriminate nonnative phonetic contrasts, even though they did so at six months of age.[32] The finding was intially interpreted as a "loss" of a subset of phonetic units initially specified, supporting the selectionist view. Further studies indicated there is not an immutable loss of phonetic discrimination for nonnative units; adults do not perform at chance when tested on foreign language contrasts.[33] These findings on adults left open the possibility that the "maintenance" aspect of the selectionist model was correct. Infants started life with a capacity for phonetic discrimination that was either maintained or decreased, depending on linguistic experience.

The alternative to an "innate-specification + maintenance" view is that infants begin with keen discriminative capacities but not an innate specification of phonetic units. What happens next, in this view, is that infants engage in an extraordinary "mapping" of language input.[34–36] This takes a different view of the developmental process, one in which exposure to ambient language is critical and infants' capacities to perceive order in language input, extraordinary.

To become a serious contender as a view of language acquisition, this view required evidence that infants exhibited genuine developmental change, not merely maintenance of an initial ability. To test the maintenance notion, studies were conducted in my laboratory on infants between 6 and 12 months of age to examine the course of development using a more sensitive measure than had been used in previous studies. The original developmental studies, conducted by Werker and Tees,[32] used a "threshold" measure of performance, determining how many infants passed threshold at each age. As shown in these early studies, a large number of infants passed the threshold measure at both 6 and 12 months of age for native-language contrasts. When nonnative contrasts were tested, large numbers of infants passed the criterion only at 6 months of age; they failed at 12 months of age. The method allowed one to show either maintenance or decline in native-language abilities. However, the method used could not demonstrate growth. In other words, the method used could not reveal whether infants listening to native-language contrasts demonstrated a pattern of developmental growth between 6 and 12 months of age. A pattern of increasing performance over time would suggest that infants are engaged in some other kind of learning process, a process that is not fundamentally subtractive in nature.

Two new studies on developmental speech perception provide evidence that infants show developmental growth, rather than maintenance, during this developmental period. The pattern of growth is seen only for native-language contrasts. In two studies, one conducted using American English /r/ and /l/ sounds and American and Japanese infants at 7 and 11 months[37] and one conducted using Mandarin Chinese fricative-affricate sounds and American and Taiwanese infants,[38] an identical pattern of findings was shown (FIG. 3). At 7 months, infants listening to native-language sounds (American infants listening to English /r/ and /l/ or Chinese infants listening to Chinese sounds) performed identically to infants for whom the contrast was foreign (Japanese infants on /r-l/ and American infants on the Chinese sounds). Both were well above chance. At 11 months, a change occurred in both native and nonnative discrimination. Infants listening to their native contrast showed a significant increase in performance, whereas infants listening to a foreign-language contrast demonstrated a decrease in performance. The data on native-language perception did not conform to a maintenance model. During this four-month window, infants appear to be engaged in a remarkable period of learning, one not accounted for by a Skinnerian learning model.

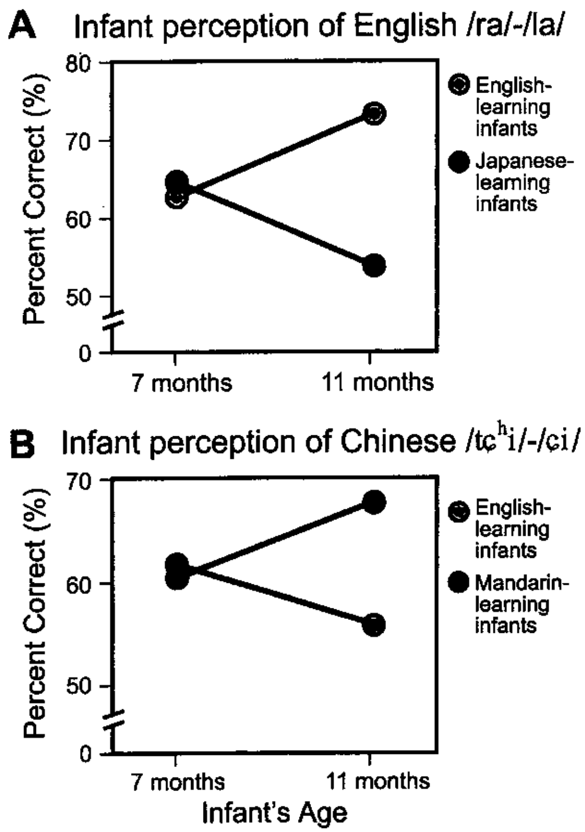


FIGURE 3. Results of two studies of infants' developmental transition in speech erception. Infants in (A) were tested in the United States and Japan on American English ›unds; infants in (B) were tested in Taiwan and the United States on Chinese sounds. In ›th cases, infants showed a significant increase in native-language phonetic perception and decrease in foreign-language phonetic perception over time.

NEW THEORIES OF LEARNING

The data on infants demonstrated that during the first 12 months of life, infants :quired an extraordinary amount of information describing their native language. ıfants' abilities required another explanation, one that described how experience ›uld have such an effect. General learning theory had been dismissed as an expla- ıtion for language acquisition on logical grounds; it could not explain the facts of nguage development.[2] At present, however, learning models—though not Skinne-

rian in nature—are at the center of current debates on language.[10,34,35,39] What has changed? The discoveries of the last decade go beyond the demonstration cited above that something other than maintenance is going on. The studies demonstrate that by simply listening to language infants acquire sophisticated information about its properties. These findings have sent theorists back to the design board to create new views of learning.

Three important organizing principles have emerged. First, infants show an extraordinary ability to detect patterns—regularities—in language input. They organize input to recognize similarities and form categories. Second, infants exploit statistical properties of the input, enabling them to detect and use distributional and probabalistic properties of the incoming signals. And third, infant perception is altered, literally "warped," by exposure to language in a way that promotes perception.

Infants Are Pattern Detectors

Critical to language perception is the ability to recognize similarity among the patterns spoken by different talkers in different contexts, a stumbling block for computer speech recognition.[40] Infants demonstrate excellent skills at pattern recognition for speech. A number of studies have shown that six-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.[41] Infants at 6 months of age are readily trained on this task (FIG. 4).

Data from studies of these kinds show, for example, that infants perceptually sort vowels that vary across talkers and intonation contours after training with one example from each category (FIG. 5).[42,43] Infants perceptually sort syllables that vary in their initial consonant (those beginning with /m/ as opposed to /n/) across variations in talkers and vowel contexts.[44,45] 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 /N/, as opposed to a set of stop consonants, /b/, /d/, and /g/.[44] Recent tests show that nine-month-old infants are particularly attentive to the initial portions of syllables.[46]

Infants' detection of patterns is not limited to phonetic units. More global prosodic patterns contained in language are also detected. At birth, infants have been shown to prefer the language spoken by their mothers during pregnancy, as opposed to another language.[47–49] 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.[50] 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's voice at birth[51] and their preference for stories read by the mother during the last 10 weeks of pregnancy.[52]

Between 6 and 9 months, infants exploit prosodic cues to detect patterns related to the stress or emphasis typical of words in their native language. In English, for example, a strong/weak pattern of stress, as exhibited in the words "baby," "mommy," and "table," is typical; whereas in other languages, a weak/strong pattern predominates. Infants tested at 6 months show no listening preference for words with

FIGURE 4. Infants being tested in a speech perception task. In (**A**), infants focused on the toys held by an assistant while a speech sound repeats from the loudspeaker on the right. In (**B**), the speech sound was changed, and infants' orienting responses were reinforced with the presentation of an animated toy positioned above the loudspeaker. Infants enjoyed the reinforcer and learned to produce head turns whenever they perceived a change in the sound.

the strong/weak as opposed to the weak/strong pattern, but by 9 months they exhibit a strong preference for the pattern typical of native language words.[53] Infants also use prosodic cues to detect major constituent boundaries, such as clauses. At four months of age, infants listen equally long to Polish and English speech samples that have pauses inserted at clause boundaries as opposed to pauses inserted within claus-

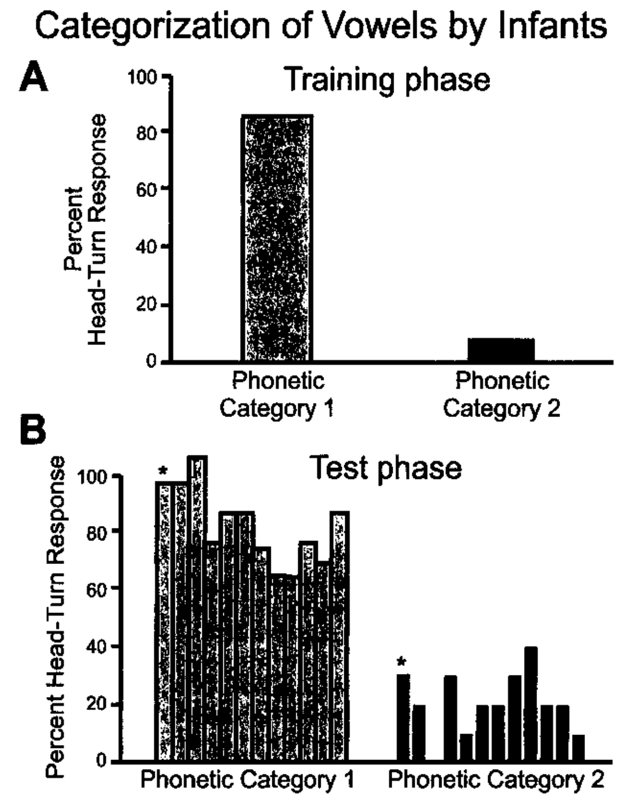


FIGURE 5. Performance on a vowel classification task by 6-month-old infants tested with the head-turn technique. Infants were trained to produce a head-turn to one of the two stimuli, the /a/ vowel (as in "cot"), and not the other, the /ae/ vowel (as in "cat") (**A**). After training, 22 novel stimuli were presented to see which elicited head-turn responses. The data show a high accuracy rate on the first presentation of the new stimuli (**B**).

es, but by 6 months, infants listen preferentially to pauses inserted at the clause boundaries appropriate only to their native language.[10,54]

By nine months of age, infants detect patterns related to the orderings of phonemes that are legal for their language. In English, for example, the combinations *zw* or *vl* are not legal, whereas in Dutch, they are permissible. By nine months of age, but not at six months of age, American infants listen longer to English lists, whereas Dutch infants show a listening preference for Dutch lists.[55]

Infants Exploit Statistical Properties of Language Input

One of the surprises of the last decade is the degree to which infants' skills go beyond perceptual organization. New research shows that infants exploit the statistical properties of the language they hear to identify likely word candidates. For example, transitional probabilities among syllables can be used as a strategy to identify word candidates. This is important because, unlike written text, there are no breaks between words. Within words, the transitional probabilities are high because the ordering of phonetic elements remains constant. In the word "potato," for example, the syllable "ta" follows the syllable "po" with a probability of 1.0. Between words, as in the string "hot potato," the transitional probability between the syllable "hot" and "po" is much lower. If infants are sensitive to transitional probabilities across syllables and the asymmetries that occur within as opposed to between words, they may recognize that points of high transitional probabilities are likely candidates for words, while points of low transitional probabilities are the boundaries between words. Goodsitt, Morgan, and Kuhl[56] demonstrated that 7-month-old infants could take advantage of transitional probabilities in artificial words.

Goodsitt *et al.*[56] 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 syllable, /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). Transitional probabilities between /ko/ and /de/ are no higher in this case than in (b).

The results confirmed the hypothesis. In two experiments, 7- to 8-month-old infants were shown to exploit the transitional probabilities in the syllable strings, allowing them to discriminate the target syllables in condition (a) significantly more accurately than in either (b) or (c), the latter of which showed equally poor discrimination (FIG. 6). This strategy has also been shown to be effective for adults presented with artificial nonspeech analogues created by computer.[57]

In further work, Saffran *et al.*[39] directly assessed eight-month-old infants' abilities to learn pseudo-words based on transitional probabilities using a listening preference technique. Infants were exposed to two-minute strings of synthetic speech composed of four different pseudo-words, like "tibudo," "pabiku," "golatu," and "daropi," which 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. The question was whether or not infants picked up the statistical regularities associated with the syllables contained in the words. During the test phase, infants listened to four isolated words, two from the original set of psuedo-words and two new words formed by combining the last syllable of one of the original words with the two initial syllables of another of the original words, an item like "tudaro," formed with the last syllable of "golatu" and the first two of

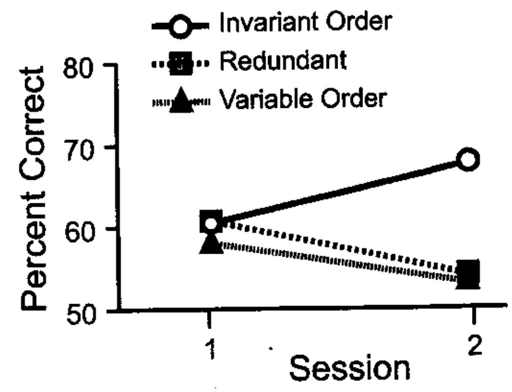


FIGURE 6. Performance on a statistical learning task by 7-month-old infants. Infants in the Invariant Order group, in which the transitional probabilities between syllables were high, showed an advantage in performance.

"daropi." Infants' listening preferences showed that they detected the statistical regularities present in the original set of words by listening longer to the new words, revealing a novelty preference. Studies show that this sensitivity is not due to infants' calculation of raw frequency of occurrence, but to the actual probabilities relating to the sequences of sounds.[58]

Additional examples of the computation of probability statistics have now been uncovered. For example, nine-month-old infants detect the probability of the occurrence of the legal sequences that occur in English.[59] Certain legal combinations of phonetic units (two consonants) are more likely to occur within words while others occur at the juncture between words. The combination "ft" is more common within words while the combination "vt" is more common between words. Nine-month-olds were tested with CVCCVC items that contained embedded CCs that were very frequent in English or with CCs that were infrequent in English. Infants tested in the listening preference paradigm were shown to listen significantly longer to the lists with the frequent within-word CCs embedded in them. However, when a prosodic cue signaling a word boundary was inserted into the sequences, either a half-second pause inserted between the two consonants or a weak/strong stress pattern synthesized instead of the strong/weak pattern that is typical for English, infants' listening preference switched to the sequences with the between-word CCs embedded in them.[59]

Taken collectively, these studies reveal infants' recognition of patterns contained in linguistic input and their ability to perform certain statistical analyses on that input. From the standpoint of theory, these forms of learning clearly do not involve Skinnerian reinforcement. Caretakers do not manage the contingencies and gradually, through reinforcement strategies, shape the pattern detection and statistical analyses performed by infants when they engage in this kind of learning. Infants do this naturally and automatically. The learning engaged in by infants acquiring language is also not well described by a selectionist model. The knowledge acquired by infants does not appear to derive from innately given options, with language input serving merely to maintain or produce atrophy in existing mechanisms. Rather, the process is characterized by a detailed analysis of language input using strategies that reveal the predictable patterns of variation in native language.

Language Experience Warps Infant Perception

Ambient language experience not only produces a change in infants' discriminative abilities and listening preferences, but also results in a "mapping" that alters their 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.[26,60–63] Many behavioral[26,60,62,64,65] and brain[66–69] studies indicate that native-language phonetic prototypes evoke special responses when compared with nonprototypes from the same phonetic category.

The notion that categories have prototypes stems from cognitive psychology.[70] Findings in that field show that the members of common categories show a graded structure such that some members of the category are more representative than others (a robin is a prototype of the category bird; an ostrich is not). Prototypes of categories are easier to remember, show shorter reaction times when identified, and are often preferred in tests that tap our favorite instances of categories (see Kuhl[26,62] for discussion).

The speech tests on infants demonstrated that when tested with a phonetic prototype, as opposed to a nonprototype from the same category, infants showed greater ability to generalize to other category members.[26,61] The prototype appeared to function as a "magnet" for other stimuli in the category, in a way similar to that shown for prototypes of other cognitive categories.[70–73]

Moreover, the perceptual magnet effect depends on exposure to a specific language.[63] 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. 7A), 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. 7B). The results show that by six months of age, perception is altered by language experience.

Infants' perceptual representations of speech in turn alter speech production. In humans (as in songbirds; see Doupe and Kuhl[74] for comparisons), sensory representations of speech serve as a guide for the motor production of speech. In the laboratory, our studies demonstrate that 15 min of listening experience is sufficient in a 20-week-old baby to induce vocal imitation of vowels.[75] These data suggest that perception and action are deeply connected at an early age and that infant babbling is not idle play. Babbling and the auditory stimulation it provides allow infants to map the relationship between speech motor movements and sound, a requirement for vocal imitation.

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.[26,62]

In adults, the distortion of perception caused by language 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

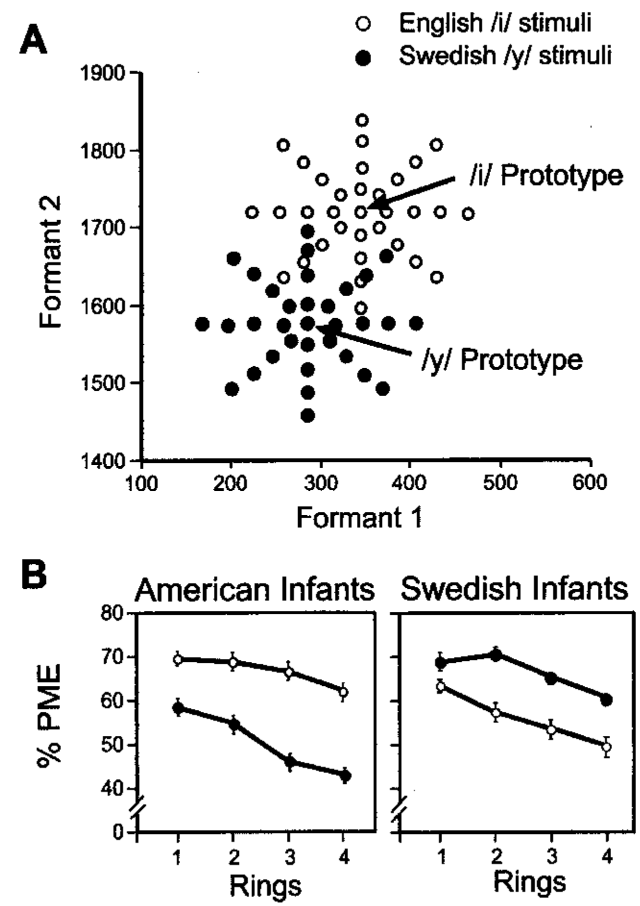


FIGURE 7. Results of a cross-language study demonstrating an effect of linguistic experience in 6-month-old infants. (A) Infants in two countries were tested with two vowel prototypes (and their variants), English (/i/) and Swedish (/y/). **(B)** Results demonstrated that infants showed greater generalization (a greater PME, or perceptual magnet effect) to variants around prototypes contained in their native language.

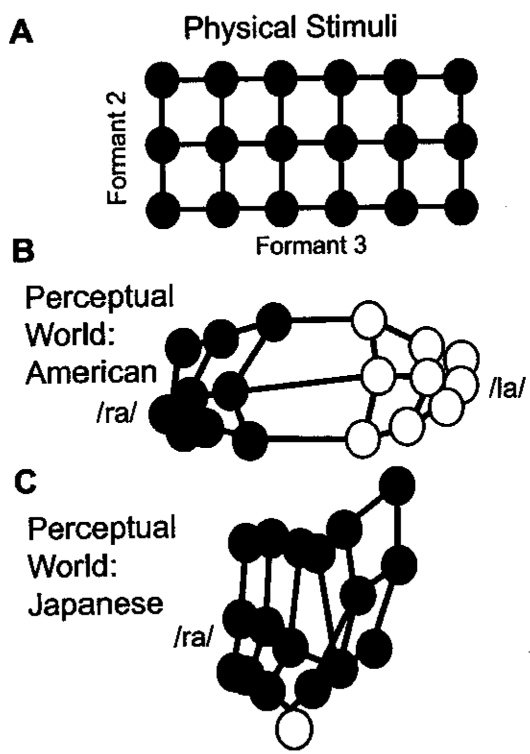


FIGURE 8. Results of tests on the perception of American English /r-l/ sounds by American and Japanese listeners. (A) Equal physical distances were created between syllables varying in Formants 2 and 3. (B) Perceptual distances as revealed by multidimensional scaling showed that American listeners perceived two distinct categories with a boundary between them. (C) Perceptual distances perceived by Japanese listeners were very different. In both cases, linguistic experience warps perception.

them at equal physical intervals in a two-dimensional acoustic grid (FIG. 8A). 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 using multidimensional scaling (MDS) 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. 8B), 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 (FIG. 8C). The map revealed by MDS 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.[36,64,79]

The important point regarding development is that the initial perceptual biases shown by infants in tests of categorical perception,[12,13,15–17] 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.[63]

A model reflecting this developmental sequence from universal perception to language-specific perception, called the Native Language Magnet (NLM) model, proposes that infants' mapping of ambient language input warps the acoustic dimensions underlying speech, producing a complex network, or filter, through which language is perceived.[36,38,82] 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. According to NLM, infants' transition in speech perception between six and twelve 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? 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 using linguistic and nonlinguistic events. What about animal tests? Thus far, the data suggest differences between animals and man on these kinds of learning. For instance, monkeys do not exhibit the perceptual magnet effect.[62] Animals do show some degree of internal structure for speech categories after extensive training,[24] but studies have not examined whether the perceptual magnet effect would be spontaneously produced in an animal after 6 months of experience listening to language, as seen in human infants. Similarly, animals are sensitive to transitional probabilities,[86] but it has yet to be determined whether an animal would spontaneously exhibit statistical learning after simply listening to language, as human infants have been shown to do. It is unlikely that animals would demonstrate these feats of unprompted learning. Empirical tests can resolve the issue.

CULTURAL ANTHROPOLOGY

One of the puzzles of language development is a behavior exhibited by adults when they talk to their children. As discovered by linguists and anthropologists in the early 1960s, parents in many of the world's cultures use a special speaking style when addressing their infants and young children. This form of speech has been dubbed "motherese" or "parentese."[87,88] Motherese is easily recognized because of its unique acoustic signature. Infant-directed utterances have, from an acoustic standpoint, a unique prosodic structure. It is characterized by a higher pitch, slower tempo, and exaggerated intonation contours. Measurements of infant-directed speech show that adults increase their habitual pitch by at least an octave and that the sing-song intonation contours contained in these signals are universal.[89,90] Motherese is used by mothers as well as fathers, and also caretakers who are not themselves parents.[91]

Whereas early anthropologists and linguists noted the unique prosodic pattern contained in motherese, it was only recently that we discovered that motherese speech is also modified at the phonetic level, in a way that aids infant learning. These data suggest a much more important role for infant-directed speech than was hypothesized by either of the historical theorists.

In the recent study, women were recorded while speaking to their two-month-old infants and to another adult in the United States, Russia, and Sweden.[92] Mothers used the vowels /i/, /a/, and /u/ in natural exchanges with their infant and with the other adult. The women were recorded, and their speech was analyzed spectrographically. The results demonstrated that the phonetic units of infant-directed speech are acoustically exaggerated. The acoustic components of the /a/, /i/, and /u/ vowels were altered such that the acoustic differences between them were stretched. This resulted in a greater acoustic area encompassing the vowel space (FIG. 9). Exaggerating speech in this way resulted in clearer sounding words, words that were more distinct from one another. The data demonstrated that all women of all cultures used an expanded vowel triangle when speaking to their infants. The expansion produced by mothers was substantial. It increased the area of the vowel space by a factor of 2:1, a factor that substantially increases the "working space" that speakers use when addressing young children.

Mothers' speech has been measured in many languages, including Mandarin Chinese, a language that uses pitch (or tone) to change the meaning of a word. In English, the pitch of the word does not change its meaning, but specifies the mood of the speaker or whether the speaker is asking a question or making a statement. In Mandarin Chinese, a change in pitch can signal a change in the meaning of the word. We therefore wondered whether Chinese mothers would exhibit the motherese pattern. Our initial experiments confirmed that Chinese mothers spoke motherese to their infants and children.[90] In a very recent study,[93] we extended that finding to confirm not only that Chinese mothers expand their vowel spaces when addressing children, but, more surprisingly, that they also expand the range used for the four tones of Mandarin Chinese when addressing their infants (FIG. 10). This finding suggests that motherese is a very general strategy that involves the exaggeration of the critical features in speech when addressing children.

Two questions frequently asked by parents, and relevant to theory, are: Why do we do this? And, is it of any consequence to the child?

Addressing the relevance issue first, the consequences to the child may be considerable. First, there is strong evidence from studies of typical and atypical speakers, that is, speakers with dysarthria, cerebral palsy, and amyotrophic lateral sclerosis, that exaggerated speech, as measured by the size of the vowel triangle in

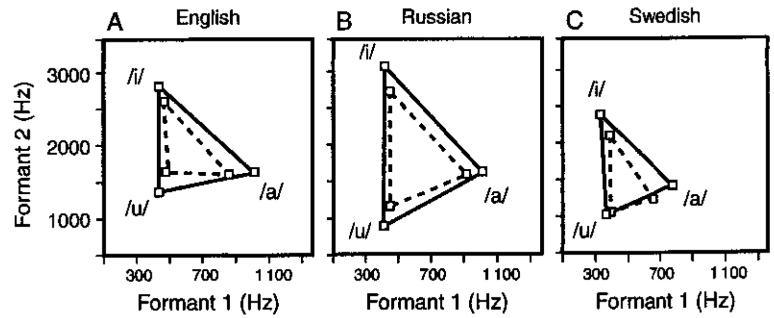


FIGURE 9. Motherese. American, Russian, and Swedish mothers speaking to their infants (*solid lines*) as opposed to other adults (*dashed lines*) exaggerated the formant frequencies and produced clearer speech.

these studies, is a strong predictor of higher speech intelligibility.[94] That is, a larger vowel space results in clearer speech.

Clear speech makes words more discriminable for infants, it highlights the critical acoustic parameters employed in the infant's native language, and it provides "prototypical" phonetic units to infants. Recall that work in our laboratory suggests that "prototypical" phonetic units aid infants' organization of speech categories.[63]

It is also relevant that, when given a choice, infants prefer to listen to motherese speech over adult-directed speech. Infants allowed to choose (by turning their heads one way or another) turn on the voices of mothers talking to infants.[95] Further studies revealed that infants' preference is dictated by the exaggerated pitch contours of the motherese speech.

Mothers addressing infants also increase the variety of exemplars they use, behaving in a way that makes mothers resemble many different talkers, a feature that has been shown to assist category learning in second-language learners.[96] 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.[97,98] Taken together, the results suggest that what adults do when talking to children makes a difference. This represents a change in theoretical perspective with regard to the role of motherese in language acquisition.

Turning to the second issue: What makes adults speak motherese? Adults are unaware of the adjustments they make when talking to children. When the phenomenon is pointed out to them, adults are often embarrassed, inquire why they do it, and ask whether it is good for their children.

Motherese may simply be one manifestation of a general listener-oriented communication system that evolved to focus on listeners. That is, if speech evolved to allow us to communicate information, adequacy is defined by the listener.[99] Speakers may have learned to automatically adjust their speech to meet the needs of the listener. A parallel that has not to our knowledge been drawn before, but one that we think apt, is the well-known "Lombard effect."[100] The Lombard effect is an automatic adjustment in the loudness of speech that speakers produced to counteract back-

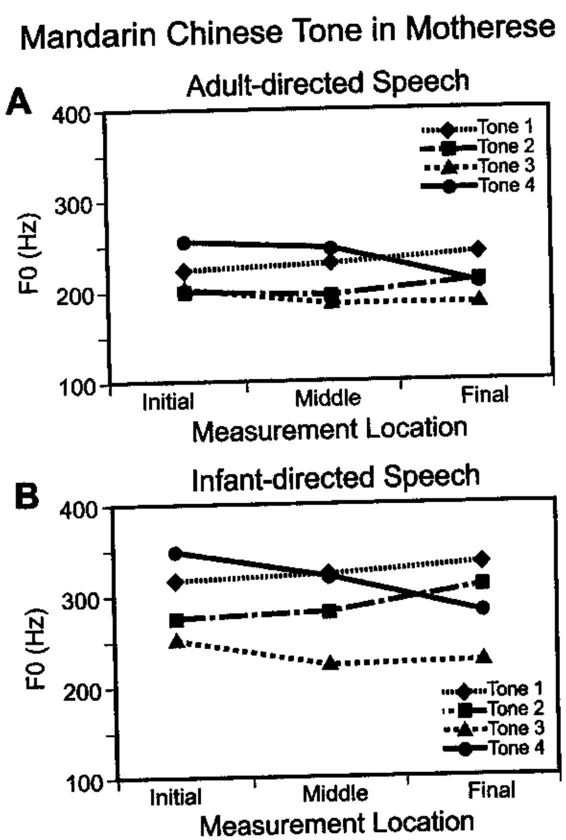


FIGURE 10. Mandarin Chinese tone in motherese. Taiwanese mothers speaking Mandarin Chinese to their infants (*bottom*) as opposed to other adults (*top*) exaggerate the tones used to convey meaning in Chinese.

ground noise. We all speak more loudly at a cocktail party when the background noise can be very high. Another example of this is that speakers listening to music over headphones increase the level of their voices (to the annoyance of listeners around them), typically without realizing that they are doing so.

The tendency to produce motherese may thus involve the same kind of unconscious adjustment to speech that is necessary to communicate with the listener. This framework would predict that speakers constantly adjust their speech (their vowel space areas, for example) to take into account the need for clarity in given situations. When giving a lecture, for example, for which clarity is critical, vowel spaces are likely quite large as we attempt to communicate with those whose understanding is not guaranteed. To friends and family in routine situations, our vowel spaces may

shrink considerably, as the communicative goal is more likely met with less effort. If this framework is correct, then "motherese" may not be a special adaptation for infants and children, but part of our biological tendency to speak as clearly as the situation demands. The prosody of motherese may play a different role than the exaggerated clear speech we have been describing. The raised pitch of motherese attracts infant attention and also conveys affection, and may be used as an acoustic hook for infants; whereas the exaggerated phonetic units of motherese may play an instructive role.

Of some import to this conception is the anecdotal information that in some cultures, mothers and fathers seldom speak directly to (or look directly at) their preverbal infants. Among the Kawara'ae of the Solomon islands, for example, it has been reported that mothers speak to infants only indirectly.[101] In New Guinea, Kaluli adults are said seldom to speak to preverbal infants.[102] Babies of Kawara'ae and Kaluli would be expected to hear a great deal of ambient language, but not language that is addressed to them directly. It is therefore unlikely that infants would hear motherese. Unfortunately, there are no acoustic data to examine and no data on language development in these cultures.

It has also been reported that there is no special prosodic register for babies and young children among American Indians speaking Quiche Mayan.[103] A raised vocal pitch in this culture is a general sign of deference and is not used with infants. It would be theoretically interesting if the voice pitch aspect of motherese, the affective, social aspect of the signal, is relatively less important than the exaggeration of the phonetic units. We know, for example, that people addressing their pets use pitch increases typical of motherese, but do not produce exaggerated vowels.[104] This suggests that the exaggeration seen in motherese is produced only when we have an expectancy about the listener's (eventual) ability to understand language.

An interesting hypothesis for future research is that the exaggerated clear speech produced when adults address children actually enhances language learning. To date, no studies have been undertaken to test this hypothesis. Ultimately, future experiments will establish the amount and kind of language input required for language learning in children.

NEUROSCIENCE

The tools of modern neuroscience are being used to address issues in language processing in ever-increasing numbers. Measures being used include event-related potentials (ERP),[105,106] microelectrode recordings,[107] and magnetoencephalography (MEG),[108] as well as metabolic studies such as positron emission tomography (PET)[109] and functional magnetic resonance imaging (fMRI).[110,111] All of these methods are providing valuable information about language processing. One general finding stemming from these studies is that a much larger number of brain areas appear to be involved in language-processing than previously thought, areas that go well beyond classic Broca's and Wernicke's areas (see Dronkers, Pinker, and Damasio[112] and Doupe and Kuhl[74] for summaries). These data are consistent with the results of studies of patients' lesions and suggest that there is not one unified area for language generation, but that different cortical systems subserve different aspects of language processing and may be activated in parallel.

Brain imaging is also providing important new evidence regarding phonological processing. What can the brain tell us about the complex mapping between sound and meaning that transpires at this level? Is there evidence regarding the brain structures responsible for the failure of Japanese adults and 11-month-old infants to discern the very prominent acoustic differences between American English /r/ and /l/, or the failure of American adults and 11-month-old infants to discern the difference between Chinese /tɕʰ/ and /ɕ/. Do brain studies reveal hemispheric specialization for speech and the age at which it is first observed?

A growing number of studies not only confirm the behavioral findings on speech previously reviewed, but also extend them in interesting ways. There are four areas in which brain studies have advanced our understanding: (a) developmental perception, (b) laterality, (c) adult brain plasticity, and (c) developmentally delayed populations.

Developmental Changes in Perception: Effects of Linguistic Experience

Studies using the mismatched negativity (MMN) response, a component of event-related potentials (ERPs), have proven highly successful, even in infants as young as 6 months of age.[82,113] The MMN response is an electrophysiological brain response elicited by a change in a repetitive auditory stimulus. The MMN response in adults has been studied extensively by Näätänen and his colleagues in Finland, who have shown that the generators involve sources in both supratemporal auditory cortex and frontal cortex and that the response is generated automatically in the absence of focal attention. (See Näätänen[114] for review; see also Näätänen and Picton.[115])

In contrast to behavioral paradigms, infants and adults can be tested while they are engaged in another activity, such as reading a book or watching television or a movie. In our laboratory, babies are tested while watching a puppet show or "toy waving" by an assistant. During the tests, parents are seated in a comfortable chair in a quiet room with their infants on their laps. Infants are distracted by the puppets or toys while the electrical recordings are being made. Babies wear a soft cap with the electrodes embedded in it and seem unaware that anything unusual is going on (FIG. 11A). Continuous waveforms are recorded from 19 electrode sites, using the standard 10/20 international system, while auditory stimuli are presented from a loudspeaker.

To conduct the tests, one speech syllable, for example, /la/, serves as the standard (85% probability) and the other, for example, /ra/, as deviant (15% probability). This is very similar to the listening situation used in the head-turn task previously described, with the exception that infants are not attending in the ERP task. A block of 1100 stimuli are presented, first 500 trials of the standard and deviant, then 100 deviants alone, then 500 more trials of standard and deviant. The MMN response is obtained by subtracting the averaged waveforms for each subject under two conditions: (a) the average of the standards and (b) the average of the deviants in the context of the standard. The difference wave, calculated by subtracting the two waveforms, reveals the MMN response, a negative peak that has its onset at about 225 msec (FIG. 11B).[82,113]

Recent studies have focused on the extent to which the MMN response shows experience-related changes, the extent to which it is elicited by speech more potently

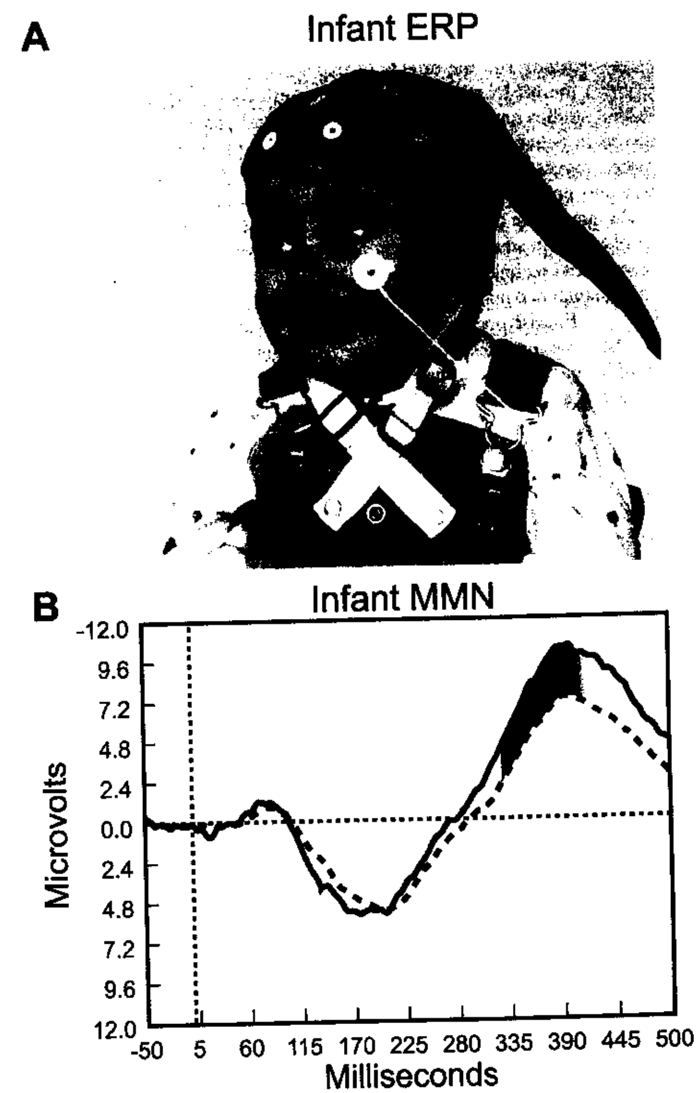


FIGURE 11. Infants' responses to speech being tested using event-related potentials. Infants wore a cap with the electrodes embedded inside (A) and listened to speech while their brain waves were recorded. The presentation of a novel syllable in a string of repeated syllables elicited the mismatched negativity, or MMN, response (*shaded area*) (B).

in the left as opposed to the right hemisphere, and the extent to which it differs in at-risk infants.

Regarding learning and the effects of experience, studies on adults and infants now verify that the MMN response reflects the effects of linguistic experience. In infants, initial studies confirm that the developmental changes in phonetic perception observed in behavioral experiments is mirrored in infant MMN measures.[116] These studies suggest that the MMN response is present in 6-month-old infants for both native and nonnative contrasts, but that by 12 months of age, the MMN response to the nonnative contrasts is no longer present. These results correspond to the results obtained in adults using native and nonnative phonetic contrasts and more sophisticated MEG techniques to measure the magnetic equivalent of the MMN response, the mismatched magnetic field (MMF). Studies conducted in Finnish-speaking adults,[68,117] English-speaking adults,[118] and Japanese-speaking adults[119] show essentially the same pattern. This indicates that infants' MMN patterns, while showing some differences when compared to those of adults, nonetheless confirm the developmental change that is so striking in behavioral tests.

Laterality for Speech

In addition to the confirmation that the transition in phonetic perception can be observed in brain measures, MMN and MMF studies are providing valuable information on hemispheric specialization. For example, the adult phonetic-perception experiments conducted on Finnish,[68,117] English,[118] and Japanese[119] people, cited previously, all demonstrated greater left- than right-hemisphere activity for native-language phonetic signals. More generally, a broad range of studies solidly support the primacy of the left hemisphere over the right in processing a variety of language stimuli in adults, although the specific areas activated by any particular kind of stimuli, for example, word processing or phonological processing, can vary substantially across studies. (See Raichle[120]; Posner and Raichle[109]; Poeppel[121]; and Demonet, Price, Wise, and Frackowiak[122] for reviews.)

Imaging studies in adults also dramatically demonstrate the lateralized processing of speech versus nonspeech signals.[123,124] For example, Zatorre and colleagues[123] examined phonetic as opposed to pitch processing using PET scans. The study employed speech signals that varied both phonetically and in their fundamental frequencies. Subjects had to judge the final consonant of the syllable in the phonetic task and the pitch (high or low) of the identical syllable in the pitch task. The results showed that phonetic processing engaged the left hemisphere, while pitch processing of the same sound engaged the right. Thus, the same stimulus can activate different brain areas depending on the dimension to which the subjects attend, providing powerful evidence of the brain's specializations for different aspects of stimuli.

While the conclusion that the left hemisphere subserves language is incontrovertible, the origin of that functional separation of the hemispheres is currently unclear. Two views have attempted to explain the functional separation. The first, proposed by Tallal and colleagues,[125,126] is that the left hemisphere specialization derives from a general tendency for the left hemisphere to engage in processing of brief, temporally complex properties of sound, of which speech is a subset.[127] Data in support of this claim derive from studies of dyslexic persons who fail on both speech and nonspeech tasks that involve rapidly changing spectral cues[98,125] and from studies of nonhuman animals that show cortical lateralization for a variety of cognitive functions, especially auditory processing of complex acoustic signals.[128] Alternatively, a second hypothesis argues that the left hemisphere specialization is due to language itself: The left hemisphere is primarily activated not by general properties of auditory stimuli, but by the linguistic significance of certain signals.[129,130] A variety of evidence supports this view, the most dramatic being from studies of deaf persons whose mode of communication involves sign language. This is a manual–spatial code that is conveyed visually, information typically thought to involve right-hemisphere analysis. A number of studies, using lesions, event-related potential methods, and PET, confirm that deaf persons process signed stimuli in the left hemisphere regions normally used for spoken language processing.[131–134] Such studies suggest that speech-related regions of the left hemisphere are well suited to language-processing independent of the modality through which it is delivered.

Given that the modality of language dominance can be specified by experience, an important question from the standpoint of development is when the left hemisphere becomes dominant in the processing of linguistic information. A variety of kinds of evidence, including lesions in children[135–139] and behavioral studies[140–143] (see Best[144] for review) as well as work using ERP methods,[67,82,145] suggests the bias toward left-hemisphere processing for language may not be present at birth, but develops rapidly in infancy. In a recent study, Dahaene-Lambertz[146] demonstrated in tests on 4-month-olds that while speech and nonspeech demonstrated distinct topographies, both signals produced greater MMN activation in the left hemisphere. Experience with linguistically patterned information may therefore be required to produce a differentiation of the hemispheres. (See Doupe and Kuhl[74] for a comparison of laterality for song in songbirds and speech in humans.) 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 nature of the signals, rather than their specific modality, that accounts for the eventual specialization. Finally, regardless of how early a bias for lateralization appears, it may be susceptible to deprivation of auditory input early in life.[134,147]

Brain Plasticity and Language: Is There a Critical Period?

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 is often 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 (as well as those on birds acquiring bird song; see Doupe and Kuhl[74]) suggest an alternative.[39,82] 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 Kuhl's NLM model, it likely "commits" neural structure in some way.[63,79] Neural commitment to a learned structure may interfere with the processing of information that does not conform to the learned pattern. In this view, 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 in production, as shown, for example, by the difficulty of the /r-l/ distinction for native speakers of Japanese, even after training.[11,78,148,149] 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 conducted by our laboratory group in collaboration with the MEG researchers at Nippon Telephone and Telegraph in Tokyo suggest that training which capitalizes on a natural strategy, "motherese," may provide the best method of teaching a second language. These studies have obvious practical value and also provide some evidence supporting a "neural commitment" view.

Support for this view derives from experiments on the effects of linguistic experience using MEG. We examined 10 Japanese subjects' perception of English /ra/ versus /la/ in comparison to that of 10 American subjects.[150] Behavioral measures included identification and discrimination of the speech syllables. While subjects listened, MEG recordings were made. Results from two representative individuals are shown in FIGURE 12. The MEG equivalent to the mismatch negativity, the MMF responses, was significantly larger for the American subjects (FIG. 12A). Moreover, the results demonstrated a left-hemisphere dominance for American subjects, indicating that significant linguistic processing was involved. In contrast, Japanese listeners showed a bilateral mismatch negativity pattern that was dominated by the right hemisphere (FIG. 12B). We interpret the study to mean that Japanese listeners were analyzing the patterns in terms of their auditory properties, whereas American listeners were responding to the signals as language.

In a follow-up training study, we asked whether Japanese listeners could be trained to respond to the the /r/ and /l/ stimuli as linguistic signals.[151] Taking our cue from the "motherese" studies, Japanese listeners heard acoustically modified /ra/ and /la/ syllables, stimuli with greatly exaggerated formant frequencies, reduced bandwidths, and extended durations, like those produced by mothers. Listeners also heard many different talkers, and the syllables were presented in many different vowel contexts. No explicit feedback was presented; the listeners presented the stimuli to themselves via computer, and their progress was measured every 50 presentations. Listeners progressed through the training program in 12 sessions. The behavioral data revealed a significant improvement in identification of these non-native speech stimuli. Correspondingly, the MEG results showed enhanced mismatch field responses in the left hemisphere and reduced activities in the right hemisphere.

These initial results suggest that the parameters of motherese may provide an excellent training method for second-language learning. The three parameters of greatest interest are the following: (a) The dimensions critical to the prosodic and segmental aspects of the language are exaggerated; (b) listeners are provided numerous instances of the critical information in an unsupervised learning situation; and (c) many talkers provide that information, giving an opportunity for the full variability of speech to be experienced, which may in turn enhance the ability to mentally represent categories of information providing listeners with multiple instances spoken by many talkers. In studies of second-language learning, these features have been shown to be more effective training methods. (See Kuhl[36] for review.) Taken together, the studies show that feedback and reinforcement are not necessary in this process; listeners simply need the right kind of listening experience.[152,153] Exagger-

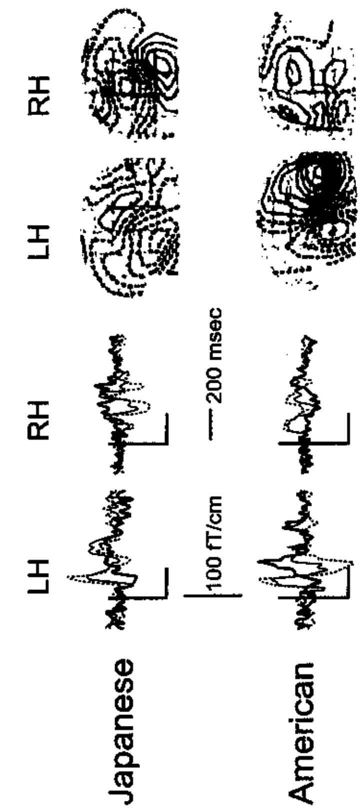


FIGURE 12. Magnetoencephalographic recordings of the MMF in American and Japanese listeners show the mismatched negativity response in the left hemisphere in response to American English /ra/ and /la/ in the American listener, but not in the Japanese listener. (A) Responses to the standard stimulus (*light dashed line*) and the deviant stimulus (*light solid line*) were subtracted to produce a difference wave (*bold line*) that peaks at about 200 msec, revealing the MMF. (B) Contour maps indicate that the sources of greatest activity are in the left hemisphere for the American listener and in the right hemisphere for the Japanese listener.

ated acoustic cues, multiple instances by many talkers, and mass listening experience—features that motherese provides infants—may represent a natural way to learn a language.

Early in life, infants have no trouble acquiring more than one language. This may be due to the fact that interference effects are minimal. We presume they acquire two separate "maps," one for each language. 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 using fMRI techniques indicate that adult bilinguals who acquire both languages early in life activate overlapping regions of the brain when processing the two languages, while those who learn the second language later in life activate two distinct regions of the brain for the two languages.[108] 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.

Developmentally Delayed Populations

Increasing evidence exists, from both behavioral and brain studies, that developmentally delayed populations, including children with autism,[154] dyslexia,[155–159] and specific language impairment, have a great deal of difficulty processing the sounds of human speech.[160,161] This initial inability to process the building blocks of language may greatly restrict infants' abilities to acquire higher levels of language. Infants' early strategies to map language are seen as providing a "bootstrap" approach to the higher levels of language.[162] Infants who cannot accomplish this initial step in language-processing may be further impaired as they attempt to process higher-order language units. Interestingly, studies in which children between 8 and 12 years of age with confirmed language disorders associated with dyslexia were treated with speech that greatly exaggerated the acoustic features of speech in a mass-practice computer game over many sessions demonstrated impressive improvements across a broad array of language measures.[97,98]

These results suggest a fundamental principle: Early measures of speech discrimination of the building blocks of speech, the consonants and vowels that make up words, should strongly correlate with later language measures. Recent studies by our laboratory team confirm this hypothesis.[93,163] In this longitudinal study, early speech discrimination data were obtained from American infants tested with nonnative vowels (Finnish /u/ vs. /y/) at 6 months of age using the head-turning technique. Nonnative vowels were used to reduce the effects of experience on infants' initial head-turn performance. At three successive ages, measurements of communication and language development were gathered using the MacArthur Communicative Development Inventory (CDI). These measures were assessed when infants were 13, 16, and 24 months old. The results demonstrated strong correlations between speech discrimination in infancy and later language development. The strong correlations were shown consistently across different ages in the second year of life (for 13, 16, and 24 months, $r = 0.70$, 0.470, and 0.656, respectively). This longitudinal study, exploring

the relation between infants' capacities of speech discrimination and early communicative development, provides the first evidence that the development of speech perception may serve as an indicator of later communicative development.[93,163] Perception of the building blocks of language, the consonants and vowels that make up words, may thus be the earliest measure of infants' eventual success at learning to communicate through language. Such measures may lead to early intervention programs among children at risk for normal language development.

COMPUTER SCIENCE AND INFORMATICS

Computer simulations of intelligent behavior, from chess playing to speech understanding, have been a goal of work in artificial intelligence for many years. The goal of such simulations was not to mimic the methods humans use to achieve the goal. Rather it was to succeed at the task. Computers could play chess differently than humans, and computers attempting to understand speech (a problem thus far unsolved) could employ different methods than humans, as long as the machine mastered the task.

This framework is changing, however, as computer scientists and informatics specialists are using computational approaches to try to mimic natural biological learning. Computers using biological approaches are now attempting to mimic a variety of feats shown by infants in the acquisition of language. Artificial neural networks are being used to model psychological results. Successful examples include a self-organizing neural network, modeled after a Kohonen[164] network, to replicate the perceptual magnet effect seen in speech,[165] a sensory-motor speech articulation device named "DIVA" that incorporates a direct link between speech perception and production and learns by "babbling,"[166,167] a temporal recurrent network (TRN) that replicates infants' pickup of the serial and temporal information in speech as shown in statistical learning,[168] and an abstract recurrent network (ARN) that mimics infants' abilities to perceive abstract structure in natural and artificial languages.[169,170]

An interesting aspect of the self-organizing systems designed by Guenther and the temporal and abstract recurrent networks designed by Dominey is that they are biologically inspired. Guenther's approach to the modeling used to generate the perceptual magnet effect is fashioned after map formation in auditory cortex, and his DIVA model is motivated by known physiological links between auditory and motor cortex in cortical language areas. Dominey's approach uses neuropsychological evidence from studies of patients with Parkinson's disease.

In my laboratory, we are approaching infant learning of the vowel systems of their languages using a neural network.[171,172] In this work, we examine the hypothesis that children can acquire vowel systems based purely on the input they receive, as long as the input mirrors what infants actually hear (i.e., motherese). The computer model is based on the Kohonen neural network, which mimics certain aspects of biological neural systems. The network learns a so-called equiprobable distribution of its input neurons. This means that each neuron has an equal probability of being activated by an input, resulting in more neurons clustering around more frequent inputs. The classification behavior of the network can then be tested using the population vector, a measure of the average activation of the neurons in the network.

The population vector for a given input tends to be biased in the direction of high densities of neurons, effectively implementing classification behavior.

The unique feature of this artificial learning system is that it will compare learning with motherese speech as opposed to adult-directed speech. We hypothesize that learning a vowel system from adult-directed utterances will prove much more difficult due to the well-known fact that adult speakers reduce their articulations in casual, rapid speech. As reviewed previously, adult speakers do not show the stretched vowel space characteristic of the exaggerated utterances used in motherese.[92] The motherese and adult-directed speech samples are those recorded during natural conversations in the previous study.[92] These samples came from women in three countries—Sweden, Russia, and the United States—and were recorded while the women addressed their young infants and another adult. The Russian, English, and Swedish speech samples will be used as input to the machine. These languages represent an interesting sample from the range of vowel systems that are used in human languages: Russian has five vowels, English has nine, and Swedish approximately sixteen. The experiments will provide an answer to whether or not an artificial learning system, without *a priori* knowledge, can derive the vowel system of a natural language by exposure to utterances designed for language learners. Will the Russian-trained network derive the vowels of Russian as opposed to Swedish, and vice versa? Moreover, the experiments will test the extent to which motherese facilitates the learning process.

Biologically inspired computer simulations are new and represent an interesting influence of developmental cognitive science on computational modeling. Only time will tell whether the answers machines provide will advance our understanding of how children learn, but the initial results look promising.

CONCLUSIONS

Cross-disciplinary work on language is changing our views of the acquisition process. Once considered an activity that began when infants learned their first words at about one year of age, language learning is now a phenomenon that begins as soon as infants hear language. Experiments on infants, along with data from cultural anthropology, neuroscience, and computer science and informatics, have revised our view of how language is acquired. The laboratory studies demonstrate that during the first year of life, infants employ strategies to "map" language that are surprising and unpredicted by historical views. By simply listening to ambient language, infants acquire information about the phonetic units employed by their language and the rules for combining sounds into words. They discover likely word candidates by statistically analyzing the serial and temporal aspects of language input, and they are attuned to surface structure components that categorize words by grammatical class. Infants accomplish this before understanding or producing a single word, and before conceiving of the fact that objects and events in the world are named. Much of infant learning has been shown to be computational. The learning that ensues in the early period before speech alters infants' perceptual systems, and this enhances the processing of a specific language. Data, originally derived from cultural anthropology studies, reveal that the unconscious speaking style that we use when addressing infants and children, "motherese," not only appears to be preferred by infants, but also is argued to be beneficial to learning. Brain-imaging studies demonstrate that language-specific processing activates the left hemisphere and that training that capitalizes on motherese is effective in teaching foreign-language processing. Biologically inspired learning networks are now being tested with natural language input to determine whether machines can derive the categories of language as infants do. Whether these divergent results will produce consilience cannot yet be known, but preliminary findings hold great promise.

ACKNOWLEDGMENTS

The research reported here was supported by grants from the National Institutes of Health and the Human Frontiers Science Program to PKK, by funding from the National Kaohsiung Normal University in Kaohsiung, Taiwan to F-MT and H-ML, by a Nippon Telephone and Telegraph research traineeship to YZ, and by research support for the Center for Mind, Brain, and Learning from the Talaris Research Institute of Seattle, Washington.

REFERENCES

1. Skinner, B.F. 1957. Verbal Behavior. Appleton Century Crofts. New York.
2. Chomsky, N. 1959. A review of B.F. Skinner's Verbal Behavior. Language **35:** 26–58.
3. Wexler, K. & P.W. Culicover. 1980. Formal Principles of Language Acquisition. MIT Press. Cambridge, MA.
4. Peterson, G.E. & H.E. Barney. 1952. Control methods used in a study of vowels. J. Acoust. Soc. Am. **24:** 175–184.
5. Liberman, A.M., F.S. Cooper, D.S. Shankweiler & M. Studdert-Kennedy. 1967. Perception of the speech code. Psychol. Rev. **74:** 431–461.
6. Miller, J.L. & A.M. Liberman. 1979. Some effects of later-occurring information on the perception of stop consonant and semivowel. Percept. Psychophys. **25:** 457–465.
7. Cutler, A., D. Dahan & W. van Donselaar. 1997. Prosody in the comprehension of spoken language: a literature review. Lang. Speech **40(2):** 141–201.
8. Wingfield, A., L. Lombardi & S. Sokol. 1984. Prosodic features and the intelligibility of accelerated speech: syntactic versus periodic segmentation. J. Speech Hear. Res. **27:** 128–134.
9. Chao, Y.R. 1948. Mandarin Primer. Harvard University Press. Cambridge, MA.
10. Jusczyk, P.W. 1997. The Discovery of Spoken Language. MIT Press. Cambridge, MA.
11. Miyawaki, K., W. Strange, R. Verbrugge, *et al.* 1975. An effect of linguistic experience: the discrimination of [r] and [l] by native speakers of Japanese and English. Percept. Psychophys. **18(5):** 331–340.
12. Eimas, P.D., E.R. Siqueland, P. Juscayk & J. Vigorito. 1971. Speech perception in infants. Science **171:** 303–306.
13. Eimas, P.D. 1974. Auditory and linguistic processing of cues for place of articulation by infants. Percept. Psychophys. **16(3):** 513–521.
14. Eimas, P.D. 1975. Auditory and phonetic coding of the cues for speech: discrimination of the /r-l/ distinction by young infants. Percept. Psychophys. **18(5):** 341–347.
15. Lasky, R.E., A. Syrdal-Lasky & R.E. Klein. 1975. VOT discrimination by four to six and a half month old infants from Spanish environments. J. Exp. Child Psychol. **20:** 215–225.
16. Streeter, L.A. 1976. Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. Nature (London) **259:** 39–41.

17. EIMAS, P.D. 1975. Speech perception in early infancy. *In* Infant Perception: Vol. 2. From Sensation to Cognition. L.B. Cohen & P. Salapatek, Eds.: 193–231. Academic Press. New York.
18. KUHL, P.K. & J.D. MILLER. 1975. Speech perception by the chinchilla: voiced-voiceless distinction in alveolar plosive consonants. Science **190:** 69–72.
19. KUHL, P.K. & J.D. MILLER. 1978. Speech perception by the chinchilla: identification functions for synthetic VOT stimuli. J. Acoust. Soc. Am. **63(3):** 905–917.
20. KUHL, P.K. 1981. Discrimination of speech by nonhuman animals: basic auditory sensitivities conducive to the perception of speech-sound categories. J. Acoust. Soc. Am. **70(2):** 340–349.
21. KUHL, P.K. & D.M. PADDEN. 1982. Enhanced discriminability at the phonetic boundaries for the voicing feature in macaques. Percept. Psychophys. **32(6):** 542–550.
22. KUHL, P.K. & D.M. PADDEN. 1983. Enhanced discriminability at the phonetic boundaries for the place feature in macaques. J. Acoust. Soc. Am. **73(3):** 1003–1010.
23. DOOLING, R.J., C.T. BEST & S.D. BROWN. 1995. Discrimination of synthetic full-formant and sinewave /ra-la/ continua by budgerigars (*Melopsittacus undulatus*) and zebra finches (*Taeniopygia guttata*). J. Acoust. Soc. Am. **97:** 1839–1846.
24. KLUENDER, K.R., R.L. DIEH & P.R. KILLEEN. 1987. Japanese quail can learn phonetic categories. Science **237:** 1195–1197.
25. RAMUS, F., M.D. HAUSER, C. MILLER, *et al.* 2000. Language discrimination by human newborns and by cotton-top tamarin monkeys. Science **288(5464):** 349–351.
26. KUHL, P.K. 1991. Perception, cognition, and the ontogenetic and phylogenetic emergence of human speech. *In* Plasticity of Development. S.E. Brauth, W.S. Hall & R.J. Dooling, Eds.: 73–106. MIT Press. Cambridge, MA.
27. KUHL, P.K. 1988. Auditory perception and the evolution of speech. Hum. Evol. **3:** 19–43.
28. MILLER, J.D., C.C. WIER, R.E. PASTORE, *et al.* 1976. Discrimination and labeling of noise-buzz sequences with varying noise-lead times: an example of categorical perception. J. Acoust. Soc. Am. **60(2):** 410–417.
29. PISONI, D.B. 1977. Identification and discrimination of the relative onset time of two component tones: implications for voicing perception in stops. J. Acoust. Soc. Am. **61(5):** 1352–1361.
30. JUSCZYK, P.W., B.S. ROSNER, J.E. CUTTING, *et al.* 1977. Categorical perception of nonspeech sounds by 2-month-old infants. Percept. Psychophys. **21(1):** 50–54.
31. LIBERMAN, A.M. & I.G. MATTINGLY. 1985. The motor theory of speech perception revised. Cognition **21:** 1–36.
32. WERKER, J.F. & R.C. TEES. 1984. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. Infant Behav. Dev. **7:** 49–63.
33. WERKER, J.F. 1995. Exploring developmental changes in cross-language speech perception. *In* An Invitation to Cognitive Science: Language. L.R. Gleitman & A.M. Liberman, Eds.: 87–107. MIT Press. Cambridge, MA.
34. KUHL, P.K. 1994. Learning and representation in speech and language. Curr. Opin. Neurobiol. **4:** 812–822.
35. KUHL, P.K. 2000. A new view of language acquisition. Proc. Natl. Acad. Sci. USA **97(22):** 11850–11857.
36. KUHL, P.K. 2000. Language, mind, and brain: experience alters perception. *In* The New Cognitive Neurosciences. 2nd ed. M.S. Gazzaniga, Ed.: 99–115. MIT Press. Cambridge, MA.
37. KUHL, P.K., T. DEGUCHI, A. HAYASHI, *et al.* 1997. Effects of language experience on speech perception: American and Japanese infants' perception of /ra/ and /la/. Paper presented at 132nd Meeting of the Acoustical Society of America, San Diego.
38. TSAO, F-M., H-M. LIU, P.K. KUHL & C-H. TSENG. 2000. Perceptual discrimination of a Mandarin fricative-affricate contrast by English-learning and Mandarin-learning infants. Paper presented at 12th Biennial International Conference on Infant Studies, Brighton, UK.
39. SAFFRAN, J.R., R.N. ASLIN & E.L. NEWPORT. 1996. Statistical learning by 8-month-old infants. Science **274:** 1926–1928.

40. GROSS, N., P.C. JUDGE, O. PORT & S.H. WILDSTROM. 1998. Let's talk! Speech technology is the next big thing in computing. Will it put a PC in every home? Bus. Week 60–72.
41. KUHL, P.K. 1985. Categorization of speech by infants. *In* Neonate Cognition: Beyond the Blooming Buzzing Confusion. J. Mehler & R. Fox, Eds.: 231–262. Erlbaum. Hillsdale, NJ.
42. KUHL, P.K. 1979. Speech perception in early infancy: perceptual constancy for spectrally dissimilar vowel categories. J. Acoust. Soc. Am. **66(6):** 1668–1679.
43. KUHL, P.K. 1983. Perception of auditory equivalence classes for speech in early infancy. Infant Behav. Dev. **6:** 263–285.
44. HILLENBRAND, J. 1983. Perceptual organization of speech sounds by infants. J. Speech Hear. Res. **26:** 268–282.
45. KUHL, P.K. 1980. Perceptual constancy for speech–sound categories in early infancy. *In* Child Phonology, Vol. 2: Perception. G.H. Yeni-Komshian, J.F. Kavanagh & C.A. Ferguson, Eds.: 41–66. Academic Press. New York.
46. JUSCZYK, P.W. 1999. How infants begin to extract words from speech. Trends Cogn. Sci. **3:** 323–328.
47. MEHLER, J., P.W. JUSCZYK, G. LAMBERTZ, *et al.* 1988. A precursor of language acquisition in young infants. Cognition **29:** 143–178.
48. MOON, C., R.P. COOPER & W.P. FIFER. 1993. Two-day-olds prefer their native language. Infant Behav. Dev. **16:** 495–500.
49. NAZZI, T., J. BERTONCINI & J. MEHLER. 1998. Language discrimination by newborns: toward an understanding of the role of rhythm. J. Exp. Psychol. Hum. Percept. Perform. **24(3):** 756–766.
50. LECANUET, J.P. & C. GRANIER-DEFERRE. 1993. Speech stimuli in the fetal environment. *In* Developmental Neurocognition: Speech and Face Processing in the First Year of Life. B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, *et al.*, Eds. The Netherlands. Kluwer, Dordrecht.
51. DECASPER, A.J. & W.P. FIFER. 1980. Of human bonding: newborns prefer their mothers' voices. Science **208(6):** 1174–1176.
52. DECASPER, A.J. & M.J. SPENCE. 1986. Prenatal maternal speech influences newborns' perception of speech sounds. Infant Behav. Dev. **9:** 133–150.
53. JUSCZYK, P.W., A. CUTLER & N.J. REDANZ. 1993. Infants' preference for the predominant stress patterns of English words. Child Dev. **64:** 675–687.
54. HIRSH-PASEK, K., D.G. KEMLER-NELSON, P.W. JUSCZYK, *et al.* 1987. Clauses are perceptual units for young infants. Cognition **26(3):** 269–286.
55. JUSCZYK, P.W., A.D. FRIEDERICI, J.M. WESSELS, *et al.* 1993. Infants' sensitivity to the sound patterns of native language words. J. Memory Lang. **32:** 402–420.
56. GOODSITT, J.V., J.L. MORGAN & P.K. KUHL. 1993. Perceptual strategies in prelingual speech segmentation. J. Child Lang. **20(2):** 229–252.
57. WOLFF, J.G. 1977. The discovery of segments in natural language. Br. J. Psychol. **68:** 97–106.
58. ASLIN, R.N., J.R. SAFFRAN & E.L. NEWPORT. 1998. Computation of conditional probability statistics by 8-month-old infants. Psychol. Sci. **9:** 321–324.
59. MATTYS, S.L., P.W. JUSCZYK, P.A. LUCE & J.L. MORGAN. 1999. Phonotactic and prosodic effects on word segmentation in infants. Cogn. Psychol. **38(4):** 465–494.
60. BOSCH, L., A. COSTA & N. SEBASTIAN-GALLES. 2000. First and second language vowel perception in early bilinguals. Eur. J. Cogn. Psychol. **12(2):** 189–221.
61. GRIESER, D. & P.K. KUHL. 1989. Categorization of speech by infants: support for speech–sound prototypes. Dev. Psychol. **25:** 577–588.
62. KUHL, P.K. 1991. Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. Percept. Psychophys. **50(2):** 93–107.
63. KUHL, P.K., K.A. WILLIAMS, F. LACERDA, *et al.* 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. Science **255:** 606–608.
64. IVERSON, P. & P.K. KUHL. 1995. Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. J. Acoust. Soc. Am. **97:** 553–562.

65. MILLER, J.L. 1994. On the internal structure of phonetic categories: a progress report. Cognition **50:** 271–285.
66. AALTONEN, O., O. EEROLA, A. HELLSTRÖM, *et al.* 1997. Perceptual magnet effect in the light of behavioral and psychophysiological data. J. Acoust. Soc. Am. **101:** 1090–1105.
67. CHEOUR-LUHTANEN, M., K. ALHO, T. KUJALA, *et al.* 1995. Mismatch negativity indicates vowel discrimination in newborns. Hear. Res. **82:** 53–58.
68. NÄÄTÄNEN, R., A. LEHTOKOSKI, M. LENNES, *et al.* 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature (London) **385:** 432–434.
69. SHARMA, A. & M.F. DORMAN. 1998. Exploration of the perceptual magnet effect using the mismatch negativity auditory evoked potential. J. Acoust. Soc. Am. **104:** 511–517.
70. ROSCH, E. 1975. Cognitive reference points. Cogn. Psychol. **7:** 532–547.
71. GARNER, W.R. 1974. The Processing of Information and Structure. Erlbaum. Potomac, MD.
72. GOLDMAN, D. & D. HOMA. 1977. Integrative and metric properties of abstracted information as a function of category discriminability, instance variability, and experience. J. Exp. Psychol. Hum. Learn. Memory **3:** 375–385.
73. MERVIS, C.B. & E. ROSCH. 1981. Categorization of natural objects. Annu. Rev. Psychol. **32:** 89–115.
74. DOUPE, A.J. & P.K. KUHL. 1999. Birdsong and human speech: common themes and mechanisms. Annu. Rev. Neurosci. **22:** 567–631.
75. KUHL, P.K. & A.N. MELTZOFF. 1996. Infant vocalizations in response to speech: vocal imitation and developmental change. J. Acoust. Soc. Am. **100:** 2425–686.
76. IVERSON, P. & P.K. KUHL. 2000. Perceptual magnet and phoneme boundary effects in speech perception: do they arise from a common mechanism? Percept. Psychophys. **62(4):** 874–886.
77. STRANGE, W. & S. DITTMANN. 1984. Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. Percept. Psychophys. **36:** 131–145.
78. GOTO, H. 1971. Auditory perception by normal Japanese adults of the sounds "l" and "r." Neuropsychologia **9:** 317–323.
79. IVERSON, P. & P.K. KUHL. 1996. Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. J. Acoust. Soc. Am. 9 1130–1140.
80. POLKA, L. & O.S. BOHN. 1996. A cross-language comparison of vowel perception English-learning and German-learning infants. J. Acoust. Soc. Am. **100(1):** 577–92.
81. MILLER, J.L. & P.D. EIMAS. 1996. Internal structure of voicing categories in ea infancy. Percept. Psychophys. **58:** 1157–1167.
82. KUHL, P.K. 1998. The development of speech and language. *In* Mechanistic Rel tionships between Development and Learning. T.J. Carew, R. Menzel & C.J. Sh Eds.: 53–73. Wiley. New York.
83. YOUNGER, B.A. & L.B. COHEN. 1985. How infants form categories. *In* The Psych ogy of Learning and Motivation. G.H. Bower, Ed.: 211–247. Academic Press, S Diego.
84. SAFFRAN, J.R., E.K. JOHNSON, R.N. ASLIN & E.L. NEWPORT. 1999. Statistical lea of tone sequences by human infants and adults. Cognition **70(1):** 27–52.
85. QUINN, P.C. & P.D. EIMAS. 1998. Evidence for a global categorical representati humans by young infants. J. Exp. Child Psychol. **69:** 151–174.
86. GALLISTEL, C.R. 1990. The Organization of Learning. MIT Press. Cambridge,
87. FERGUSON, C.A. 1964. Baby talk in six languages. Am. Anthropol. **66:** 103–11
88. BLOUNT, B.G. 1990. Parental speech and language acquisition: an anthropolo perspective. Pre- Peri-Natal Psychol. J. **4(4):** 319–337.
89. FERNALD, A. & T. SIMON. 1984. Expanded intonation contours in mothers' sp newborns. Dev. Psychol. **20(1):** 104–113.
90. GRIESER, D. & P.K. KUHL. 1988. Maternal speech to infants in a tonal langu port for university prosodic features in motherese. Dev. Psychol. **24(1):** 14

91. JACOBSON, J.L., D.C. BOERSMA, R.B. FIELDS & K.L. OLSON. 1983. Paralinguistic features of adult speech to infants and small children. Child Dev. **54:** 436–442.
92. KUHL, P.K., J.E. ANDRUSKI, I.A. CHISTOVICH, *et al.* 1997. Cross-language analysis of phonetic units in language addressed to infants. Science **277:** 684–686.
93. LIU, H-M., F-M. TSAO, E.B. STEVENS & P.K. KUHL. 2000. Maternal speech to infants in Mandarin Chinese: support for an expanded vowel triangle in Mandarin Motherese. Paper presented at 27th International Congress of Psychology, Stockholm, Sweden.
94. TURNER, G.S., K. TJADEN & G. WEISMER. 1995. The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. J. Speech Hear. Res. **38:** 1001–1013.
95. FERNALD, A. 1985. Four-month-old infants prefer to listen to motherese. Infant Behav. Dev. **8:** 181–195.
96. LIVELY, S.E., J.S. LOGAN & D.B. PISONI. 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.
97. MERZENICH, M.M., W.M. JENKINS, P. JOHNSTON, *et al.* 1996. Temporal processing deficits of language-learning impaired children ameliorated by training. Science **271(5245):** 77–81.
98. TALLAL, P., S.L. MILLER, G. BEDI, *et al.* 1996. Language comprehension in language-learning impaired children improved with acoustically modified speech. Science **271(5245):** 81–84.
99. LINDBLOM, B. 1991. The status of phonetic gestures. *In* Modularity and the Motor Theory of Speech Perception. I.G. Mattingly & M. Studdert-Kennedy, Eds.: 7–24. Lawrence Erlbaum Associates. Hillsdale, NJ.
100. LANE, H.L. & B. TRANEL. 1971. The Lombard sign and the role of hearing in speech. J. Speech Hear. Res. **14:** 677–709.
101. WATSON-GEGEO, K.A. & D.W. GEGEO. 1986. Calling-out and repeating routines in Kwara'ae children's language socialization. *In* Language Socialization across Cultures. B. Schieffelin & E. Ochs, Eds.: 17–50. Cambridge University Press. Cambridge, UK.
102. SCHEIFFELIN, B.B. & E. OCHS. 1983. A cultural perspective on the transition from prelinguistic to linguistic communication. *In* The Transition from Prelinguistic to Linguistic Communication. R.M. Golinkoff, Ed.: 115–131. Lawrence Erlbaum Associates. Hillsdale, NJ.
103. RATNER, N.B. & C. PYE. 1984. Higher pitch is not universal: acoustic evidence from Quiche Mayan. J. Child Lang. **11:** 515–522.
104. BURNHAM, D.K., U. VOLLMER-CONNA & C. KITAMURA. 2000. Talking to infants, pets, and adults: what's the difference? Paper presented at 12th Biennial International Conference on Infant Studies, Brighton, UK.
105. NEVILLE, H.J., J. NICOL, A. BARSS, *et al.* 1991. Syntactically based sentence processing classes: evidence from event-related brain potentials. J. Cogn. Neurosci. **3:** 155–170.
106. OSTERHOUT, L. & P. HOLCOMB. 1992. Event-related brain potentials elicited by syntactic anomaly. J. Memory Lang. **31(6):** 785–806.
107. OJEMANN, G.A. & M.J. SCHOENFIELD. 1999. Activity of neurons in human temporal cortex during identification and memory for names and words. J. Neurosci. **19(13):** 5674–5682.
108. HÄMALÄINEN, M.S., R. HARI, R.J. ILMONIEMI, *et al.* 1993. Magnetoencephalography: theory, instrumentation, and applications to noninvasive studies of the working human brain. Rev. Mod. Phys. **65:** 413–497.
109. POSNER, M.I. & M.E. RAICHLE. 1995. Precise of images of mind. Behav. Brain Sci. **18(2):** 327–383.
110. BINDER, J.R. 1997. Neuroanatomy of language processing studied with functional MRI. Clin. Neurosci. **4:** 87–94.
111. KIM, K.H.S., N.R. RELKIN, K.M. LEE & J. HIRSCH. 1997. Distinct cortical areas associated with native and second languages. Nature **388(6638):** 171–174.

112. DRONKER, N.F., S. PINKER & A. DAMASIO. 2000. Language and the aphasia. *In* Principles of Neural Science. 4th ed. E.R. Kandel, J.H. Schwartz & T.M. Jessell, Eds.: 1169–1187. McGraw-Hill. New York.
113. PANG, E., G.E. EDMONDS, R. DESJARDINS, *et al.* 1998. Mismatch negativity to speech stimuli in 8-month-old infants and adults. Int. J. Psychophysiol. **29(2):** 227–236.
114. NÄÄTÄNEN, R. 1990. The role of attention in auditory information processing as revealed by event-related potentials and other brain measure of cognitive function. Behav. Brain Sci. **13:** 201–288.
115. NÄÄTÄNEN, R. & T.W. PICTON. 1987. The N1 wave of human electric and magnetic response to sound: a review and analysis of the component structure. Psychophysiology **24:** 375–425.
116. CHEOUR, M., K. ALHO, R. CEPONIENE, *et al.* 1998. Maturation of mismatch negativity in infants. Int. J. Psychophysiol. **29(2):** 217–226.
117. RINNE, T., K. ALHO, P. ALKU, *et al.* 1999. Analysis of speech sounds is left-hemisphere predominant at 100–150 ms after sound onset. Neurorep. Rapid Commun. Neurosci. Res. **10(5):** 1113–1117.
118. ALHO, K., J.F. CONNOLLY, M. CHEOUR, *et al.* 1998. Hemispheric lateralization in preattentive processing of speech sounds. Neurosci. Lett. **258(1):** 9–12.
119. KOYAMA, S., A. GUNJI, H. YABE, *et al.* 2000. Hemispheric lateralization in an analysis of speech sounds: left hemisphere dominance replicated in Japanese subjects. Cogn. Brain Res. **10:** 119–124.
120. RAICHLE, M.E. 1992. Mind and brain. Sci. Am. **267:** 48–57.
121. POEPPEL, D. 1996. A critical review of PET studies of phonological processing. Brain Lang. **55(3):** 317–351.
122. DEMONET, J.F., C. PRICE, R. WISE & R.S. FRACKOWIAK. 1994. Differential activation of right and left posterior sylvian regions by semantic and phonological tasks: a positron-emission tomography study in normal human subjects. Neurosci. Lett. **182(1):** 25–28.
123. ZATORRE, R.J., A.C. EVANS, E. MEYER & A. GJEDDE. 1992. Lateralization of phonetic and pitch discrimination in speech processing. Science **256(5058):** 846–849.
124. POEPPEL, D., C. PHILLIPS, E. YELLIN, *et al.* 1997. Processing of vowels in supratemporal auditory cortex. Neurosci. Lett. **221(2-3):** 145–148.
125. SCHWARTZ, J. & P. TALLAL. 1980. Rate of acoustic change may underlie hemispheric specialization for speech perception. Science **207:** 1380–1381.
126. TALLAL, P., S. MILLER & R.H. FITCH. 1993. Neurobiological basis of speech: a case for the preeminence of temporal processing. *In* Temporal Information Processing in the Nervous System: Special Reference to Dyslexia and Dysphasia. P. Tallal, A. M. Galaburda, R.R. Llinas & C.V. Euler, Eds. Ann. N.Y. Acad. Sci. **682:** 27–47.
127. FITCH, R.H., S. MILLER & P. TALLAL. 1997. Neurobiology of speech perception. Annu. Rev. Neurosci. **20:** 331–353.
128. HEFFNER, H.E. & R.S. HEFFNER. 1984. Temporal lobe lesions and perception of species-specific vocalizations by macaques. Science **226(4670):** 75–76.
129. GATHERCOLE, S.E. & A.D. BADDELEY. 1990. Phonological memory deficits in language disordered children: is there a causal connection? J. Memory Lang. **29(3):** 336–360.
130. MODY, M., M. STUDDERT-KENNEDY & S. BRADY. 1997. Speech perception deficits in poor readers: auditory processing or phonological coding? J. Exp. Child Psychol. **64(2):** 199–231.
131. BELLUGI, U., H. POIZNER & E.S. KLIMA. 1983. Brain organization for language: clues from sign aphasia. Hum. Neurobiol. **2:** 155–170.
132. KLIMA, E.S., U. BELLUGI & H. POIZNER. 1988. Grammar and space in sign aphasiology. Aphasiology **2:** 319–327.
133. NEVILLE, H.J., D.L. MILLS & D.S. LAWSON. 1992. Fractionating language: different neural subsystems with different sensitive periods. Cerebr. Cortex **2:** 244–258.
134. NEVILLE, H.J., D. BAVELIER, D. CORINA, *et al.* 1998. Cerebral organization for language in deaf and hearing subjects: biological constraints and effects of experience. Proc. Natl. Acad. Sci. USA **95(3):** 922–929.

135. DENNIS, M. & B. KOHN. 1975. Comprehension of syntax in infantile hemiplegics after cerebral hemidecortication: left-hemisphere superiority. Brain Lang. **2(4):** 472–482.
136. DENNIS, M. & H.A. WHITAKER. 1976. Language acquisition following hemidecortication: linguistic superiority of the left over the right hemisphere. Brain Lang. **3:** 404–433.
137. KOHN, B. & M. DENNIS. 1974. Selective impairments of visuo-spatial abilities in infantile hemiplegics after right cerebral hemidecortication. Neuropsychologia **12(4):** 505–512.
138. TALLAL, P., R.L. SAINBURG & T. JERNIGAN. 1991. The neuropathology of developmental dysphasia: behavioral, morphological, and physiological evidence for a pervasive temporal processing disorder. Read. Writ. **3(3-4):** 363–377.
139. WITELSON, S.F. 1987. Neurobiological aspects of language in children. Child Dev. **58:** 653–688.
140. BEST, C.T., H. HOFFMAN & B.B. GLANVILLE. 1982. Development of infant ear asymmetries for speech and music. Percept. Psychophys. **31:** 75–85.
141. BEVER, T.G. 1971. The nature of cerebral dominance in speech behavior of the child and adult. *In* Language Acquisition: Models and Methods. R. Huxley & R. Ingham, Eds. Academic Press. London.
142. ENTUS, A.K. 1977. Hemispheric asymmetry in processing of dichotically presented speech and nonspeech stimuli by infants. *In* Language Development and Neurological Theory. S.J. Segalowitz. & F.A. Gruber, Eds. Academic Press. New York.
143. VARGHA-KHADEM, F. & M.C. CORBALLIS. 1979. Cerebral asymmetry in infants. Brain Lang. **8:** 1–9.
144. BEST, C.T. 1988. The emergence of cerebral asymmetries in early human development: a literature review and a neuroembryological model. *In* Brain Lateralization in Children: Developmental Implications. D.L. Molfese & S.J. Segalowitz, Eds.: 5–34. Guilford Press. New York.
145. MOLFESE, D.L., L.M. BURGER-JUDISCH & L.L. HANS. 1991. Consonant discrimination by newborn infants: electrophysiological differences. Dev. Neuropsychol. **7:** 177–195.
146. DEHAENE-LAMBERTZ, G. 2000. Cerebral specialization for speech and non-speech stimuli in infants. J. Cogn. Neurosci. **12(3):** 449–460.
147. MARCOTTE, A.C. & D.A. MORERE. 1990. Speech lateralization in deaf populations: evidence for a developmental critical period. Brain Lang. **39(1):** 134–152.
148. FLEGE, J.E., N. TAKAGI & V. MANN. 1995. Japanese adults can learn to produce English /r/ and /l/ accurately. Lang. Speech **38:** 25–55.
149. YAMADA, R.A. & Y. TOHKURA. 1992. The effects of experimental variables on the perception of American English /r/ and /l/ by Japanese listeners. Percept. Psychophys. **52:** 376–392.
150. ZHANG, Y., P.K. KUHL, T. IMADA, *et al.* 2000. MMFs to native and nonnative syllables show an effect of linguistic experience. Paper presented at the 12th International Conference on Biomagnetism, Helsinki, Finland.
151. ZHANG, Y., P. KUHL, T. IMADA, *et al.* 2000. Neural plasticity revealed in perceptual training of a Japanese adult listener to learn American /l-r/ contrast: a whole-head magnetoencephalography study. Paper presented at the 6th International Conference on Spoken Language Processing, Beijing, China.
152. MCCLELLAND, J.L., A. THOMAS, B.D. MCCANDLISS & J.A. FIEZ. 1999. Understanding failures of learning: Hebbian learning, competition for representational space, and some preliminary experimental data. *In* Brain, Behavioral, and Cognitive Disorders: The Neurocomputational Perspective. J. Reggia, E. Ruppin & D. Glanzman, Eds.: 75–80. Elsevier. Oxford, UK.
153. PISONI, D.B. 1992. Talker normalization in speech perception. *In* Speech Perception, Production and Linguistic Structure. Y. Tohkura, E. Vatikiotis-Bateson & Y. Sagisaka, Eds.: 143–151. Ohmsha Press. Tokyo.
154. BALTAXE, C. & J.Q. SIMMONS. 1985. Prosodic development in normal and autistic children. *In* Communication Problems in Autism. E. Schopler & G. Mesibov, Eds. Plenum Press. New York.

155. DE GELDER, B. & J. VROOMEN. 1998. Impaired speech perception in poor readers: evidence from hearing and speech reading. Brain Lang. **64:** 269–281.
156. MCBRIDE-CHANG, C. 1995. Phonological processing, speech perception, and reading disability: an integrative review. Educ. Psychol. **30:** 109–121.
157. REED, M.A. 1989. Speech perception and the discrimination of brief auditory cues in dyslexic children. J. Exp. Child Psychol. **48:** 270–292.
158. WERKER, J.F. & R.C. TEES. 1987. Speech perception in severely disabled and average reading children. Can. J. Psychol. **41:** 48–61.
159. WOOD, C. & C. TERRELL. 1998. Poor readers' ability to detect speech rhythm and perceiving rapid speech. Br. J. Dev. Psychol. **16:** 397–413.
160. LEONARD, L.B., K.K. MCGREGOR & G.D. ALLEN. 1992. Grammatical morphology and speech perception in children with specific language impairment. J. Speech Hear. Res. **35:** 1076–1085.
161. STARK, R. & J. HEINZ. 1996. Vowel perception in language-impaired children and language normal children. J. Speech Hear. Res. **39:** 860–869.
162. MORGAN, J.L. & K.D. DEMUTH, Eds. 1996. Signal to Syntax: Bootstrapping from Speech to Grammar in Early Language Acquisition. Lawrence Erlbaum Associates. Mahwah, NJ.
163. TSAO, F-M., H-M. LIU & P.K. KUHL. 2001. The relationship between speech perception and early language development. Manuscript in preparation.
164. KOHONEN, T. 1982. Self-organized formation of topologically correct feature maps. Biol. Cybernet. **43:** 59–69.
165. GUENTHER, F.H. & M.N. GJAJA. 1996. The perceptual magnet effect as an emergent property of neural map formation. J. Acoust. Soc. Am. **100:** 1111–1121.
166. GUENTHER, F.H. 1994. A neural network model of speech acquisition and motor equivalent speech production. Biol. Cybernet. **72:** 43–53.
167. GUENTHER, F.H. 1995. Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. Psychol. Rev. **102:** 594–621.
168. DOMINEY, P.F. 1998. A shared system for learning serial and temporal structure of sensori-motor sequences? Evidence from simulation and human experiments. Cogn. Brain Res. **6(3):** 163–172.
169. GUENTHER, F.H. & D.M. BARRECA. 1997. Neural models for flexible control of redundant systems. *In* Self-organization, Computational Maps and Motor Control. P.G. Morasso & V. Sanguineti, Eds.: 383–421. Elsevier-North Holland. Amsterdam.
170. DOMINEY, P.F., T. LELEKOV, J. VENTRE-DOMINEY & M. JEANNEROD. 1998. Dissociable processes for learning the surface structure and abstract structure of sensorimotor sequences. J. Cogn. Neurosci. **10(6):** 734–751.
171. DE BOER, B. & P.K. KUHL. 2001. Investigating the role of motherese in vowel acquisition with computer models. Paper presented at Speech Recognition as Pattern Classification Workshop, Nijmegen, the Netherlands.
172. DE BOER, B. 2001. Instinctive learning of vowel categories. Paper submitted for International Joint Conference on Artificial Intelligence, Seattle, WA.
173. PICKETT, J.M. 1999. The Acoustics of Speech Communication: Fundamentals, Speech Perception Theory, and Technology. Allyn and Bacon. Needham Heights, MA.