

# THE ORGANIZATION AND REORGANIZATION OF HUMAN SPEECH PERCEPTION

*Janet F. Werker and Richard C. Tees*

Department of Psychology, University of British Columbia,  
Vancouver, BC V6T 1Z4, Canada

KEY WORDS: infancy, speech perception, development

One of the most fundamental attributes of being human is the ability to perceive and produce language. Although humans communicate by sign and writing, by far the most common and enduring form of human communication is spoken language, or speech. The unique set of abilities that characterizes speech perception, and the early appearance of these abilities during human ontogeny, suggests that these capacities may be deeply rooted in our biology. In this chapter we review selective behavioral work on the development of speech perception in humans and try to relate these empirical findings both to the ontogeny of auditory/communicative abilities in infrahumans and to their neurobiological substrates.

We start with an overview of some of the fundamental characteristics of speech perception that make it an intriguing area of inquiry and suggest the involvement of specialized biological predispositions. We then survey research on developmental changes in speech perception. This section begins with a characterization of the “initial state” of the human infant’s ability to process speech sounds (including a consideration of potential prenatal environmental influences on this state), and then reviews empirical work examining postnatal changes in speech processing, with a focus on our own work in cross-language speech perception. In the course of this analysis we relate this work, albeit selectively, to what we believe is parallel research involving the processing of auditory/communicative signals by nonhuman animals. In the final section of the chapter, we briefly speculate about possible neurobiological processes that might account for the nature of the recent developmental evidence.

377

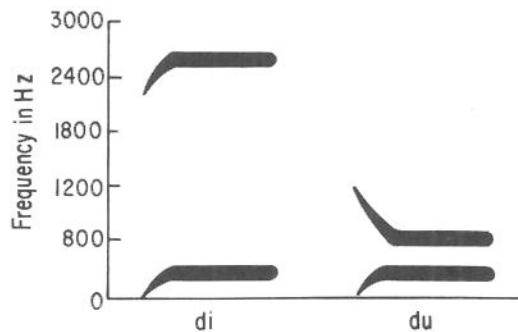
0147-006X/92/0301-0377\$02.00

## SPEECH PERCEPTION

We typically produce approximately 12–14 individual speech sounds per second, three to four times the number of sequential arbitrary sounds that can be perceived (Warren et al 1969), and can perceive 50 or 60 segments per second with little loss in intelligibility. The speed with which speech is produced and perceived is difficult to explain, particularly given the complex computational requirements for perceiving human speech. This complexity takes two forms (Liberman et al 1967). The first is a segmentation problem. Although we think of speech as linear, there are, in reality, no clear breaks between words, syllables, and the basic elements called phonemes. Rather, the information that specifies a given phoneme is spread over several surrounding segments and often crosses syllable and even word boundaries. This is a result of coarticulation (the influence of one phoneme on another in the articulatory process).

The second problem is a lack of invariance. Simply stated, to date, no single (or even set of) invariant property(ies) in the acoustic signal has been successfully identified as necessary and sufficient for a given phoneme. Indeed, the acoustic characteristics of a given phoneme vary tremendously in different contexts (see Figure 1). Attempts to understand speech perception in terms of acoustic information have now moved to a consideration of relational information, but substantial ambiguity still exists.

The lack of equivalence between the acoustic signal and the perceived segments of an utterance has generated considerable theoretical controversy (see Mattingly & Studdert-Kennedy 1991 for discussion), which has been the catalyst for much of the empirical work. The controversy pits “specialized speech-specific” explanations against “generalized auditory” theories of speech perception. Proponents of the *speech-is-special* point of



*Figure 1* Spectrographic patterns sufficient for the synthesis of /d/ before /i/ and /u/. In /di/ the second formant transition rises, in /du/ it falls. (Adapted from Liberman et al 1967.)

view identify instances of different data patterns in the perception of speech vs. nonspeech sounds, whereas adherents of the *general auditory* approaches attempt to show that such differential data patterns do not exist.

One of the first pieces of evidence for specialized speech processing was the finding that human adults show "categorical" perception for speech sounds, and thus respond to only those acoustic variations which signal a difference in meaning (Liberman et al 1967). This helps to explain the perceptual invariance in light of the enormous acoustic variability. More recent work has shown that our perceptual system actually adjusts to contextual influences on speech production. For example, when one speaks rapidly, the formant transitions in a stop consonant will be shortened. Our perceptual systems calibrate for that (Miller & Liberman 1979) and other contextual influences.

Studies revealing that speech is perceived bimodally through our eyes as well as our ears provide further evidence that speech is not perceived like other acoustic signals (Summerfield 1991). For example, when shown a "talking head" producing the articulatory movements for the syllable /ba/, and presented with an acoustic /ga/, human adults typically report perceiving an instantaneous, unambiguous /da/ syllable (McGurk & MacDonald 1976). Other combinations of articulatory and acoustic information lead to different percepts, but the immediacy of the integrated speech percept is what is compelling.

The occurrence of duplex perception has also been interpreted as highlighting the independence of speech perception from the more general processing of other acoustic signals. Basically, when a speech signal is taken apart such that the base component is presented to one ear, and two of the formant transitions (representing rapid changes in frequency) to the other ear, adult listeners report two simultaneous percepts. One corresponds to the original fully integrated syllable, and the other to a non-speech glissando—the sound that results from the transitions alone (Liberman & Mattingly 1989). Experimental manipulations indicate that when the transitional and base components are of equal intensity, only an integrated syllable is heard. The transitions have to be presented at a considerably higher intensity than the base in order for the duplexity effect to be evident (Whalen & Liberman 1987), thus suggesting that the speech percept "preempts" the non-speech percept.

Virtually all of these "special" aspects of speech perception have been shown to hold true for at least some non-speech sounds as well. The perception of non-speech, tone-onset-tone continua is partially categorical. The context effect with respect to "speaking rate" has been shown to extend to sine-wave stimuli, and duplex perception has been shown to

extend to several domains, including musical stimuli and nonsignificant auditory events such as the slamming of a door (for reviews, see Fowler & Rosenblum 1991, Pisoni & Luce 1986). Thus, it is clear that these effects are not unique to the perception of speech. Speech, in particular consonant-vowel (CV) syllables, however, constitutes the only kind of signal for which all these effects can be demonstrated. It can be argued, therefore, that it is not the particularity of any of these effects, but rather their universality that provides the weight of evidence pointing to biological preparedness and perhaps biological specialization.

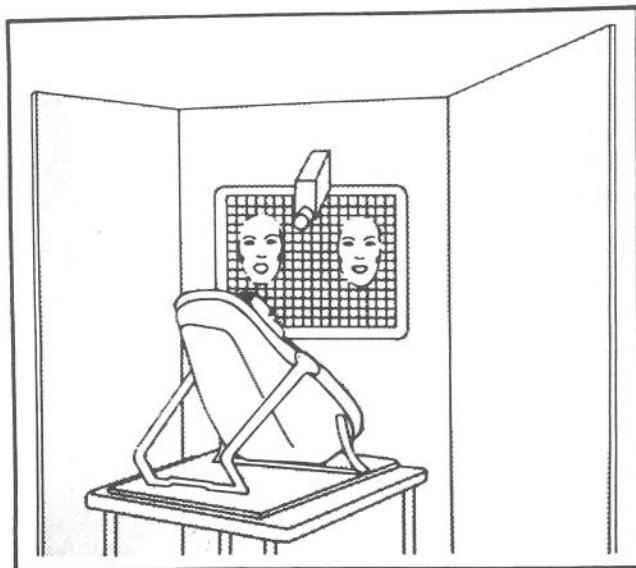
Neuropsychological studies provide more direct support for differential processing of speech sounds. Considerable evidence has shown that the rule-based aspects of language processing (phonology, syntax) are predominantly a left-hemisphere function, even, interestingly, in the signing of deaf adults with damage to one of the hemispheres (Poizner et al 1990). Well-known demonstrations of left-hemisphere processing include the early studies of aphasic patients showing a loss of speech production abilities following left, but not right, hemisphere damage, and of split-brain patients. Studies of regional brain blood-flow, auditory evoked responses, carotid sodium amytal injection, cortical stimulation, and dichotic listening indicate a right-ear, left-hemisphere advantage in the perception of consonant-vowel syllables (e.g. Kimura 1967; for a review see Kolb & Whishaw 1990).

## SPEECH PERCEPTION IN HUMAN INFANTS

Perhaps the strongest evidence in support of some kind of biological preparedness comes from studies indicating that many of these behavioral and neuropsychological indices are evident even in very young infants. Prelinguistic infants as young as one month of age show categorical-like perception of speech sounds (Eimas et al 1971). When tested in a high amplitude sucking procedure, after repeated presentations of consonant-vowel syllables from one phonetic category, the infants will show an increase in sucking rate to a new stimulus, but only if it is from a contrasting (adult) phonetic category. More recent research has shown this categorical-like speech perception to be quite general (e.g. Kuhl 1987) and even to extend to phones the infant has not experienced in his/her language-learning environment (for a review, see Werker 1991), thus providing quite convincing evidence that the phonetic-relevance of early infant speech perception does not rely on specific language experience. In other words, perceptual biases exist from birth that predispose an infant to be able to discriminate acoustic variability that signals phonetic distinctiveness, and to be less able to discriminate equal-sized acoustic variability that exists

between stimuli from within a single phonetic category. Such perceptual constancy is evident even across variability in intonation contour and speaker identity.

Infants also display several other quite remarkable capacities with respect to speech perception. They show context effects with respect to speaking rate (Eimas & Miller 1980). Thus the boundary between /ba/ and /wa/ shifts in accordance with the overall duration of the syllable. Perhaps even more surprisingly, infants recognize bimodal equivalences in speech. For example, in Kuhl's work, infants of 4 months of age were shown two filmed faces side by side, with a single vowel sound presented from a loudspeaker located midway between the two facial images (see Figure 2). One of the faces articulated the /a/ vowel while the other articulated the /i/ vowel (side of presentation was counterbalanced). On half the trials the vocal stimulus was /a/, and on half the trials it was /i/. These neonates clearly demonstrated recognition of auditory-visual equivalence in speech by selecting to look at the face that "matched" the vowel being spoken 73% of the time (Kuhl & Meltzoff 1982).



*Figure 2* Infants' cross-modal speech-perception abilities are tested by presenting them with two facial images, one articulating the vowel /a/ and the other the vowel /i/. One or the other of the sounds is presented from a loudspeaker midway between the two facial images. The results show that infants look longer at the face that "matches" the speech sound they hear. (Adapted from Kuhl & Meltzoff 1982.)

These same infants also showed a rudimentary detection of auditory-motor equivalences by tending to match their vocal productions with the female speaker's vowel production. The match was apparent both in terms of prosodic characteristics (pitch control and duration) and the formant characteristics that distinguish the vowels /a/ and /i/.

Finally, neuropsychological studies reveal some evidence of left-hemispheric specialization for speech perception in early infancy. Infants display a right-ear advantage in the processing of stop consonants and a left-ear advantage for the perception of musical sounds (Best et al 1982). Both full-term and premature infants display asymmetric left-hemisphere auditory evoked responses to at least some consonant signals (Molfese & Molfese 1980). The comparability in such "test results" between infants and adults argues strongly against an interpretation that the specialized nature of speech perception is acquired through extensive experience of hearing speech in early infancy, and suggests instead a propensity that exists from birth. Whether this propensity is specific to speech or reflects more general auditory or sensorimotor abilities that are also applied to speech remains uncertain. There is no doubt, however, that the young infant is born with abilities that greatly enhance the speed, efficiency, and accuracy with which he/she can process linguistic input. At the very least, this suggests the existence of neural circuitry that is ideally suited for processing speech.

The existence of neural circuitry that is utilized rapidly to disambiguate complex signals has certainly been documented in other perceptual/motor domains. The "sophisticated" neural characteristics of the barn owl's inferior colliculus that support the animal's ability to localize sound and react appropriately (Knudsen & Konishi 1978) is perhaps the most widely cited example. The features of the neural circuitry necessary to disambiguate ambiguous visual objects, including faces (e.g. Perrett et al 1982), represent another example of a biologically significant signal that is likely to be processed by specialized central neural mechanisms in humans. The general proposition that organisms might be adapted to process, attend, or associate particular kinds of behavior with such specific stimuli is well established (e.g. Garcia & Koelling 1966, Marler 1990).

Several different neurobiological models have been proposed for explaining the specialized characteristics of human speech perception. As one example, Sussman (1989) extrapolates from the sound-localizing system of the barn owl and provides a speculative conceptualization of how a collective processing network could analyze speech sounds. Sussman proposes that hierarchically arranged arrays of biologically specialized combination-sensitive neurons might be able to account for perceptual constances both in the perception of place cues for stop consonants and for vowel normalization.

We think two points are important to keep in mind with respect to this and other auditory-based neural models of speech perception. First of all, topographical maps of auditory features such as sound localization in both cortical and subcortical structures are, in fact, sensorimotor maps. For example, both the inferior and the superior colliculus contain ordered maps of motor space that are in sensory register with visual, auditory and somatosensory maps. The little neurobiological work that has been conducted with infrahumans with respect to species-specific calls/speech suggests an intimate link between perceptual and motor processes. For example, Williams and Nottebohm (1985) report evidence that the motor system responsible in part for song production in male infant zebra finches is selectively tuned for perceiving the critical elements of song. Similar observations can be made for the colliculi of rodents, barn owls and birds in connection with the analysis of visual, auditory, and spatial location. The perception of biologically significant information across a variety of species thus seems to be organized in overlapping sensorimotor systems. As far as speech itself is concerned, electrical stimulation of the same cortical sites disrupts both the production of sequential oral facial movements and the ability to discriminate phonemes (Ojemann 1983). At some level it seems undeniable, then, that articulatory and perceptual neural systems should be highly integrated in the case of human speech (Mattingly & Liberman 1989).

Second, the evidence that is available (e.g. Tees 1990a,b) makes it clear that these perceptual-motor systems are altered by changes in auditory, visual, and tactile stimulation history. The impact of the environment is most dramatic neonatally, but there is unequivocal evidence (e.g. King & Moore 1991, Merzenich et al 1984) that perceptual-motor neural representations, and the related behavioral competences, remain somewhat plastic throughout life. Thus at the outset, we propose that both neonatal abilities and postnatal changes in speech perception competences might be best explained by reference to a specialized but flexible perceptual-motor system.

### PERCEPTION OF SPECIES-SPECIFIC CALLS BY ANIMALS

An evolutionary perspective, in which one predicts continuity among phenomena and mechanisms across species, focuses attention on the evidence that many nonhuman animals also have specialized mechanisms for the perception of their own species calls.<sup>1</sup> This evidence can be examined

<sup>1</sup> The questions addressed by this research are distinct from those examined in investigations of the perception of human speech by nonhuman animals. That work is not reviewed here (see Kuhl 1988 for an overview).

in the context of the speech-processing characteristics mentioned above and helps put specialized speech processing into a biological context (see Petersen 1982 for a review).

One interesting set of experiments (reviewed by Ehret 1987) focuses on evidence of *categorical perception* of species-typical communicative signals. Both labeling and discrimination tests involving synthetic series of calls relevant to mice, vervet and macaque monkeys, and pygmy marmosets have been undertaken and, in each case, good evidence for categorical perception has been found.

Japanese macaque (*m. fuscata*) monkeys not only show categorical perception, but also show the kind of perceptual constancy for their species-specific calls that humans demonstrate for speech sounds. The Japanese macaque produces 80 to 90 different vocalizations. Two of these calls can be distinguished on the basis of whether a frequency-modulated sweep occurs early or late in the signal. Using operant techniques, Green (1975) found that the Japanese monkeys are able to discriminate the "linguistically distinctive" frequency-modulated segment of the calls and ignore other, noncritical acoustic information. Other, non-Japanese, "comparison" (pigtailed and bonnet macaque) monkeys who do not utter these calls require extensive training to master this task. In a subsequent experiment, all monkeys were required to differentiate these same calls on the basis of a nonlinguistic acoustic cue—high vs. low fundamental frequency—instead of temporal location of frequency-modulated sweep within segments. Although the comparison monkeys quickly learned the discrimination, the Japanese monkeys had great difficulty. These findings are consistent with the possibility that Japanese monkeys use neuronal mechanisms different from those of the "comparison" monkeys in perceiving Japanese macaque calls.

Direct support for the possibility of specialized neural processing for macaque calls comes from measurements of lateralization. Japanese monkeys show a right-ear advantage, indicative of left hemisphere dominance in their ability to discriminate pairs of macaque stimuli presented alternatively on a random basis to one ear at a time. The comparison monkeys exhibit no such right-ear advantage. Although comparison and Japanese monkeys can and do use the same features of the calls when performing the discriminations (Petersen et al 1984), the cortical lesions that disrupt macaques' ability to discriminate macaque calls are specific to the left temporal cortex, and are distinct from other regions that, when lesioned, yield general hearing losses (Heffner & Heffner 1989). These results from Japanese monkeys have a parallel in comparative studies of Thai and English speakers. In Thai, changes in fundamental frequency serve a critical linguistic function by specifying semantic distinctions among

words. In English such changes serve only paralinguistic functions such as providing information about the speaker's emotional state. Thai speakers demonstrate a right-ear advantage for fundamental frequency whereas English speakers may not (Van Lancker & Fromkin 1973).

Another intriguing parallel with features of human speech involves vocalizations directed primarily at infant animals by squirrel monkey mothers (Biben et al 1989). Like humans, squirrel monkeys rarely use vocalizations typical of adult "conversations" when addressing their youngsters, and instead use "caregiver" calls with several unique features (see Fernald 1984 for a review of the human work). However, whereas infant-directed speech in humans is typically higher pitched than adult-directed speech, caregiver calls in squirrel monkeys have most of their power concentrated at several kilohertz lower than most squirrel monkey calls. A special "baby talk" register is used, as in humans (Ferguson 1964).

### EXPERIENCE, GENETIC PREDISPOSITION, AND THE NEONATE

If we accept the notion that there is specialization for the perception of biologically significant communicative signals, it becomes of interest to explore the ontogenetic roots of such specialization (see also Miller & Jusczyk 1989). The most obvious explanation for the remarkable capabilities of human infants is that genetic and maturational factors leave the neonate with a specialized processing system for human speech. We would argue, however, that the research to date is more consistent with the hypothesis that the abilities that exist in early infancy reflect the probabilistic outcome of both endogenous and environmental factors. The idea that we would like to entertain (as have others) is that normally inevitable pre- and postnatal experiences (i.e. species-typical) are not only influential but necessary for normal development (e.g. Gottlieb 1985a, Hebb 1980). Psychologists would characterize the resulting process as *innately guided learning* (Jusczyk & Bertoni 1988), while neurobiologists might want to describe it in terms of *activity-dependent neurogenesis* (e.g. Greenough et al 1987). Both reflect a probabilistic epigenetic viewpoint that the phenotype or endpoint is a complex outcome of both endogenous and exogenous factors, including species-invariant sensory input. The same kind of argument for probabilistic epigenesis has been made in the case of the ontogeny of early-appearing visual competences such as stereopsis, binocularly, and avoidance of heights. For each of these basic visual abilities, the manipulation of normally "inevitable" input results in significant performance deficits (Tees 1990a,b, Mitchell 1989).

There is now strong evidence that the auditory system of the human fetus is partially functional during the last prenatal trimester. The mother's voice is transmitted to the uterus by bone conduction, as well as, perhaps, other human voices via air and fluid conduction. Although there is still considerable controversy concerning the precise acoustic information that is available to the human fetus, it is generally agreed that at the very least, low-frequency information (<400 Hz) is available (Fifer & Moon 1988). The attenuated linguistic input that thus reaches the fetus may well play a part in sculpturing the developing speech or nonspeech-related auditory systems and facilitating the neonate's processing of speech and speech-like sounds. We know from a variety of studies (e.g. Smotherman & Robinson 1989) that nonhuman mammalian fetuses are responsive to sensory stimuli encountered during gestation, and that prenatal stimulation may alter their postnatal behavior (Turkewitz 1988). The first persuasive demonstration of prenatal influences on speech processing in humans was provided in a study by DeCasper & Fifer (1980), who demonstrated that on the first postnatal day, human babies of either gender will suck (an artificial nipple) preferentially to hear a tape of their mother's voice over that of another female. (Neonates showed no such preference for their father's voices.) There is also data indicating that speech stimuli with familiar vs. unfamiliar melodic and/or temporal characteristics elicit different cardiac patterns in human fetuses and that neonates will suck preferentially to a song or a story heard prenatally (Fifer & Moon 1988).

More recently, Mehler et al (1988) have shown that infants can discriminate global characteristics of native from nonnative speech samples within days after birth. Two-month-old American-English-learning infants in Oregon were compared to 4-day-old French-learning infants in Paris on their ability to discriminate French vs. Russian and Italian vs. English samples of speech. The four-day-old French-learning infants discriminated the French vs. Russian samples, but not the English vs. Italian speech. In contrast, the English-learning infants discriminated English from Italian, but not French from Russian. Thus, each group of infants was able to discriminate its native language from an unfamiliar language, but was unable to discriminate two unfamiliar languages. These results are interpreted by the authors as indicating that prenatal experience with a language facilitates processing of the global properties of that language, thus allowing infants to distinguish or prefer the native language over an unfamiliar one.

Together, these findings help us understand the human postnatal preference for familiar stimuli such as mother's voice and native language. We would like to suggest that they may also help explain the young infant's abilities to perceive the phonetic aspects of speech in a categorical-like fashion, the evidence of trading relations in infant perception, and perhaps

even the infant's ability to discriminate nonnative sounds at birth. Our argument is as follows. Although it is very likely that detailed phonetic information is not available prenatally, some aspects of the speech signal do reach the fetus. The input that is available presumably provides enough information to enable infants to have greater facility in processing variability that corresponds to that experienced prenatally. Thus, the linguistic relevance of speech perception in the newborn may reflect an epigenetically determined processing specialization for stimuli that share global characteristics with those experienced prenatally. There is undoubtedly an innate propensity for this bias, but we are arguing that there may also be an experiential component. In terms of Greenough et al's (1987) distinction between *experience-expectant processes* and *experience-dependent processes*<sup>2</sup> regarding the roles played by early sensory experience/learning, we are identifying the experiential impact in terms of experience-expectant processes.

## DEVELOPMENTAL STUDIES OF NONHUMANS

Although primates are *the* choice for investigating possible parallels in the ontogeny of systems for species-specific calls, data on the development of such abilities in primates is limited. In contrast, there have been many developmental studies on the effects of experience on vocal communication in birds. Most of the evidence for experiential effects on vocal communication rely on data on vocal output. Marler and his colleagues have conducted a number of studies revealing that for many songbirds, experience "hearing" a species-specific song during a sensitive period in early development is required in order to sing the species-specific song correctly as an adult (see Marler 1990). The requisite amount and timing of exposure varies tremendously across avian species, but in most cases the birds are not capable of correctly singing the song of another species even if given considerable early experience with such songs. Thus, in general, the data pattern is consistent with the notion that experience-expectant processes play an important role in the development of vocal communication.

In addition to simple exposure to the song or calls of conspecifics, other types of "nonobvious" experience are also possible (e.g. Johnston 1988). These include self-produced sounds (including calls), sounds of siblings,

<sup>2</sup> Greenough categorized two ways experience might influence the developing (and mature) organism in terms of the type of information stored and the brain mechanism involved. Experience-expectant involves incorporation of ubiquitous early environmental information by means of selective sculpturing of intrinsically overproduced synaptic connections between sensory/motor neurons. Experience-dependent involves the storage of other experiences by means of activity-associated generation of new synapses.

parents, and other conspecifics (both prenatally and postnatally), and nonvocal social stimulation. These nonobvious sources of experience can also exert a profound impact on vocal communication. For example, West and King (1988) have shown that the vocal behavior exhibited by the adult cowbird depends not only on early acoustic experiences, but also on the bird's social environment throughout its life. In illustration, an adult male cowbird can learn a new dialect as an adult when in the company of a (nonsinging, but potentially responsive) female cowbird familiar with only the new dialect.

Gottlieb (1985a,b) has also examined nonobvious experiential influences on communication development in several different species of ducklings, using responsiveness to species-specific calls rather than vocal production as the dependent variable. Gottlieb has shown that this species-specific responsiveness required relevant exposure in embryo. In the Peking duck, this experience can be either the calls of conspecifics in the environment or self-produced vocalizations—but the experience must occur during a relatively constrained period of development for the duckling to show the typical species-specific preference. Interestingly, in devocalized ducklings, the embryonic experience need not match the calls precisely. In fact, for the Peking duck, variability in the input produces greater preference than does the "ideal" repetition rate in the case of the contact call. Similarly, for the wood duck, prehatching exposure to a frequency range different from that of the assembly call is adequate for ensuring a post-hatching preference for their assembly call. These findings confirm the role that experiential influences play even in species-specific behaviors. In addition, they provide a concrete instance of our suggestion that non-identical but related exposure in embryo can canalize the perception of species-specific communicative signals in human infants.

Drawing precise parallels between most of these studies and work on human speech perception is difficult because the studies of these birds rely on production<sup>3</sup> and preference to reflect perceptual competence. Recent work on birdsong by Sinnott (e.g. 1987) comparing blackbirds, cowbirds, and humans after training to discriminate both birds' full songs and song elements is more promising in this regard. In this instance, species-specific coding was more apparent when birds were processing information in

<sup>3</sup>The problem with using vocal output as a sole index of perceptual processing is that it may reflect only the properties of the receptive "apparatus" that interact or guide the motor systems involved in producing vocalization. The perceptual system may well be a complex multistage one, able to process much more about the signals than is reflected in the animal's vocal behavior and only some of the stages involved in directing the motor program, e.g. for song acquisition. In any event, if we relied only on vocal output we certainly would grossly underestimate the sensory and perceptual capacities of prelinguistic human infants.

the complex, full-song context than it was when birds were specifically "tutored" with song elements out of context. These results are strikingly analogous to the pattern of data from speech perception studies with adult humans showing auditory and phonetic processes in development.

## CROSS-LANGUAGE SPEECH PERCEPTION IN INFANTS

One way of collecting evidence on the interaction between biological endowment and experiential factors in human speech development is to examine speech perception in a cross linguistic framework. This allows an evaluation of the effects of naturally occurring variations in input on speech perception abilities. Developmental processes can be assessed by comparing the effects of variation in input at different ages. We began to address this question by examining language-specific influences on the perception of the phonetic aspects of speech in infants and adults.

Previous work had shown that although infants discriminate both native and nonnative speech contrasts according to phonetically relevant boundaries (Trehub 1976), adults often show difficulty discriminating nonnative speech contrasts. Thus, the existing work (Strange & Jenkins 1978) had suggested a change between infancy and adulthood in the ease with which listeners discriminate unfamiliar phonetic contrasts. Our work has been designed to investigate the validity and meaning of this claim. In a series of experiments conducted over the last ten years, we have confirmed that there is a profound developmental change between infancy and adulthood in the ease with which listeners can differentiate phonetic contrasts that are not used in their native language. Furthermore, we have shown that this developmental change from broad-based to language-specific phonetic perception is evident as early as 10–12 months of age. We briefly review this work below.

In an early experiment, English-speaking adults, Hindi-speaking adults, and English-learning infants aged six–eight months were compared on their ability to discriminate two Hindi speech contrasts: the retroflex/dental place-of-articulation contrast, /Ta/-/ta/, and the voiceless aspirated vs. breathy voiced dental stops, /t<sup>h</sup>a/-/d<sup>h</sup>a/. Natural rather than synthetic stimuli were used in this original study to allow us to assess discrimination and categorization of phonetic categories within the context of at least some naturally occurring variation. Infants and adults were tested in a variation of the head-turn procedure (see Figure 3). The basic logic of this procedure is that the infant is conditioned to turn her head when she detects a change in the speech stimulus. Correct (but not incorrect) head turns are reinforced with the activation of single or multiple mechanical



Figure 3a An infant attending to the experimental assistant during the "no-change" intervals in the head-turn procedure.



Figure 3b An infant turning her head toward the visual reinforcer upon detecting a change in the speech sound. (Notice that the correct head turn is "reinforced" by the activation of the toy animals as well as by clapping and praise by the experimental assistant.)

toy animals. Adults and older children indicate detection of a change by pushing a button (for fuller details of this procedure, see Kuhl 1987). The results indicated that although all the Hindi adults and most of the young English-learning infants could reach discrimination criterion on the two Hindi contrasts, the majority of the English-speaking adults could not, showing particular difficulty on the retroflex/dental distinction. Indeed, when a second group of English-speaking adults was given 25 training trials on each of the Hindi contrasts, their performance improved on the voicing distinction, but training did not affect performance of the more difficult retroflex/dental contrast (although we have subsequently developed procedures in which sensitivity to even the retroflex/dental contrast is maintained). These results confirmed that although young infants are equally sensitive to both native and nonnative phonetic contrasts, adult perception is modified by language experience, and the impact of experience is more profound for some nonnative contrasts than it is for others (Werker et al 1981).

A series of experiments was run to try to identify the age at which the developmental change in sensitivity is first apparent. After first finding that children aged 12, 8, and even as young as 4 years have difficulty with some nonnative contrasts (Werker & Tees 1983), we eventually discovered evidence of a developmental change by 10–12 months of age. Briefly, English-learning infants of 6–8, 8–10, and 10–12 months of age were compared on their ability to discriminate two non-English phonetic contrasts as well as the English bilabial/alveolar contrast, /ba/-/da/. The non-English contrasts were the Hindi retroflex/dental and an Nthlakampx (a Northwest, Interior Salish language) glottalized velar vs. glottalized uvular contrasts. The youngest English-learning infants could discriminate all three sets of contrasts, but the infants aged 10–12 months could only discriminate the native language /ba/-/da/ distinction (see Figure 4). To ensure that the performance of the older infants was not simply a general age-related performance decline, we tested a few Hindi- and Nthlakampx-learning infants aged 11–12 months and found they could quickly reach a 9 out of 10 discrimination criterion on their native contrast (Werker & Tees 1984a).

We have replicated this finding of developmental change between 6–12 months of age with a synthetically produced Hindi retroflex/dental contrast involving voiced rather than voiceless stimuli (/Da/ vs. /da/ rather than /Ta/ vs. /ta/) (Werker & Lalonde 1988). Also, Best and McRoberts have replicated the developmental change between 6 and 12 months of age for the Nthlakampx contrast by using a different procedure (C. T. Best, in preparation). They have also recently reported data showing a similar age-related reorganization for three other click contrasts.

**INFANT SUBJECTS REACHING CRITERION  
ON HINDI AND SALISH CONTRASTS**

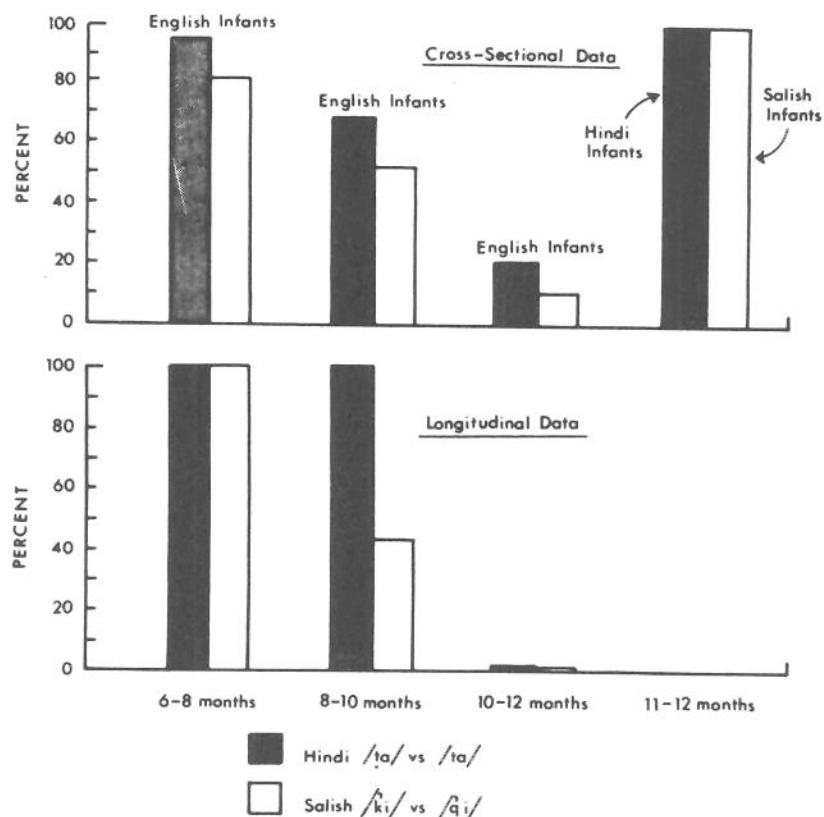


Figure 4 The proportion of infants at three different ages able to discriminate the Hindi and Nthlakampx (non-English) phonetic contrasts. (Adapted from Werker & Tees 1984a.)

One very interesting study by Best et al (1988) did not replicate our results. In this study, English-learning infants of 6-8, 8-10, 10-12, and 12-14 months of age were compared to both Zulu- and English-speaking adults on their ability to discriminate the difference between a medial vs. lateral Zulu click contrast. Not only does this contrast not occur in English, clicks do not have lexical status at all. Of interest, the English adults and the infants of all four ages were able to discriminate this nonnative contrast. Best has begun to develop a scheme for predicting which kinds of nonnative contrasts will be easy and which difficult to discriminate.

Polka is currently conducting a set of experiments with one of us that examines developmental changes in cross-language vowel perception (L. Polka and J. F. Werker, *in preparation*). Infants of 6–8 and 10–12 months of age and English- and German-speaking adults are being compared on their ability to discriminate two sets of German (non-English) vowel contrasts as well as one English vowel contrast. In all cases the vowels are embedded in between the consonants /d/ and /t/. The English distinction contrasts /deet/ and /dot/. The German contrasts both involve front vs. back distinctions between rounded vowels. One contrast involves two lax vowels /dvt/ vs. /dyt/ and the other, two tense vowels /bv:t/ vs. /by:t/. The results to date show that English-speaking adults can discriminate both of these contrasts with relative ease; however, English-learning infants of 10–12 months<sup>4</sup> show considerably more difficulty. These data show that even when the nonnative contrast is quite acoustically distinct, if the stimuli map onto native-language phones, there will be evidence of a prelinguistic reorganization in perceptual performance.

To find out more about the nature of the discriminative abilities of young (6–8-month-old) infants and whether they “categorized” even nonnative stimuli according to phonetic identity, we utilized synthetically produced retroflex/dental stimuli. A 16-step /ba/-/da/-/Da/ continuum was synthesized varying in 16 equal steps according to the starting frequency of the first and second formant transitions (Werker & Lalonde 1988). English-learning infants of both 5–8 and 11–13 months of age were tested in three conditions (see Figure 5 for stimulus selections). In the first two, adjacent sets of three stimuli were drawn from a section of the continuum such that they were identified as /ba/ and /da/, respectively, by adult (English and Hindi) listeners. This condition would be phonemic to either a Hindi- or an English-speaking adult listener. In the second, two sets of adjacent stimuli were selected that are identified as dental /da/ and retroflex /Da/ by adult Hindi listeners but are all identified as alveolar /da/ by English listeners. This condition would be “phonemic” to a Hindi but not an English listener, and would clearly correspond to a “universal” phonetic category. In the third condition, two sets of adjacent stimuli were selected that are identified as primarily all retroflex by Hindi listeners (and as all alveolar by English listeners) and thus do not correspond to any phonetic category.

It was reasoned that if perception is phonetically relevant in early

<sup>4</sup>The data actually suggest that the English-learning infants of 6–8 months also have difficulty with the two German vowels, which suggests that the reorganization in vowel perception might occur earlier than that for consonant perception. P. Kuhl (personal communication) has also recently found native-language influences on vowel perception by six months.

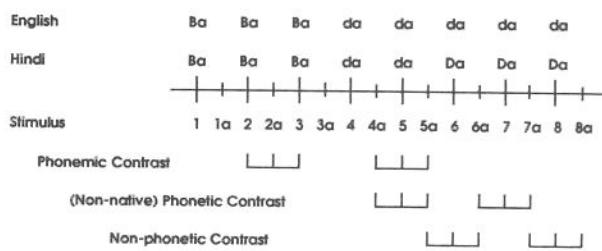


Figure 5. The top part of this figure shows the way English- vs. Hindi-speaking adults divide the synthetic /ba/-/Da/-/da/ continuum. The bottom half of the figure indicates the stimuli that were used as phonemic, phonetic, and nonphonetic contrasts in Werker & Lalonde (1988).

infancy, the infants of 6–8 months of age should be able to perform successfully in the first two, but not the third condition. That is, they should be able to categorize sets of varying stimuli according to native or “universal” phonetic boundaries, but should be unable to categorize stimuli according to an arbitrary point along the continuum that does not conform to a phonetic boundary. This is precisely what was obtained. The result confirms the phonetic relevance of speech perception at the more abstract level of categorization in infants as young as 6 months of age.

This research is similar to the research by Kuhl outlined above showing that by 6 months of age, infants can perceptually categorize vocalic stimuli differing in speaker and intonation on the basis of vowel color, and that they organize vowel categories around phonetic prototypes (e.g. Kuhl 1987). Our cross-language research complements this pattern of data by showing that young infants categorize stimuli, even stimuli that they have not heard before, in terms of phonetically relevant boundaries. The results for the 11–13-month-old infants replicated the findings of Werker & Tees (1984a). These older infants were only able to categorize stimuli according to a phonetic boundary that has functional, phonemic status in their own language but not according to a phonetic boundary that is ignored by adult speakers of their language.

Auditory experience continues to influence competence in speech perception in the months following birth. For example, although infants can detect violations of natural clausal boundaries in both native and non-native speech by 4 months of age (Hirsh-Pasek et al 1987), by 6 months of age they are only sensitive to clausal boundaries in their native language. Further, although infants at this age can detect prosodic differences (melody, stress, tempo) between native and nonnative low-frequency words by as early as 6 months of age, they show no evidence of preferring native

**Table 1** Experimental influences on speech perception

Age	Speech perception competence
Birth	Preference for mother's voice Preference for story and song heard prenatally Ability to discriminate overall prosody of native language from that of a nonfamiliar language Can discriminate "universal" set of phonetic contrasts
4 Months	Sensitivity to clausal boundaries in native and unfamiliar languages
6 Months	Decline in sensitivity to clausal boundaries in unfamiliar languages
9–10 Months	Preference for listening to words with melodic and rhythmical characteristics of native language Decline in sensitivity to vocalic contrasts in unfamiliar languages
10–12 Months	Preference for listening to words that conform to native-language phonotactic <sup>a</sup> rules Decline in sensitivity to consonant contrasts in unfamiliar languages

<sup>a</sup>Phonotactic refers to the phonological rules specifying the letter/sound sequences that are acceptable in a given language.

words on the basis of phonotactic information (the rules for which phonemes can occur together in a sequence) until about 9 months of age (Jusczyk 1992). Overall, the data suggest that the preference of a native over a nonnative language in infants follows a developmental progression in sensitivity from the global features of language-specific sound patterning such as melody and rhythm, to sensitivity to smaller units such as phrases, clauses, individual words, and ultimately individual syllables and phonemes (see Table 1).

## CROSS-LANGUAGE PERCEPTION IN ADULTS

Early research in cross-language speech perception led to the conclusion that the difficulty adults have in discriminating nonnative phonetic contrasts is quite general and irreversible (Strange & Jenkins 1978). More recent research, however, has made clear that the situation is much more complex (for a review see Werker 1991). Developmental changes in sensitivity do not apply equally to all nonnative distinctions (Best et al 1988) and do not indicate an absolute loss of the ability to discriminate nonnative distinctions (Werker & Logan 1985). Furthermore, even when adults clearly have initial difficulty with a nonnative distinction, training studies indicate that they can typically improve after practice or feedback (Logan et al 1991, Tees & Werker 1984) or after intensive study of the relevant language (MacKain et al 1981, Tees & Werker 1984). In many cases,

however, the performance levels they obtain fall far short of those obtained by native speakers (Polka 1991). Finally, early (pre-linguistic) exposure to a language also seems to help maintain sensitivity to the phonetic distinctions (Tees & Werker 1984).

Of perhaps even greater interest, there is evidence that adults may show a latent capacity for discriminating even apparently quite difficult contrasts if the testing procedure is adequately sensitive. For example, we have shown that although adults most readily discriminate CV speech stimuli in terms of native language phonemic categories, sensitivity to nonnative phonetic contrasts is maintained throughout adulthood *even without training* (Werker & Logan 1985, Werker & Tees 1984b). This sensitivity to nonnative phonetic category differences exists in addition to the already known latent sensitivity to acoustic differences within a phonetic category. This led us to propose that adults can process speech in several ways, depending upon task conditions. The most readily available strategy seems to be to perceive speech in terms of native-language phonemic categories. Adults maintain a sensitivity to nonnative phonetic category differences, however, and can even be shown to be sensitive to changes within a phonetic category under some circumstances.

What do these results from cross-language studies with adults tell us about universal capabilities? They tell us that although there are clear experimentally based changes in the ease with which nonnative contrasts can be discriminated, the underlying sensitivity in both universal phonetic and to nonphonetic acoustic differences remains. Thus, it would be incorrect to conclude that lack of listening experience during the first year of life leads to some permanent "loss" in either ability. Nevertheless, these results make clear that there is a substantial developmental change in the ease with which listeners discriminate nonnative contrasts.

## POSSIBLE NEUROBIOLOGICAL MECHANISMS

In previous work, we have outlined several different classes of explanations that might be useful for understanding developmental changes in speech perception (Werker 1991). These include perceptual tuning, modular recalibration, articulatory mediation, phonological development, and cognitive categorization (see also Jusczyk 1991 for an alternate cognitive model). In this chapter, we briefly identify specific processes of neurogenesis that might be related to developmental changes in human speech perception. In this endeavor, we purposely avoid a discussion of general models, such as that proposed by Edelman (1987), and focus instead on what can be said about the involvement of specific events in neurogenesis.

Before considering possible neurobiological mechanisms that might be

related to developmental changes in speech perception, we briefly review what we think is a reasonable description of the key evidence.

1. Young infants, even neonates, show highly developed speech perception abilities, some of which have been influenced by prenatal experiential factors.
2. Speech perception in both adults and infants is complex, multi-dimensional, and linguistically relevant and likely to involve a highly integrated but distributed network of perceptual and motor components.
3. Postnatal linguistic experience during the first year of life influences speech perception competences—first for global characteristics of language (prosody, clausal marking) and later for vowel and consonant perception.
4. This reorganization reflects the sound patterning of the native language.
5. Developmental changes are more pronounced for some nonnative phones than for others and in all cases are reversible.
6. Even without extensive listening experience, adults are able to use different processing strategies—phonemic, phonetic, and acoustic—for categorizing speech sounds under particular task conditions.

The evidence of prenatal and life-long influence on speech processing makes clear at the outset that speech processing is unlikely to be a “closed” developmental program. The genetic program undoubtedly sets *constraints* on the kinds of stimuli that are most easily processed and that can influence the emerging perceptual and neural system, but experience undoubtedly also plays a role in influencing perceptual abilities (Jusczyk & Bertoni 1988, Miller & Jusczyk 1989). At the neural level, *activity-dependent* and *reactive neurogenesis* must underlie the fine-tuning of the perceptual-motor network to these characteristics of the human voice that are available to the fetus. We suggest that these experience-expectant, prenatal processes yield a stable base sensitivity to the broad set of “species-specific” characteristics of human speech. The profound experiential influences evident during the first year of life reflect changes in the ease with which the child can access or use these various species-specific sensitivities. This latter tuning must be more malleable because at least some aspects of this developmental program remain “open” throughout much of the lifespan.

The behavioral evidence for high neonatal ability followed by selective postnatal decline in sensitivity suggests the involvement of regressive processes in postnatal neurogenesis (see also Cowan et al 1984, Kolb 1989). The class of regressive neuronal events that we believe to be most relevant involves competitive sparing of synapses produced in early development (Greenough et al 1987). Such *experience-expectant* plasticity has already

been invoked to help explain the emergence of self-produced locomotion (Bertenthal & Campos 1987). If we invoke this process to help explain the developmental change between language-general and language-specific sensitivities seen across the first year of life, several caveats need to be considered. First, the postnatal developmental changes in language-specific sensitivity are by no means absolute or irreversible, whereas we have suggested that some aspects of prenatal tuning may be resistant to change. Thus any neural mechanism that results in irreversible loss of connections is not adequate to explain the changes seen in the first year of life. Second, it is not enough to say simply that regressive neurogenesis might be the mechanism accounting for particular developmental change. Different kinds of synaptic connections in different areas of the brain proliferate and regress at different points in development. Thus it is necessary to specify the location and timing of regressive neuronal events in relation to the appropriate developmental changes in speech-related behavior.

In considering the reversibility in sensitivity to nonnative speech contrasts, it is important to remember that competitive sparing is not irreversible in the same way that, for example, cell death is. In fact, recent research confirms that new synapses can be generated throughout much of the lifespan (e.g. Greenough et al 1987) and that dysfunctional synapses can, under certain circumstances, regain functionality. For example, when one monocularly deprives a binocular cat, the resulting apparent "loss" in "synaptic control" of binocular cortical neurons is recoverable (Mitchell 1989). Appropriate regimes of visual exposure through both eyes and each eye can reverse the neural and behavioral consequences of early competitive disadvantages. The recovery of sensitivity in adulthood to even difficult nonnative speech contrasts could also be accounted for by such *experience-dependent* neural plasticity. The importance of changes in the inhibitory connections of the neural substrates of such emerging abilities has been discussed elsewhere (e.g. Tees 1990b).

Our ability to relate behavioral changes in speech perception to specific sites of regressive neurogenesis is more problematic. This involves a consideration of the areas of the brain that might be most intimately involved in this particular skill, as well as a consideration of the developmental point in the process of neurogenesis at which behavioral consequences would most likely emerge. Data on human cortical development (Conel 1939–1967) shows different patterns of growth across brain regions. Generally, according to Kolb & Fantie (1989), postnatal changes in dendritic complexity within speech areas (though simple at birth) are among the most impressive in the brain; a dramatic increase in the density of synaptic connections of the left temporal/parietal cortex begins at 8 and declines at 20 months of age. In the frontal cortex, the peak period begins a little

later; the increase in the density of synaptic connections occurs between 15 and 24 months of age, followed by a gradual decline.

To the extent that developmental changes in speech perception reflect tuning of a perceptual-motor system specialized to speech, it is likely that developmental changes in behavior will be related to neurobiological events in speech-related cortical areas. If this is the case, one would predict a tight relationship between developmental changes in speech perception and speech production. In previous writings (Werker & Pegg 1992), we have reviewed the data that show a relationship between developmental changes in babbling and in speech (e.g. Locke 1990).

At least some of the developmental changes in speech perception—particularly those showing a mapping to native-language phonological contrasts—might be related to changes in either cognitive mechanisms or to the construction of a phonological rule system (see Goldman-Rakic 1987, Werker 1991). These changes could be related to prefrontal cortical mechanisms. Goldman-Rakic (1987) reports a significant increase in synaptic proliferation in humans in the frontal and prefrontal cortex beginning around 8 months of age and continuing until about 2 years of age (see also Huttenlocher 1979). Lecours (1975) also suggests that intra- and interhemispheric cortical association bundles begin to myelinate at 7 months postnatally.

Before definitive links can be made, it is necessary to specify more precisely what kinds of changes in behavior might be related to specific synaptic proliferations and what kinds might be related to regressive events. Whether the initial reorganization and emergence of competences reflect a widespread *concurrent* period of excess synaptogenesis and subsequent pruning throughout the cortex, representing the elements of the distributed perceptual-motor network, or whether the neural unfolding is more piecemeal and hierarchical, remains uncertain. We have detailed the timing of important age-related changes in speech-related behavior. The differences in speech-related perceptual/motor competences that are environmentally influenced must, one way or another, yield changes in the dendritic fields of neurons within these neural systems. We are hopeful that the data and constructs reviewed in this chapter will help direct thinking and research toward further specification of the link between neural and behavioral events in human speech perception.

#### ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada (OGP0001103 to J. F. Werker and OGP000179 to R. C. Tees). The stimuli for our experiments were prepared

with assistance from NICHD Contract No. 1-HD-5-2910 to Haskins Laboratories.

#### Literature Cited

- Bertenthal, B. L., Campos, J. J. 1987. New directions in the study of early experience. *Child Dev.* 58: 560-67.
- Best, C. T., Hoffman, H., Glanville, B. B. 1982. Development of infant ear asymmetries for speech and music. *Percept. Psychophys.* 31: 75-85.
- Best, C. T., McRoberts, G., Sithole, N. 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *J. Exp. Psychol. Human Percept. Perform.* 14: 345-60.
- Biben, M., Symmes, D., Bernhards, D. 1989. Contour variables in vocal communication between squirrel monkey mothers and infants. *Dev. Psychobiol.* 22: 617-31.
- Conel, J. L. 1939-1967. *The Postnatal Development of the Human Cerebral Cortex*, Vol. 1-8. Cambridge, MA: Harvard Univ. Press.
- Cowan, W. M., Fawcett, J. W., O'Leary, D. M., Stanfield, B. B. 1984. Regressive events in neurogenesis. *Science* 225: 1258-65.
- DeCasper, A. J., Fifer, W. P. 1980. Of human bonding: Newborns prefer their mothers' voices. *Science* 208: 1174-76.
- Edelman, G. 1987. *Neural Darwinism: The Theory of Neuronal Group Selection*. New York: Harper & Row.
- Ehret, G. 1987. Categorical perception of sound signals: Facts and hypotheses from animal studies. In *Categorical Perception*, ed. S. Harnad, pp. 301-31. New York: Cambridge Univ. Press.
- Eimas, P. D., Miller, J. L. 1980. Contextual effects in infant speech perception. *Science* 109: 1140-41.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P., Vigorito, J. 1971. Speech perception in infants. *Science* 171: 303-6.
- Ferguson, C. A. 1964. Baby talk in six languages. *Amer. Anthropol.* 66: 103-14.
- Fernald, A. 1984. The perceptual and affective salience of mothers' speech to infants. In *The Origins and Growth of Communication*, ed. L. Feagans, C. Garvey, R. Golinkoff, pp. 5-29. Norwood, NJ: Ablex.
- Fifer, W. P., Moon, C. 1988. Auditory experience in the fetus. In *Behavior of the Fetus*, ed. W. Smotherman, S. Robertson, pp. 175-88. West Caldwell, NJ: Telford Press.
- Fowler, C. A., Rosenblum, L. D. 1991. The perception of phonetic gestures. In *Modularity and the Motor Theory of Speech Perception*, ed. I. G. Mattingly, M. Studert-Kennedy, pp. 33-59. Hillsdale, NJ: Erlbaum.
- Garcia, J., Koelling, R. 1966. Relation of cue to consequence in avoidance learning. *Psychonom. Sci.* 4: 123-24.
- Goldman-Rakic, P. S. 1987. Development of cortical circuitry and cognitive function. *Child. Dev.* 58: 601-22.
- Gottlieb, G. 1985a. On discovering significant acoustic dimensions of auditory stimulation for infants. In *Measurement of Audition and Vision in the First Year of Postnatal Life: A Methodological Overview*, ed. G. Gottlieb, N. A. Krasnegor, pp. 3-29. Norwood, NJ: Ablex.
- Gottlieb, G. 1985b. Development of species identification in ducklings: XI. Embryonic critical period for species-typical perception in the hatching. *Animal Behav.* 33: 225-33.
- Green, S. 1975. The variation of vocal pattern with social situation in the Japanese monkey (*Macaca fuscata*): A field study. In *Primate Behavior*, ed. L. Rosenblum, 4: 1-102. New York: Academic.
- Greenough, W. T., Black, J. E., Wallace, C. S. 1987. Experience and brain development. *Child Dev.* 58: 539-59.
- Hebb, D. O. 1980. *Essay on Mind*. Hillsdale, NJ: LEA Press.
- Heffner, H. E., Heffner, R. S. 1989. Effect of restricted cortical lesions on absolute thresholds and aphasia-like deficits in Japanese macaques. *Behav. Neurosci.* 103: 158-69.
- Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Wright Cassidy, K., Druss, B., et al. 1987. Clauses are perceptual units for young infants. *Cognition* 26: 268-86.
- Huttenlocher, P. R. 1979. Synaptic density in human frontal cortex—developmental changes and effects of aging. *Brain Res.* 163: 195-205.
- Johnston, T. D. 1988. Developmental explanation and the ontogeny of birdsong: Nature/nurture redux. *Behav. Brain Sci.* 11: 617-63.
- Jusczyk, P. W. 1991. Developing phonological categories from the speech signal. In *Phonological Development: Models, Research, and Implications*, ed. C. E. Fer-

- guson, L., Menn, C., Stoel-Gammon. Parkton, MD: York Press. In press.
- Jusczyk, P. W., Bertoni, J. 1988. Viewing the development of speech perception as an innately guided learning process. *Lang. Speech* 31: 217-38.
- Kimura, D. 1967. Functional asymmetry of the brain in dichotic listening. *Cortex* 8: 163-78.
- King, A. J., Moore, D. R. 1991. Plasticity of auditory maps in the brain. *Trends Neurosci.* 14: 31-37.
- Knudsen, E., Konishi, M. 1978. Space and frequency are represented separately in auditory midbrain of the owl. *J. Neurophysiol.* 41: 870-84.
- Kolb, B. 1989. Brain development plasticity in behavior. *Amer. Psychol.* 44: 1203-12.
- Kolb, B., Fantie, B. 1989. Development of the child's brain and behavior. In *Handbook of Clinical Child Neuropsychology*, ed. C. R. Reynolds, E. Fletcher-Janzen, pp. 17-39. New York: Plenum.
- Kolb, B., Whishaw, I. A. 1990. *Fundamentals of Human Neuropsychology*. New York: Freeman, 3rd ed.
- Kuhl, P. K. 1987. Perception of speech sounds in early infancy. In *Handbook of Infant Perception*, ed. P. Salapatek, L. Cohen, 2: 275-382. New York: Academic.
- Kuhl, P. K. 1988. Auditory perception and the evolution of speech. *Human Evol.* 3: 19-43.
- Kuhl, P. K., Meltzoff, A. N. 1982. The bimodal perception of speech in infancy. *Science* 218: 1138-44.
- Lecours, A. R. 1975. Myelogenetic correlates of the development of speech and language. In *Foundations of Language Development: A Multidisciplinary Approach*, ed. E. H. Lenneberg, E. Lenneberg, I: 121-35. New York: Academic.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., Studdert-Kennedy, M. 1967. Perception of the speech code. *Psychol. Rev.* 74: 431-61.
- Liberman, A. M., Mattingly, I. G. 1989. A specialization for speech perception. *Science* 243: 489-94.
- Locke, J. L. 1990. Structure and stimulation in the ontogeny of spoken language. *Dev. Psychobiol.* 23: 621-43.
- Logan, J. S., Lively, S. E., Pisoni, D. B. 1991. Training Japanese listeners to identify /r/ and /l/: A first report. *J. Acoust. Soc. Amer.* 89: 874-86.
- MacKain, K. S., Best, C. T., Strange, W. 1981. Categorical perceptions of English /r/ and /l/ by Japanese bilinguals. *Appl. Psycholinguist.* 2: 369-90.
- Marler, P. 1990. Innate learning preferences: Signals for communication. *Dev. Psychobiol.* 23: 557-68.
- Mattingly, I. G., Liberman, A. M. 1989. Speech and other auditory modules. In *Signal and Sense: Local and Global Order in Perceptual Maps*, ed. G. M. Edelman, W. E. Gall, W. W. Cowan, pp. 775-93. New York: Wiley.
- Mattingly, I. G., Studdert-Kennedy, M. 1991. *Modularity and the Motor Theory of Speech Perception*. Hillsdale, NJ: Erlbaum.
- McGurk, H., MacDonald, J. 1976. Hearing lips and seeing voices. *Nature* 264: 746-48.
- Mehler, J., Jusczyk, P. W., Lambert, G., Halsted, N., Bertoni, J., et al. 1988. A precursor of language acquisition in young infants. *Cognition* 29: 143-78.
- Merzenich, M. M., Nelson, R. J., Stryker, M. P., Cynader, M. S., Schoppmann, A., et al. 1984. Somatosensory cortical map changes following digit amputation in adult monkeys. *J. Comp. Neurol.* 224: 592-605.
- Miller, J. L., Jusczyk, P. W. 1989. Seeking the neurobiological bases of speech perception. *Cognition* 33: 111-37.
- Miller, J. L., Liberman, A. J. 1979. Some effects of later-occurring information on the perception of stop consonant and semivowel. *Percept. Psychophys.* 25: 457-65.
- Mitchell, D. E. 1989. Normal and abnormal visual development in kittens: Insights into the mechanisms that underlie visual perceptual development in humans. *Canad. J. Psychol.* 43: 141-64.
- Molfese, D. L., Molfese, V. J. 1980. Cortical responses of preterm infants to phonetic and nonphonetic speech stimuli. *Dev. Psychol.* 16: 574-81.
- Ojemann, G. 1983. Brain organization for language from the perspective of electrical stimulation mapping. *Behav. Brain Sci.* 6: 189-230.
- Perrett, D. I., Rolls, E. T., Caan, W. 1982. Visual neurones responsive to faces in the monkey temporal cortex. *Exp. Brain Res.* 47: 329-42.
- Petersen, M. R. 1982. The perception of species-specific vocalizations by primates: A conceptual framework. In *Primate Communication*, ed. C. Snowden, C. Brown, M. Petersen, pp. 171-211. Cambridge, UK: Cambridge Univ. Press.
- Petersen, M. R., Beecher, M. D., Zoloth, S. R., Green, S., Marler, P. R., et al. 1984. Neural lateralization of vocalizations by Japanese macaques: Communicative significance is more important than acoustic structure. *Behav. Neurosci.* 98: 779-90.
- Pisoni, D. B., Luce, P. A. 1986. Speech perception: Research, theory, and the principal issues. In *Pattern Recognition by Humans and Machines: Speech Perception*,

- Vol. 1, ed. I. C. Schwab, H. C. Nusbaum, pp. 1–50. New York: Academic.
- Poizner, H., Bellugi, U., Klima, E. S. 1990. Biological foundations of language: Clues from sign language. *Annu. Rev. Neurosci.* 13: 283–307.
- Polka, L. 1991. Cross language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *J. Acous. Soc. Amer.* 89: 2961–77.
- Sinnott, J. M. 1987. Modes of perceiving and processing information in birdsong (*Agelaius Phoeniceus*, *Molothrus ater*, and *Homo sapiens*). *J. Comp. Psychol.* 101: 355–66.
- Smotherman, W. P., Robinson, S. R. 1989. Cryptopsychobiology: The appearance, disappearance, and reappearance of a species-typical action pattern during early development. *Behav. Neurosci.* 101: 246–53.
- Strange, W., Jenkins, J. J. 1978. Linguistic experience and speech perception. In *Perception and Experience*, ed. R. D. Walk, H. L. Pick, pp. 125–69. New York: Plenum.
- Summerfield, Q. 1991. Visual perception of phonetic gestures. In *Modularity and the Motor Theory of Speech Perception*, ed. I. G. Mattingly, M. Studdert-Kennedy, pp. 117–30. Hillsdale, NJ: Erlbaum.
- Sussman, H. M. 1989. Neural coding of relational invariance in speech: Human language analogs to the barn owl. *Psychol. Rev.* 96: 631–42.
- Tees, R. C. 1990a. Plasticity and change. In *The Cerebral Cortex of the Rat*, ed. B. Kolb, R. C. Tees, pp. 475–81. Cambridge, MA: MIT Press.
- Tees, R. C. 1990b. Experience, perceptual competences, and rat cortex. In *The Cerebral Cortex of the Rat*, ed. B. Kolb, R. C. Tees, pp. 507–36. Cambridge, MA: MIT Press.
- Tees, R. C., Werker, J. F. 1984. Perceptual flexibility: Maintenance or recovery of the ability to discriminate non-native speech sounds. *Can. J. Psychol.* 38: 579–90.
- Tréhub, S. 1976. The discrimination of foreign speech contrasts by infants and adults. *Child Dev.* 47: 466–72.
- Turkewitz, G. 1988. A prenatal source for the development of hemispheric specialization. In *Brain Lateralization in Children: Developmental Implications*, ed. D. L. Molfese, S. J. Segalowitz, pp. 73–81. New York: Guildford.
- Van Lancker, D., Fromkin, V. A. 1973. Hemispheric specialization for pitch and "tone": Evidence from Thai. *J. Phonet.* 1: 101–9.
- Warren, R. M., Obusek, C. J., Farmer, R. M., Warren, R. P. 1969. Auditory sequence: Confusion of patterns other than speech and music. *Science* 196: 586–87.
- Werker, J. F. 1991. The ontogeny of speech perception. In *Modularity and the Motor Theory of Speech Perception*, ed. I. G. Mattingly, M. Studdert-Kennedy, pp. 91–110. Hillsdale, NJ: Erlbaum.
- Werker, J. F., Gilbert, J. H. V., Humphrey, K., Tees, R. C. 1981. Developmental aspects of cross-language speech perception. *Child Dev.* 52: 349–55.
- Werker, J. F., Lalonde, C. E. 1988. Cross-language speech perception: Initial capabilities and developmental change. *Dev. Psychol.* 24: 1–12.
- Werker, J. F., Logan, J. 1985. Cross-language evidence for three factors in speech perception. *Percept. Psychophys.* 37: 35–44.
- Werker, J. F., Pegg, J. E. 1992. Infant speech perception and phonological acquisition. In *Phonological Development: Models, Research, and Implications*, ed. C. E. Ferguson, L. Menn, C. Stoel-Gammon. Parkton, MD: York. In press.
- Werker, J. F., Tees, R. C. 1983. Developmental changes across childhood in the perception of non-native speech sounds. *Canad. J. Psychol.* 37: 278–86.
- Werker, J. F., Tees, R. C. 1984a. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behav. Dev.* 7: 49–63.
- Werker, J. F., Tees, R. C. 1984b. Phonemic and phonetic factors in adult cross-language speech perception. *J. Acous. Soc. Amer.* 75: 1866–78.
- West, M. J., King, A. 1988. Female visual displays affect the development of male song in the cowbird. *Nature* 334: 244–46.
- Whalen, D., Liberman, A. 1987. Speech perception takes precedence over nonspeech perception. *Science* 237: 169–71.
- Williams, H., Nottebohm, F. 1985. Auditory responses in avian vocal motor neurons: A motor theory for song perception in birds. *Science* 229: 279–82.