Development of Hearing. Part III. Postnatal Development

James E. Peck*

Abstract
As humans, we hear the way we do because of at least three major forces. The first is phylogeny, the evolutionary changes in the auditory system since its beginnings. The second is embryology, the formation of the system in each individual. Finally, there is the interaction between the biologically determined mechanism we are born with and the environment. This series of articles reviews each aspect in turn, so that we may have a fuller appreciation of how it is we come to hear the way we do. Part I examined how the auditory system evolved, and Part II outlined the sequence of ear formation in prenatal life. In this concluding article, we examine the development of auditory perception from when we first hear (in utero) through 12 months of age, which is when speech usually appears. Humans seem to come "programmed" to perform all the processes necessary to learn the language of their environment. Interaction with the environment hones our abilities to process sound and thereby optimizes language acquisition and production.

Key Words: Auditory development, auditory discrimination, auditory localization, auditory perception, experience, feature detectors

This article is the concluding segment to this three-part series on the development of hearing. In the first part, the evolution of the hearing apparatus was traced from the origins of life to the large-brained humans (Peck, 1994a). The second outlined the embryonic development of the various parts of the auditory mechanism (Peck, 1994b). This third part focuses on the development of hearing from the earliest auditory experience to the onset of speech.

Earliest Auditory Experience. This article necessarily overlaps with the preceding one on embryology because, in all likelihood, we experience sound before we are born. Indeed, the fascinating question arises: when do we first hear? As we saw in Part II, all auditory mechanisms seem ready at 6 to 7 months of gestation. What is the final event that allows us to hear? Is it mechanical, neural, chemical? Various structures have at one time or another been assigned the role of "trigger" of onset of cochlear function. Closer to the truth seems to be a "simultaneous and synchronous maturation of many mechanical and neural properties" (Rubel, 1985, p. 57). Moreover, various factors influence one another for most of pregnancy (see Rubel, 1978, 1985). For example, certain central auditory ganglia, deprived of input from the peripheral receptors, will not develop fully (Rubel, 1985). In short, there is probably no one final cause.

Clinical observations suggest that some very premature infants born at 25 weeks react to sound. Support for these observations comes from high-resolution ultrasound imaging. Birnholtz and Benacerraf (1983) studied response to vibroacoustic stimuli in human fetuses by ultrasound. Blink responses were first detectable at between 24 and 25 weeks and were consistent at 25 weeks. One may speculate, based on such observations and our knowledge of embryology, that fetuses are capable of experiencing sound by the seventh month of intrauterine life.

An equally intriguing question is what do we first hear? Since the baby's sound environment consists of fluid insulated by layers of tissues, outside sounds are less intense than inside ones. Sound levels inside the womb average 85 dB SPL (Walker et al, 1971). Regardless of origin, low frequencies travel more efficiently than high frequencies through such dense media, so that sounds in the 20- to 200-Hz range predominate...
(Walker et al, 1971). Even though external sounds are attenuated, specially above 800 Hz, nearby voices or loud sounds may reach the fetus (Armitage et al, 1980).

To study this matter, Armitage et al (1980) measured sound inside sheep. Hydrophones were placed inside intact amniotic sacs of pregnant ewes. A loudspeaker was strapped to but did not touch the ewe's skin. Comparing external and internal levels showed attenuation ranging from 16 to 37 dB with peak attenuation occurring just below 1000 Hz, although generally high frequencies were attenuated more than low frequencies. Further, the examiners could hear internal sounds, such as the mother's eating, drinking, ruminating, breathing, and muscular movement. They could also hear, but not always understand, outside conversation. Surprisingly, heart beat and blood flow were not perceptible. Despite this information, one must keep in mind that the sound spectrum inside a uterus may be different from that which arrives at a fetus's inner ear due to altered transmission characteristics of an outer and middle ear filled with liquid (Rubel, 1985). What such signals "sound like" to the fetus may not be how they would sound to a person listening by air conduction via earphones or a loudspeaker. Rather, fetal sound experience may be more akin to bone-conduction hearing.

In sum, the sounds we hear first are probably those that are strongest during the entire pregnancy (i.e., low-frequency components of internal sounds, principally the mother's vocalizations and physiologic noise). Fetal sound experience in general may help fine tune the developing auditory system (Fifer and Moon, 1988). More specifically, prenatal hearing of maternal voice may influence postnatal speech perception, an issue discussed below (Aslin et al, 1983; Fifer and Moon, 1988).

**Postnatal Hearing.** With birth comes a major shift from fluid- and tissue-conducted sound to air-conducted sound (somewhat analogous to amphibians shifting from an aquatic to a terrestrial environment, as we saw in Part I [Peck, 1994a]). The infant is suddenly exposed to a very different sound spectrum. Also with birth, the organism is much more accessible for investigation, both by electrophysiologic and behavioral techniques.

**Sensitivity.** Perhaps the most basic auditory function is the ability to detect acoustic energy; how well it is done is described as sensitivity. Among the electrophysiologic techniques to study sensitivity in infants, the most often utilized has been auditory brainstem response (ABR) audiometry. ABR results have shown that babies' thresholds for clicks are about 17 dB higher than those of adults (Hecox, 1975; Schuman-Galambos and Galambos, 1979). Also, the latency of the infant's brainstem response decreases from before birth until the second year of life. This decreasing latency reflects a maturing auditory system, attributed to the process of myelination (Starr et al, 1977).

It must be kept in mind, however, that the ABR is a neural response and that ABR testing is a measure of synchronous neural discharge. A synchronous neural response is not to be equated with hearing or with the ability of an organism to make meaningful interpretations about its acoustic environment (Rubel, 1978), a capacity without which a child cannot acquire auditory language.

Even more productive than electrophysiologic means have been behavioral techniques for learning about what and how babies hear. Merely through observation, one sees obvious reflexive responses (usually "startle") to intense sounds in term babies and in many "preemies." Regardless of level, broad-band sounds are more likely than narrow-band sound to elicit a response from infants. Speech, a broad-band sound, has been shown to be more response eliciting than a narrow band or discrete frequency, perhaps due to listening experience in utero (Eisenberg, 1976), as mentioned above in the discussion of what we first hear.

Responses to intermediate sound levels are subtle or absent and are highly influenced by physiologic states of hunger and fatigue. While it is true that one can observe behavioral changes to very low-level signals, even in newborns, these changes have traditionally been considered unreliable.

Recently, however, the observer-based psychoacoustic procedure (OPP) was developed for behavioral study of auditory ability of very young infants down to 3 months of age (Olsho et al, 1987). Under the OPP paradigm, no particular response is specified (e.g., head turn). Rather, any behavioral change is accepted (eye widening, activity suppression, sucking changes). Sounds are presented to the infant via loudspeakers. To control for observer bias, the observer wears earphones and hears all trials, whether genuine or no-signal control trials. Thus, the observer is ignorant of when signals are actually presented to the baby. The observer decides
whether a given behavior occurred as a result of a signal. If the observer is correct, visual reinforcement is supplied to the baby for whatever the behavior was. Such reinforcement also helps the observer to become a more accurate judge of infant responses. Using OPP, infants in the 2- to 5-week age range showed thresholds at 500 and 4000 Hz of about 55 dB and 60 dB SPL, respectively (Werner and Gillenwater, 1990).

Visual reinforcement audiometry (VRA) has been tremendously successful with most babies above about 5 months of age. Using VRA, Moore and Wilson (1978) showed that infants 6- and 12-months of age have quite low sensitivity thresholds. They plotted sound field thresholds of 15 to 20 dB SPL at 500, 1000, and 4000 Hz for infants of 6 to 12 months of age.

While infant sensitivity is very good, it still is not quite as good as that of adults. Trehub et al (1980) compared infants and adults (using VRA) for six octave-band noises from 200 to 10,000 Hz in sound field. In general, 6-month-old infants' thresholds were 20 to 25 dB poorer than those of adults. By 12 months, the difference was about 10 dB less.

The gap between child and adult sensitivity continues to close with age but at different rates for different frequencies (see review in Trehub et al, 1988, 1989). The most rapid improvement in sensitivity occurs in the region of 10,000 Hz, where adult levels are attained as early as 4 or 5 years of age. The next frequency region to reach maximal detectability is 20,000 Hz, which occurs by ages 6 to 8 years. Indeed, 7-year-old children have more sensitive hearing at 20,000 Hz than do adults; around age 10 years, children begin to “lose” hearing until adult levels are reached. For stimuli in the 2000- to 4000-Hz range, maximal sensitivity is achieved by 8 years, and finally the 400- to 1000-Hz range at 10 years of age.

The bases for these changes are not clear. Speculation has included mechanical changes in the outer, middle, or inner ear. For example, the smaller pinna and canal of children, as compared to adults, might enhance high-frequency detection. Also, earlier research suggested that children had wider critical bands than adults (Schneider and Trehub, 1981), but recent evidence shows that critical band width does not change with age (Schneider et al, 1990).

**Localization.** In addition to detecting soft sound, infants can localize sound. While all sound processing may have survival value, localization is probably one of the more basic auditory abilities. Virtually all organisms that hear can localize sound in space (Erulkar, 1972). This harkens back to the first paper in this set of three, which referred to the survival value of hearing in vertebrates (Peck, 1994a). Hearing is a distance sense; thus, the ability to detect a sound is even more helpful when the organism can tell where the sound is coming from (consider hunting or escape). A great deal has been learned about human infants' abilities to resolve auditory space.

As with most auditory abilities, resolution of auditory space improves with age. Improvement in localization accuracy among infants occurs first in the horizontal plane and then in the vertical plane. Newborns make head turns to the right or left of midline in an orienting response to sound (Clifton et al, 1981). The newborn places sounds in a long narrow band in front and just above horizontal with almost no vertical dimension. This unidimensional region “disappears” around 2 months and “reappears” in two dimensions at about age 4½ months with a wider horizontal and some vertical aspect. Also, infant response becomes brisker. For example, head-turn latency is 7 sec in newborns and decreases to about 2 sec by 5 months of age (Morrongiello and Clifton, 1984). These changes in localization are interpreted as a shift from reflexive behavior (subcortical) at birth to voluntary/controlled behavior (cortical) after the neonatal period (see review in Morrongiello and Gotowiec, 1990).

At 8 weeks of age, the smallest angle of change that can be detected (minimum audible angle or MAA) is about 27 degrees. Performance improves systematically, so that by 6 months, infants' MAA is about 14 degrees (Morrongiello et al, 1990). After 6 months of age, improvement in localization is slower but still systematic. The MAA is 8 degrees at 12 months and only 4 degrees at 18 months (Morrongiello, 1988b). (The difference is not due to limitation of head control; infants at 6 months can visually locate a light to within about 5 degrees [Morrongiello and Rocca, 1987a].)

Throughout life, our sense of auditory space is most keen directly in front. In other words, we can detect smaller right-left distinctions in front than changes in sound location within the hemifield on the right or left sides. In infancy, the MAA for sounds referenced to 60 degrees to the side is about 5 degrees poorer than for noting right-left shifts directly in front (Morrongiello and Rocca, 1990). As with other skills, localizing within hemifields improves with age. At 6
months of age, the MAA is 19 degrees, while at 12 months, it is 15 degrees (Morrongiello and Rocca, 1990).

One of the two main cues used in horizontal localization is time of arrival (the other being relative intensity at the two ears). When a sound is on one side, stimulation of the near ear precedes, by a fraction of a second, stimulation of the far ear. Auditory systems give precedence to the first ear stimulated and identify the sound as coming from the near-ear side. The effect is greatest in the low frequencies. Even if the first sound is followed by reverberations of this waveform, the auditory system regards only the first wavefront in deciding about sound location. An auditory illusion of sidedness, known as the precedence effect, can be created by producing a sound through two separated loudspeakers with very small differences in time of onset. A listener experiences the sound as coming from the side of the first loudspeaker.

The illusion of location from the precedence effect is strongly age related and can be exploited for studying infants’ processing of time cues for localizing. If a single-source sound occurs on a particular side, neonates can orient to that sound as well as 5-month-old infants. However, if a sound is presented through two separated loudspeakers with a 7-msec delay at one loudspeaker, neonates fail to orient, while 5-month-old babies can orient to the lead side (Clifton, 1985). By 6 months, infants consistently turn toward the side of the leading loudspeaker (Clifton et al, 1984). In turn, preschool children can localize smaller delays than infants, and adults can localize still smaller delays than preschoolers. Investigators speculate that improving precedence perception is related to progressing myelination of the cortex (see literature review in Morrongiello and Gotowiec, 1990).

As mentioned above, infants resolve auditory space more accurately in the horizontal than in the vertical plane. For example, at 6 and 12 months of age, infants have a larger angle of error by 1 or 2 degrees for vertical versus horizontal tasks (Morrongiello and Rocca, 1987b). This implies that there are different cues for sound location. The explanation seems to involve high-frequency reception and pinna shape. The sense of verticality is strongly influenced by high frequencies, and the pinnae strongly influence high-frequency reception (Morrongiello and Gotowiec, 1990). Perhaps the changing size and shape of the child’s pinnae help resolve the high-frequency cues of everyday sounds.

In short, at first, the infant can respond to gross laterality (right-left) changes; later, the infant is increasingly accurate in placing sound in space (Morrongiello and Rocca, 1987a; Morrongiello, 1988a).

Localization is not purely auditory; it subserves visual searching and finding (Morrongiello and Gotowiec, 1990). After 2 months, infants begin to integrate auditory and visual space (e.g., the mother’s voice and her facial movements coinciding in a unified space) (Muir, 1985). Also, Muir (1985) observed that a blind infant was less likely than a sighted child to localize at 0 to 6 months of age. It seems, then, that localization is an integrated, bimodal function with audition supporting vision in locating the source of a sound.

The phenomenon of localization holds relevance for the clinical audiologist in assessing hearing of very young children. Generally, such children do not tolerate earphones well, so their hearing is assessed in a sound field. The goal is to determine whether or not there is at least one normally hearing ear, since having a hearing loss in both ears will impair communication development. However, it is now clear that a long-standing hearing loss in even just one ear can have serious consequences, prompting the need to identify unilateral hearing deficits (Bess and Tharpe, 1986; Oyler et al, 1988). Testing an infant’s ability to localize sound can be helpful in suggesting the presence or absence of a unilateral hearing loss (Morrongiello, 1989; Auslander et al, 1991). See Figure 1 for examples of localization behaviors one might see in informal clinical assessment of infants.

**Recognition.** It seems a universal belief that very young babies recognize and prefer their own mother’s voice. Data support this popular belief. Neonates no older than 3 days have been taught that if they produced a particular sucking pattern, they could listen to their own mother’s voice. Even when the criterion to hear one’s mother’s voice was reversed, the neonates still “worked” to hear their mother’s voice rather than the voices of the other mothers (DeCasper and Fifer, 1980). Moreover, newborns are able to distinguish between their mother speaking in her native tongue versus speaking in an unfamiliar language, even when the signals have been low-pass filtered (Mehler et al, 1988). Perhaps babies are sensitized prenatally to respond to their mother’s voice (DeCasper and Fifer, 1980).
Newborn period to 4 mo.
Normal infant is aroused from sleep by sound signals of 90 dB (SPL) in a noisy environment, 50-70 dB (SPL) in quiet.

2 to 4 mo.
Normal infant begins to make a rudimentary head turn toward a sound signal 50-60 dB (SPL).

4 to 7 mo.
Baby turns head directly toward the side of a signal 40-50 dB (SPL) but cannot find it above or below.

7 to 9 mo.
Baby directly locates a sound source of 30-40 dB (SPL) to the side and indirectly below.

9 to 13 mo.
Baby directly locates a sound source of 25-35 dB (SPL) to the side and below.

13 to 16 mo.
Toddler localizes a direct sound signal of 25-30 dB (SPL) to the side and below, indirectly above.

16 to 21 mo.
Toddler localizes directly a sound signal of 25-30 dB (SPL) on the side, below and above.

21 to 24 mo.
Child sounds directly a sound signal of 25 dB (SPL) at all angles.

Figure 1 Illustrations of typical development of sound localization in infants. Left panel shows infant responses from newborn to 9 months. Right panel shows infant responses from 9 to 24 months. (Adapted with permission from Northern and Downs, 1991.)

Not only do babies recognize their own mother's voice, but they also recognize and prefer certain kinds of vocalizations. Interestingly, the vocal patterns they prefer are the sorts grown-ups use in speaking to babies (Aslin et al., 1983; Fernald, 1992). Indeed, infant responses probably reinforce our adult behavior.

Universally, when adults speak to infants, they use slow tempo, high pitch, and marked intonation changes. Adults seem to have an almost irresistible urge to address infants in this specific way. This speech style is called "motherese," even though it is used by males and females, parents and nonparents. Adults spontaneously use "motherese" because its features are highly salient auditory stimuli for recruiting infants' attention (Fernald, 1984, 1992).

It is clear that young babies can discern the acoustic parameters that typify infant-directed speech. Spring and Dale (1977) demonstrated that young infants in the 1- to 4-month age range can detect acoustic cues that serve to mark intonation (e.g., changes in fundamental frequency, intensity, and duration of voice). The basis for these phenomena is undetermined, but speculation includes intrauterine exposure to such sounds and the attention-getting property of variable sounds. Apparently, infants arrive ready to be responsive to those acoustic characteristics of the "motherese" signal (Fernald, 1992).

"Motherese" is not merely a social interaction to delight adults and babies alike. Rather, it plays an important role in caregiver-infant bonding. From the outset, mothers use their voice to arouse their babies, to attract and maintain their attention, and to soothe and warn them. In short, a mother's vocalizations communicate affect, that is, her feelings and mood. Later in infancy, as a result of experience, the vocalizations take on additional meaning as "signs" of events around them. Finally, the infant attributes specific significance to speech and arrives at linguistic meaning. Concomitantly, the infant uses sound initially to convey affect and ultimately to express lexical meaning through words (Fernald, 1984).

Summarizing this section, there is a mutual interaction at work here. On the one hand, adults spontaneously speak to infants in a highly stylized way. On the other, infants prefer those vocal patterns.
Why should humans engage in such elaborate and ritualistic vocal-auditory behavior? Taking a highly Darwinian approach, Fernald (1992) argues persuasively that infant-directed speech is biologically determined through evolution by natural selection. The basis, she claims, is the lengthening period of dependency in primates. Humans are the most neotenous of all, born the least developed and the most helpless. Prolonged immaturity among hominids favored certain parenting behaviors. Among them was adopting vocal behaviors that were suited to the offsprings' limited auditory perceptual systems, thereby promoting attention and bonding.

Effective, early mother-infant communication had adaptive advantages for hominids extending far past infancy, reasons Fernald (1992). The quality of attachment in infancy impacts on social competence in childhood and on emotional relationships and parenting in adulthood. Good nurturing produces a socially competent member of the group and an effective parent. The nurtured becomes the nurturer. In keeping with natural selection, such behaviors improve the chances of survival of both the individual and of the group and thereby tend to be passed along to subsequent members (Fernald, 1992).

**Discrimination.** It has been known for 30 years that not only can neonates respond to sound, but that they can also discriminate among different sounds (Papousek, 1961; Siqueoland and Lipsitt, 1966). We saw above that neonates can discriminate between their mother's voice and the voices of other mothers. In a classic study of speech sound discrimination, Eimas et al. (1971) showed that infants aged 1 to 4 months could distinguish between voiced and voiceless sounds of /b/p and /d/t. Infants were presented synthetic syllables, which varied in voice onset time (VOT), the main cue for voiced-voiceless distinction. The high amplitude sucking paradigm was used, based on the phenomenon that when infants are presented a novel stimulus, their sucking rate increases suddenly. Those infants listening to syllables with the same VOT did not alter sucking rate, whereas those infants presented with syllables of differing VOTs consistently accelerated sucking rates.

Since the early investigations, a considerable body of research has emerged showing that infants in the first year of life have the ability to make virtually all of the auditory discriminations necessary to learn language (see Kuhl, 1987). They have also shown that infants can discriminate other aspects of speech, such as time and pitch changes. For example, Olsho (1984) compared difference limens for frequency (DLF) of babies 5 to 8 months of age with those of adults. Tones were octave points 250 to 8000 Hz. For the infants, she conditioned head-turn responses to frequency changes. For both groups, relative frequency discrimination limens (change in frequency/frequency) at 2000 Hz and above were similar at about 1 percent (i.e., detected a change of as few as 20 cycles per second in a 2000-Hz signal). For frequencies below 2000 Hz, infants' DLF was about 1 to 3 percent larger than those of adults. For adults, the DLFs at 250, 500, and 1000 Hz were about 2 percent, 1.5 percent, and 0.4 percent, respectively. In contrast, infants' DLFs were approximately 3.5 percent, 2.5 percent, and 1 percent for the same tones. As other examples of discrimination ability, 6-month-old infants can detect small temporal changes on the order of 20 msec (Morrongiello and Trehub, 1987), and 9- to 11-month-old infants can detect changes in melodic contour in different contexts (Trehub et al., 1987).

Being able to detect time and pitch changes (i.e., suprasegmental features) enhances speech discrimination. Karzon (1985) studied the ability of 1- to 4-month-old infants to discriminate between three-syllable utterances contrasted by one phoneme. Vocal stress was varied by intensity, duration, and fundamental frequency. Karzon (1985) found increased ability to discriminate when the target syllable was exaggerated in the highly intonated manner in which speech is directed to infants. Perhaps prenatal listening to low-frequency sound heightened their response to prosody.

**Categorization.** Languages have categories of speech sounds called phonemes. The ability to perceive speech sounds as belonging to categories denotes the ability to detect a difference between stimuli, even if varying in only one dimension (see Kuhl, 1987). An example is noting the difference in VOT that distinguishes /p/ from /b/. Numerous studies since the early 1970s demonstrate that infants group speech sounds into phonetic categories in adult-like fashion (see Kuhl, 1987).

However, categorizing speech sounds requires not only noting differences but also similarities. This means that the listener must treat sounds that are discriminatively different as equivalent, a phenomenon known as equivalence classification. Equivalence classification
Innate Abilities and Special Mechanisms.

With infants exhibiting such extensive speech "processing" abilities, investigators, like Eimas et al. (1971), hypothesized that human infants were born with special auditory processors, which he dubbed feature detectors. In other words, it was felt that babies came ready to make linguistically relevant auditory discriminations even before they had been exposed to them. It was as if babies' auditory systems had a "speech mode" for listening to speech. Two related hypotheses surfaced that this speech mode was species specific—only humans had it—and that it was speech specific (i.e., the ability evolved specifically to perceive speech). However, research in the 1970s called into question the speech-specific idea that humans were programmed to process speech sounds. Studies showed that adults made categorical distinctions between nonspeech sounds much as they did with speech sounds (Cutting and Rosner, 1974; Pisoni, 1977). (Here, the term nonspeech sounds means sounds that mimic some property or properties of speech without being identified as speech.) Later studies involving infants found that they, also, could make fine discriminations between nonspeech sounds (Jusczyk et al., 1977, 1980; Clarkson and Clifton, 1985).

By the late 1970s, the species-specific notion that we humans are the only species capable of such exacting auditory behavior had also been seriously weakened. Reports appeared of various mammals making many categorical speech sound distinctions (Kuhl, 1978). With both the speech-specific and species-specific notions in doubt, Kuhl (1986) saw no need to invoke "special mechanisms" to explain categorical perception or speech sound discrimination. On the contrary, she hypothesized that as speech evolved, certain "phonetic" sounds were selected for production because they possessed characteristics that the auditory system could detect easily. Thus, Kuhl argued, infant speech perception was based on certain properties of the mammalian auditory system, which the human baby draws upon in processing speech.

Recent evidence from Kuhl herself suggests that, indeed, there may be a degree of species-specific speech processing ability after all. Kuhl (1991) demonstrated that some speech sounds are more typical of a given speech sound category than others. In other words, there is an ideal sound for each category, or a "prototype." This typicality strongly influences speech perception. Kuhl found, for both human adults and 6-month-old infants, that the closer a sound is to the prototype, the easier it is to recognize it as a member of that category. However, in testing monkeys, Kuhl found that they did not show a difference in discriminating speech sounds based on prototype or nonprototype sounds. She concluded that human adults and infants have an auditory image of what constitutes the center of a speech sound category.

How humans come to have this internal auditory map of categories is uncertain. Kuhl (1991) suggests two possibilities. First, infants may be biologically endowed with these images of prototypes. Alternatively, experience may have a strong determining force, even by 6 months of age. Finally, it could be that there is a mix of these two factors. Whichever the case, the critical point here is that there seems to be at least some species-specific ability to process speech. In the larger context of this series of papers, the relevance of the preceding paragraphs is that our remarkable perceptual abilities, however special or general, facilitate the acquisition of language.

Experience. As impressive as our inborn perceptual capabilities may be, they are not enough for postnatal development of all auditory skills and speech production. Experience seems critical in refining and maintaining speech perception. As an obvious example, profound congenital deafness, without intervention, imposes a nearly insurmountable difficulty in acquiring aural-oral communication. The biologic systems underlying auditory perception cannot fulfill their roles unless sound is at least perceptible. As we have seen, experience begins prenatally. Indeed, experience is critical for proper neurogenesis; it finetunes our perceptual-motor networks to those characteristics of human voice that are available to the fetus (Werker and Tees, 1992). Prenatal experience shapes postnatal
sound preference and enhances the ability to process sound in general and speech in particular. In other terms, speech perception has an endogenous basis with an experiential overlay (Werker and Tees, 1992).

Despite general agreement that interaction with our environment is important in communication development, the way and degree to which it is important are not at all clear. Views run the range of extremes from experience being absolutely crucial to being completely unnecessary for each aspect of development. Perhaps the role of environment is multidimensional and involves several parallel processes at work simultaneously (Aslin et al, 1983; Werker, 1989). Figure 2 shows five hypothetical roles of experience on development (see Werker, 1989).

In maturation, the first process shown, the environment has no role at all in development. Accordingly, certain perceptual skills emerge "on their own" regardless of experience. In a second process, perceptual skills are completely ready at birth, but experience is necessary to maintain the perceptual skills. Some would offer the above case of deafness as an example of the maintenance/loss process. It implies that skills are fully developed but will not be maintained unless sounds of language are experienced by the individual. Next, experience facilitates perceptual development. Facilitation does not alter the final level of ability but rather it speeds the rate of reaching that level. Fourth, experience attunes the perceptual mechanisms to become sharper in their auditory discriminations. Finally, experience induces the skills. In the induction hypothesis, skills are latent but undeveloped, and their appearance is wholly dependent upon environmental exposure. This process is similar to imprinting. A well-known example of imprinting is ducklings developing a strong and instant recognition of the first sound they hear—usually their mother's call—to the exclusion of later-heard animal calls.

In at least one aspect, experience can seem, paradoxically, to diminish rather than aid auditory skills. Werker (1989) and her colleagues tested persons from infancy to adulthood in their ability to discriminate sound contrasts that occur in their language and sound contrasts that occur in other languages. Not unexpectedly, they found that adults can distinguish between phones used in their language but not between phones used in an unfamiliar language. In contrast, infants can distinguish between phones in a language they have never heard. It seems that infants are born with a universal phonetic sensitivity, which fades between ages 6 and 12 months.

These initial findings had been interpreted as an example of maintenance/loss in experience. However, later investigations showed that, given considerable training, adults could regain some ability to make these distinctions (Werker, 1989; Werker and Tees, 1992). In other words, their capacity to make those distinctions is not lost entirely.

Werker (1989) suggests that decreased phonetic sensitivity at 10 to 12 months of age is not a decline in auditory skill but rather a language-based reorganization of phone categories. She sees a parallel in oral language development. Babies seem to be able to make all sounds of any language, but as soon as spoken language emerges, vocal productions narrow to approximate the sounds of that language. In other words, phonetic sensitivity gives way to language sensitivity. Werker (1989) considers these changes as perceptual and productive streamlining, by limiting perception and production to categories useful for communicating
in the language of one's environment and ignoring the rest.

Data on the overall role of experience in auditory development are limited. Part of the problem is the familiar “nature versus nurture” controversy. Resolving it in humans is difficult because it is not ethically possible to perform the studies, which would isolate the critical variables. For the foreseeable future, we may need to be content with the notion that experience is crucial in the development of audition and communication, but to what extent is conjectural.

**SUMMARY**

This series of three articles has given an overview of the development of hearing. The papers addressed in turn evolution, embryology, and postnatal development of hearing. While much has been learned, “the mechanisms underlying auditory development are largely unknown” (Aslin et al, 1983, p. 669). Here is a brief synopsis of what is currently believed about auditory development. Our auditory system has evolved over a span of nearly 500 million years. The human ear’s internal fluid-filled labyrinth has its origin in an equilibril organ in fish. This early internal balance organ was adapted for hearing separately by fish, amphibia, and reptiles. Amphibia adapted their auditory organ for air-conduction hearing on land by coupling it to the outside world with a “middle ear” mechanism. The middle ear plan, a large membrane mechanically coupled to the inner ear, became part of the genetic heritage of later vertebrates. Early reptiles evolved, perhaps in more than one line, an auditory labyrinth from an equilibril one. The auditory organ of at least two reptilian species served as the basis for the auditory organ of later mammals and birds. Mammals developed two distinct types of auditory sensory cells, the outer and inner hair cells. Eventually, through primate and hominid development, humans came to possess an exquisite, large brain with which to process the neural codes from the peripheral receptor system.

In the case of individual humans, the embryo rapidly develops an intricate auditory system according to genetic instructions inherited over the ages. The inner ear has its beginnings at the end of the third gestational week, when cells in tissue from the outer layer of the embryo fold upon themselves, forming a fluid-filled cyst. By 10 weeks, the cyst has transformed into a cochlea of adult shape and by 5 months it attains adult size. Between 3 and 5 months, some of the cells lining the cyst specialize to become a mature organ of Corti. The combined sensory and neural systems are ready to respond to sound early in the seventh month. The middle ear originates in the fourth week as outpouchings of the embryonic pharynx. Ossicles form during months 2 through 4, and the middle ear space is open (but fluid filled) near the end of 7 months. Finally, the ear canal and outer ear originate in the fourth and fifth weeks, respectively. Tissue from six tiny bulges blend to form the pinna into its familiar shape by about 5 months. The ear canal deepens and widens and terminates in a thin tympanic membrane in the seventh month.

By birth, we humans possess biologic systems, some ready and others soon to be ready, to process complex sounds in a linguistically relevant manner. Our biology is influenced by exposure to sound, particularly the sounds and content of language. A superb central nervous system detects, discriminates, recognizes, categorizes, sequences, and stores sounds and, in a way surpassing that in all other species, extracts meaning from those sounds. Ultimately, near 12 months of age, we lay claim to perhaps our most human characteristic and begin to speak.

**Acknowledgment.** I am grateful to Drs. David Lipscomb, Noel D. Matkin, and Robert Oyler for their editorial help in preparing this series of papers. Thanks to Drs. Barbara A. Morrongiello, Sandra E. Trehub, and Janet F. Werker for their copious and generous suggestions for this third paper. Particular appreciation is given to L. Judson Farmer, friend and colleague, who suggested I undertake this project and urged its completion.

**REFERENCES**


Clifton RK. (1985). The precedence effect: its implications for developmental questions. In: Trehub SE,
opmental look at an auditory illusion: the precedence
Newborns’ orientation toward sound: possible implications
categorical look at an auditory illusion: the precedence


