Chapter 4

Exploring Developmental Changes in Cross-Language Speech Perception

Janet F. Werker

One of the most remarkable attributes of being human is the ability to acquire language, and considerable research in cognitive science has been directed toward understanding the processes and capabilities that make this feat possible. In order to acquire a spoken language, the listener must be able to parse the continuous stream of speech into sentences, clauses, words, syllables, and phonemes and to represent the information at each level. A phoneme is a subsyllabic unit, and the difference in a single phoneme can convey different meanings (for example, *big* versus *dig* or *bat* versus *bet*).

There are differences between languages in the phoneme inventory. For instance, the distinction between /r/ and /l/ is important in English: the word "rate" means something quite different than the word "late"; however, this difference is not phonemic—it does not signal a difference in meaning—in Japanese. Similarly, the difference between a dental /t/ and retroflex /T/ is not phonemic in English, although it can signal a difference in meaning in Hindi. Thus, for an English-learning child, the word *table* refers to a single class of items irrespective of whether the /t/ is produced with the tongue behind the front teeth (dental), on the ridge behind the teeth (alveolar), or with the underside of the tongue tip against the roof of the mouth (retroflex). For a Hindi-learning child, though, the dental and retroflex /t'/ are both phonemes and thus can signal different words. Clearly, then, English or Hindi children—like all children in the world—require some working knowledge of the phonology (the sound structure) of their native language in order to map sound on to meaning.

The focus of this chapter is to describe the quest by psychologists and linguists to understand the mechanisms and processes by which children acquire sensitivity to the sound structure, particularly the phoneme inventory, of their native language. The chapter will begin with a brief review of the remarkable speech perception abilities of young infants. This will be followed by a description of the research showing how these initial capacities are shaped and influenced by exposure to different languages. Finally,

some of the issues raised in attempting to explain these findings will be discussed. In reading this chapter, you will find that by the end of the first year of life, long before many infants have even acquired their first word, speech perception capacities have been modified to match the properties of the sound structure of the native language.

4.1 Speech Perception Sensitivities in Very Young Infants

In 1971 Eimas et al. published an influential paper in *Science* showing that infants 1–4 months of age are able to discriminate computer-synthesized examples of the syllables /ba/ and /pa/, which differ in only the initial phoneme. Eimas and colleagues also tested infants on their ability to discriminate various instances of the syllable /ba/ that differed by exactly the same amount as the /ba/ versus /pa/ stimuli. This clarified whether infants are exquisitely sensitive to any small acoustic differences or whether they are most sensitive to those differences that have been used by languages of the world to distinguish meaning. The infants were unable to discriminate these equal-sized acoustic differences from within the /ba/ category. Subsequent studies by the original authors and several other researchers confirmed this basic finding. Infants are able to discriminate nearly every phonemic contrast on which they have been tested, but they typically fail to discriminate equal-sized acoustic differences that specify two variations of the same phoneme.

This ability to discriminate phonemic contrasts extends to syllables the infant has never before heard. For example, young Kikuyu-learning infants can discriminate phonemic contrasts that are used in English but not in Kikuyu (Streeter 1975), and young Japanese-learning infants can discriminate contrasts that are used in English but not in Japanese (Tsushima et al. 1994). Young English-learning infants can discriminate consonant contrasts that are used in Czech (Trehub 1976), Hindi (Werker et al. 1981), Nthlakampx (Werker and Tees 1984a), Spanish (Aslin et al. 1981), and Zulu (Best, McRoberts, and Sithole 1988), to name but a few. They also seem to discriminate nonnative vowel contrasts (Trehub 1973; Polka and Werker 1994). These results suggest that rather than having to learn about important differences in speech, infants are equipped—undoubtedly through evolution—with sensitivities that make them "universal listeners," ready to learn any of the world's languages.[1]

In contrast to these language-general sensitivities shown by young infants, under many circumstances adults have difficulty discriminating

1. See Werker and Tees (1992) for a discussion of how evolution propels epigenetic processes.

syllables differing by only a single phoneme if that particular phonemic contrast is not used in their native language. Without training, Japanese adults have difficulty discriminating the difference between /ra/ and /la/ (Strange and Jenkins 1978), English adults have difficulty discriminating certain Hindi (Werker et al. 1981), Nthlakampx (Werker and Tees 1984a), and Czech (Trehub 1976) contrasts, and may even need short familiarization periods to discriminate acoustically quite salient nonnative distinctions (Pisoni et al. 1982; Werker and Logan 1985).

4.2 Early Effects of the Language Environment

If infants are "universal listeners" but adults sometimes have difficulty discriminating nonnative phonemic contrasts, then there must be a decline across age in cross-language speech perception performance. This contradicts the typical expectation that there will be age-related increases in ability and age-related improvements in performance. To see if the pattern of broad-based abilities in infancy and subsequent decline is accurate, a series of studies tested infants and adults on the same nonnative contrasts in the same procedure.

In an initial study Werker et al. (1981) compared English-learning infants aged 6–8 months with both English-speaking adults and Hindi-speaking adults on their ability to discriminate the English (and Hindi) /ba/-/da/ phoneme distinction plus two (non-English) Hindi contrasts selected to vary on their potential difficulty. The Hindi retroflex versus dental contrast, /Ta/-/ta/, involves two consonants that vary on the precise location of the tongue. The Hindi breathy voiced and voiceless unaspirated contrast, /t^ha/-/d^ha/ involves a difference in the timing and shape of the opening of the vocal cords upon release of the consonant. Significantly, neither of these Hindi phoneme contrasts is used to distinguish meaning in English.

All stimuli were produced by a native Hindi speaker who was also a trained linguist. Several exemplars of each phoneme were recorded, and eight from each category were selected. The final stimuli that were selected were uniformly rated as excellent exemplars by native speakers. Also, care was taken to ensure that the exemplars specifying each contrast (such as the /t^ha/'s and the /d^ha/'s) were overlapping in intensity, duration, fundamental frequency, and intonation contour so that listeners could not use these nonphonetic acoustic cues to distinguish the contrasts.

A method of testing was adopted that can be implemented in very similar forms with infants (5 1/2 months or older), children, and adults to ensure that subjects of different ages are tested in an equally sensitive procedure. The procedure used with infants is called the Conditioned Head

Turn, and the adult procedure is a close variate. Basically, this is a category change task in which the subject has to monitor a continuous background of syllables from one phonetic category (for example, /ba/), and signal when the stimuli change to a contrasting phonetic category (for example, /da/). Infants are conditioned to turn their head toward the speaker when a change is presented. Correct head turns are reinforced with interesting, lighted toys (such as drumming bears) and verbal praise from the experimenter. Adults and children signal detection of this change by pressing a button. Their correct responses are reinforced with a flashing light. Incorrect responses are not reinforced. A picture of an infant performing in the Conditioned Head Turn procedure is shown in figure 4.1.

The results from this early study comparing Hindi-speaking adults, English-speaking adults, and English-learning infants were consistent with the prediction of language-universal infant sensitivities and subsequent decline. Virtually all subjects in all groups could discriminate the /ba/-/da/ contrast—a distinction that is common across the world's languages and used in both English and Hindi. The six–eight-month-old English-learning infants and the Hindi-speaking adults could easily discriminate both Hindi

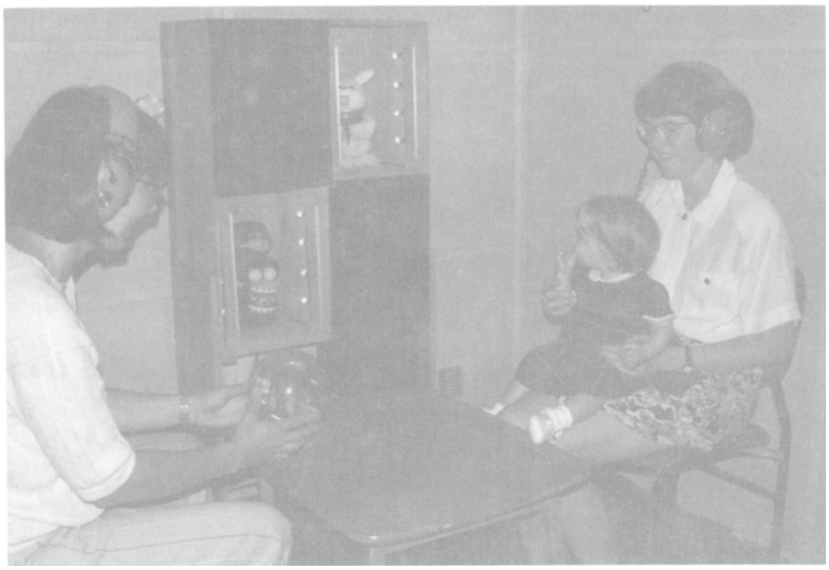

Figure 4.1
Madeleine performing in the Conditioned Head Turn Procedure. In this case, she has just successfully turned her head upon detecting a change in the speech sound stimulus. Her correct head turn is being reinforced with the lighting and activation of the toy animal, and with praise from the experimenter. (Photo courtesy Steve Heine)

contrasts. The English-speaking adults, however, had trouble discriminating the Hindi contrasts, showing particular trouble with the more difficult retroflex/dental place-of-articulation distinction (Werker et al. 1981). These results are summarized in figure 4.2.

To ascertain the age at which language experience first influences phonetic perception, Werker and Tees began a series of experiments testing children of different ages. First, ages were selected to test Lenneberg's hypothesis of a critical period for language acquisition (Lenneberg 1967). Lenneberg argued for a critical period for acquiring spoken language ending around the onset of puberty on the basis of data from children acquiring a second language and from reacquisition of spoken language after brain injury at different ages. His writings were so influential that it was reasoned that a similar critical period might apply to language perception. To test for this critical period, children of 12 and 8 years were tested, with the expectation that the 8-year-olds but not the 12-year-olds would be able to discriminate nonnative contrasts. English-speaking children of both ages, however, performed like English-speaking adults and were unable to discriminate the Hindi non-English phonemes. The study was extended to 4-year-old children, who actually performed most poorly of

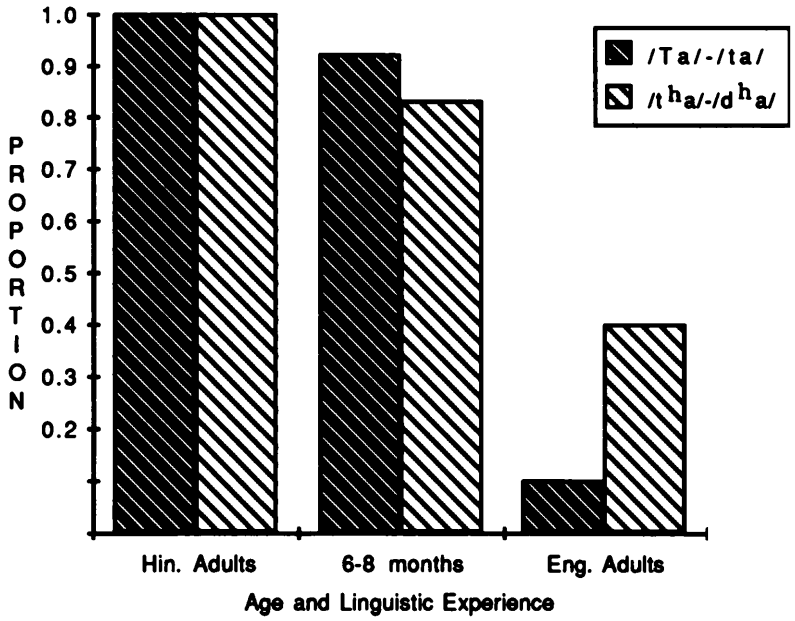

Figure 4.2
Proportion of subjects reaching criterion as a function of age and language contrast. Adapted from Werker et al. 1981.

all on the nonnative contrasts. Importantly, their poor performance was not due to task difficulties as they performed well on an English /ba/-/da/ distinction, and as 4-year-old Hindi-speaking children performed well on the Hindi contrasts (Werker and Tees 1983). These findings revealed that experience must begin to influence speech perception long before age 4, certainly well before the critical period suggested by Lenneberg.

In 1984 Werker and Tees continued a series of experiments designed to test even younger children to establish the age at which the decline is first evident. English-learning infants of 6–8, 8–10, and 10–12 months of age were tested on their ability to discriminate the Hindi (/Ta/-/ta/) contrast and the Nthlakampx /k'i/-/q'i/ contrast. Almost all the infants aged 6–8 months could discriminate both non-English contrasts, but the pattern was entirely different for older infants. As shown in the top part of figure 4.3, very few of the infants aged 10–12 months could discriminate either the Hindi or the Nthlakampx contrast (Werker and Tees 1984a). The infants aged 8–10 months showed an intermediate pattern of performance—with about half discriminating and half not discriminating the nonnative contrasts. As shown in the lower part of figure 4.3, the same basic pattern was revealed in a longitudinal study in which a small group of infants (six) were retested at two-month intervals (Werker and Tees 1984a) from ages 7 through 11 months. Finally, to ensure that the effect was due to experience with the language, a small number of Hindi and Nthlakampx-learning infants aged 11–12 months were tested and were found to be able to discriminate their native contrasts.[2]

This pattern of change between 6 and 12 months of age has been reported for a different retroflex/dental distinction (/Da/-/da/) (Werker and Lalonde 1988) and for three Zulu contrasts: a bilabial plosive/implosive distinction, a lateral fricative voiced/voiceless contrast, and a velar voiceless/ejective stop distinction (Best 1994). The change for the Nthlakampx contrast has also been replicated by Best et al. (in press). In a comparable series of experiments Tsushima et al. (1994) showed an analogous pattern for Japanese-learning infants. The Japanese infants of 6–8, but not 10–12 months of age can discriminate the English (non-Japanese) /ra/-/la/ distinction. Finally, although there is some evidence that language specific influences might begin to influence vowel perception at an

2. Before being tested on the non-English contrasts, the infants were required first to show that they could perform in the Head Turn procedure on the English /ba/-/da/ distinction. All infants were then given twenty-five trials on which to reach discrimination criterion on a non-English contrast. Before concluding that any infant who failed to reach criterion in that number of trials could really *not* discriminate a nonnative contrast and had not just lost interest in the procedure, all such infants were subsequently retested on the English /ba/-/da/ distinction. The data were only retained as meaningful from infants who passed the /ba/-/da/ test both before and after failing the nonnative contrast.

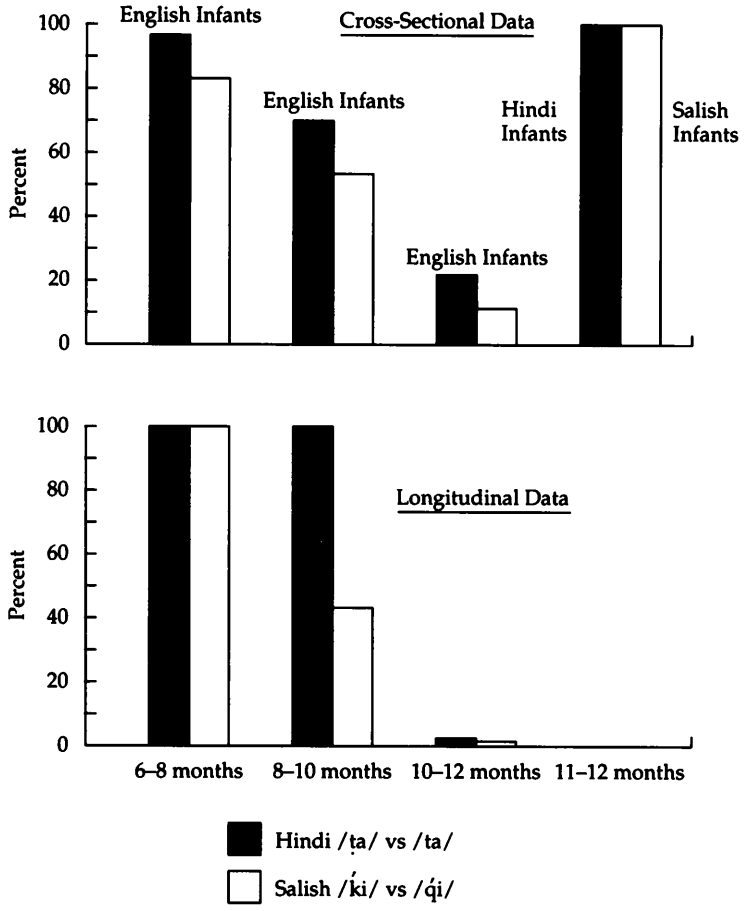

Figure 4.3
Proportion of infant subjects from three age groups and various language backgrounds reaching criterion on the Hindi and Salish contrasts. From Werker and Tees 1984a, 61.

earlier age—by 6 months—than they do consonant perception (Kuhl et al. 1992; Polka and Werker 1994), there is evidence of further change between 6 and 12 months of age, by which time infants' performance on nonnative vowel contrasts declines to the level of their performance on nonnative consonant contrasts (Polka and Werker 1994).

What can we conclude from these findings? First, infants stop being "universal listeners" for phoneme distinctions by the end of the first year of life. Second, it appears that the role of experience is to "maintain" those perceptual sensitivities that are already evident in the young infant. Without such exposure, initial abilities will be lost. The classic example of loss of perceptual ability following lack of exposure is the demonstration by

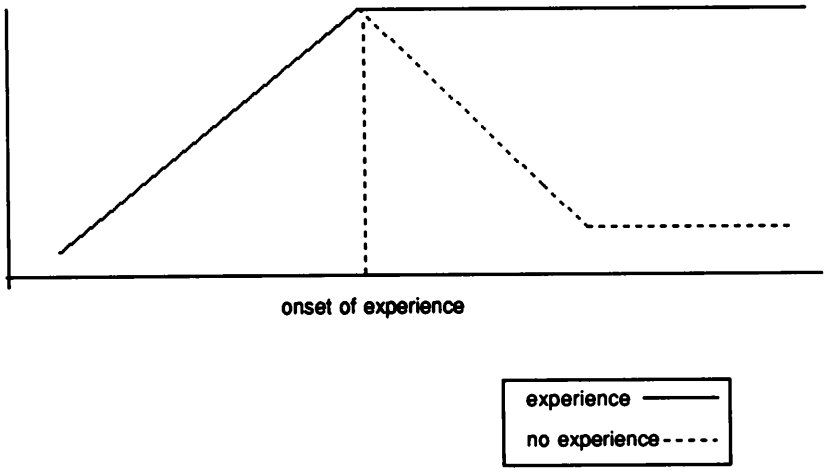

Figure 4.4
A representation of the Maintenance/loss model of experiential influences on perceptual development. Adapted from Gottlieb 1981.

Hubel and Wiesel that some cells in the visual cortex are designed to respond quite selectively to certain kinds of input. On the basis of early data, it was believed that if deprived of the appropriate input (such as vertical lines) for a sufficiently long time during a critical or sensitive period of development (for example, Wiesel and Hubel 1965), these cells atrophy. Irrespective of what happens later in life, these cells can no longer fire to the kind of information they were designed to detect (see Tees 1986 for a critique). Extrapolating to the field of speech perception, it was similarly believed that lack of exposure for a criterial amount of time should lead to permanent loss of the ability to discriminate nonnative phonemic contrasts.

Maintenance/loss is only one of the many possible roles experience might play in perceptual development (see Gottlieb 1981 for other possible roles). The Maintenance/loss model is shown in figure 4.4. The central assumption of this model when applied to speech perception is that the neonate is equipped with sensitivity to all possible phonemes of all languages. The role of experience is to maintain sensitivity to only those phonemes that the infants hears. Sensitivity to nonexperienced (nonnative) contrasts declines and/or disappears.

4.3 Problems for the Maintenance View

The data reviewed above are most consistent with a Maintenance role of experience, but there are fundamental problems with this view. First, the

Maintenance/loss model suggests that declines in sensitivity should follow lack of exposure and should be permanent and absolute. Although adults perform more poorly than infants on nonnative phoneme contrasts, there are conditions under which continuing adult sensitivity can be demonstrated. In one of the more intriguing demonstrations, it was shown that if the critical acoustic information in the speech contrast is presented alone so that the syllables no longer sound like speech, adults can discriminate nonnative contrasts. To illustrate, Werker and Tees (1984b) presented adult English speakers with either the ejective portion alone from the Nthlakampx /k'i/-/q'i/ contrast (telling listeners that it was water dropping into a bucket and that they should signal when the bucket was switched) or a truncated portion of the Hindi retroflex/dental (/Ta/-/ta/) contrast. In each case adult English speakers discriminated the shortened pairs with ease, but they still failed to distinguish the full syllables even when tested on them immediately after being presented with the shortened portions. Furthermore, adults can be "taught" to discriminate the full syllables if given enough training trials, or if tested in sensitive procedures with low memory demands (Pisoni et al. 1982; Werker and Logan 1985). Thus, it appears that the age-related decline in performance is evident only for stimuli that sound speechlike, and only in testing contexts that have processing demands similar to the memory and uncertainty demands required in everyday language processing. The age-related decline is not permanent and absolute, as would be predicted by a Maintenance/loss view.

A second problem for the Maintenance view is that neither the ages nor the form of the findings maps on to that predicted. According to a Maintenance model, one would predict that total duration of "misuse" would best predict poor performance—or, at least, that performance would decline to some "floor" level at which it would remain. It would be incompatible with a Maintenance model to expect a decline and then a recovery. Recall, however, that Werker and Tees (1983) found that English 4-year-old children perform more poorly than older children and adults do. Specifically, performance on the Hindi retroflex/dental (/Ta/-/ta/) contrast was found to be poor from age 10–12 months through adulthood; however, the pattern of findings for a perceptually easier Hindi voicing contrast was quite different. English-learning infants aged 6–8 months showed high levels of discrimination. Children of 10–12 months and 4 years of age showed poor discrimination. Of critical importance, adults and children 8 and 12 years of age showed better discrimination than the younger children did. A similar pattern has been reported more recently by Best (1994) for a different consonant contrast. Finally, Polka and Werker (1994) reported that adult English speakers can discriminate the German front/back vowel contrasts, whereas infants aged 10–12 months cannot. These findings present a serious challenge to the Maintenance model.

The Maintenance model's requirement of permanent change early in life might also predict that declines in perceptual sensitivity will be reflected in production difficulties. This prediction is, however, incompatible with the known fact that children far older than 10–12 months of age can move to a new country and acquire a new language with no trace of an accent (see the chapter by Gleitman and Newport in this volume). Indeed, the best estimates are that children up to at least 6 years of age, and possibly later, remain accent-free when acquiring a new language. An adequate model of age-related changes in speech perception has to be able to account for these data.

Another problem for the Maintenance view is that some nonnative contrasts are easy for adults to discriminate, even with no training at all. In a classic example Best, McRoberts, and Sithole (1988) tested infants and adults on their ability to discriminate a nonnative contrast that not only is absent from the native language but also comprises phones that are completely unlike any occurring in the native language. Best and colleagues selected an apical/lateral click contrast from the Zulu language because clicks are not used at all in the English language; thus, it can be safely assumed that English speakers have not been exposed to these phones—at least not in a linguistic context.[3]

Testing was conducted with English-learning infants 6–8, 8–10, 10–12, and 12–14 months of age, as well as with Zulu- and English-speaking adults. As expected, Zulu adults easily discriminated the click contrast. According to a Maintenance/loss explanation, older infants and adults should be unable to discriminate a phonetic contrast comprised of phones not heard in the input. Contrary to this prediction, Best and colleagues found that English subjects of all ages could easily discriminate the apical/lateral click contrast. Infants of 10–12 and even 12–14 months of age performed as well as both Zulu adults and infants aged 6–8 months. Similarly, English adults discriminated this contrast with ease.

On the basis of these results, Best and colleagues proposed that it is not just experience or lack thereof that leads to a change in speech perception performance; instead, the experience needs to map on to the phonology of the language of input. Phones that are similar to those used in the native language will be susceptible to reorganization, but phones that are completely unlike the phonetic repertoire of the language of input will not be assimilated to native-language phones and will remain discriminable. These nonassimilable phones will be treated as any other nonspeech signal, and

3. Subjects may have been exposed to these sounds in a nonlinguistic context. The apical click sounds something like the 'tsk tsk' nonspeech sound we make, and the lateral click is something like the click one makes to urge a horse to run faster. Even so, it can be safely assumed that these sounds would be extremely rare in the input.

will be discriminable on the basis of any available acoustic information. (See Best 1994 for a fuller description of the proposed model and for a review of more recent research.)

A recent series of experiments by Pegg and Werker (1994) provides a new kind of challenge to the Maintenance/loss view. In this research subjects were tested on their ability to discriminate phones that infants "hear" in their native language but that do not have functional phonemic significance (are not used to distinguish meaning). The phones investigated were two allophone variates. Briefly, allophones are variations on phonemes that occur within a language but do not function to contrast meaning. Some allophonic variation is random (speakers might vary in how far they push their tongue between their front teeth when pronouncing [th]), whereas other allophonic variation is systematic and predictable. Either way, if the variation exists in the input, listeners have had "experience" hearing it. Thus, if simple exposure to phonetic variation is adequate to maintain discriminability, then listeners should be able to discriminate the phonetic variation that occurs in allophones (see MacKain 1982 for an early discussion of this issue).

To test this hypothesis, Pegg and Werker (1994) compared adult English speakers with English-learning infants of 6–8 and 10–12 months of age on their ability to discriminate a systematic allophone distinction that occurs in English. The phonetic distinction selected involves allophonic variations in the English phonemes /d/ and /t/. In English, the /t/ produced in initial position in a word is classified as voiceless, aspirated (phonetically transcribed [t^h]). The /d/ in initial position is voiced, unaspirated, and is transcribed phonetically as [d]. The /t/ following /s/ in "st" words has characteristics of both initial position /t/ and initial position /d/; nevertheless, when the [s] is removed from (s)t words (hereafter referred to as (s)/t/),[4] English speakers perceive it as a /d/, and, indeed, consistently rate it as a "good" token of /d/.

To examine the effect of experience on perception of an allophone distinction, Pegg and Werker tested infants of 6–8 and 10–12 months of age, as well as English-speaking adults, on their ability to discriminate multiple exemplars of (s)/ta/ from multiple exemplars of /da/ using the Conditioned Head Turn procedure. The results revealed that the majority (but not all) of English adults discriminate this allophone difference, as do the majority of infants aged 6–8 months; however, almost none of the infants aged 10–12 months who were tested were able to discriminate this comparison. (For a fuller description, see Pegg 1995.)

4. It is still transcribed phonemically as a /t/ because it patterns phonologically as a /t/. For example, voiceless stop consonants such as /t/ can occur before /r/ in words such as *tree* or *street*, but we cannot have voiced consonants such as /d/ occur before /r/ in words.

These results are quite interesting. First, the results from the adults show that adult listeners can perceive a systematic allophonic distinction, even when both allophones are categorized as variations in /da/. Second, the results from the 6–8-month-old infants broaden the meaning of "language-universal" phonetic sensitivity in young infants. Here many young infants discriminate a phonetic distinction that is systematic and natural, but that is not used (with this precise form of phonetic variation) as a phonemic contrast in any of the world's languages. Finally, the results from the 10–12-month-old infants are most clear. Infants this age are unable to discriminate the allophone variates, even though they hear them in their everyday language and even though the majority of adult speakers of the language can discriminate them. This finding confirms that simply "hearing" phonetic forms in the input language is not necessarily sufficient to maintain discriminability at 10–12 months.

Taken together, these studies suggest that the Maintenance hypothesis is inadequate to account for the results to date. Indeed, instead of a pattern such as that shown in figure 4.4, the effects of experience on perceptual performance seem to show high levels of ability in the young infant, a dip from 10 months to at least 4 years of age, and recovery of function some time after. (See figure 4.5.) Thus, an alternative explanation is required.

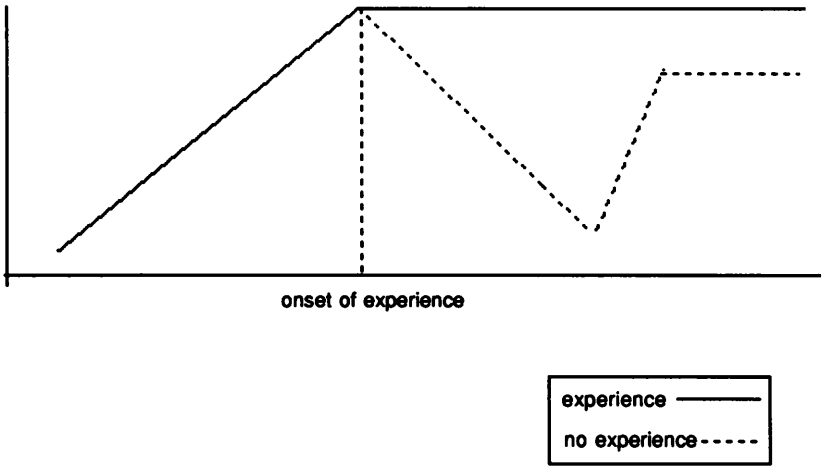

Figure 4.5
A representation of the data from studies of cross-language phoneme discrimination with subjects of different ages showing a decline in performance toward the end of the first year of life followed by an improvement in discrimination performance between early childhood and adulthood.

4.4 A New Approach: Functional Reorganization

If the change in phoneme discrimination does not involve an absolute "loss" of sensitivity as a function of experience, what might it involve? One possibility that has been suggested is that the change in nonnative phoneme discrimination involves some kind of functional reorganization (Werker and Pegg 1992). The central assumption of this Functional Reorganization Hypothesis is that children's changing sensitivities reflect not absolute, hardware changes in the auditory system, but rather the needs of the developing linguistic system. According to this explanation, the "universal" sensitivities seen in the newborn infant may continue to be present across the life span, and thus may show up in adults under certain kinds of testing conditions. The changes seen in perceptual performance across the first year of life do not reflect changes in the hard-wiring of the brain; instead, they reflect the operation of postperceptual processes that allow the infant to attend to only phonetic information that distinguishes meaning.

The Functional Reorganization Hypothesis can account nicely for much of the data. It appears that very young infants respond to any detectable phonetic variation in the speech signal—a response bias that prepares the child to acquire any of the world's language. Young infants are therefore listening functionally to as many aspects of the phonology as they can pick up. Adults, on the other hand, have a bias for phonemic information and most readily and reliably discriminate phones that distinguish meaning; however, adults are also able to listen strategically for nonphonemic variation if required by particular task conditions. It is infants 10–12 months of age, and possibly even children up to at least 4 years who seem constrained to listen only to that information that might be used to distinguish meaning. Why might the older infant and young child be so constrained?

Starting at around 1 year of age, infants are poised to begin to learn words, a task they will devote considerable energy to over the next several years. A language-specific bias to attend to only those differences that are used to contrast meaning in the native language will help the child map sound on to meaning. Sensitivity to too much variation could result in errors, making the child map different meanings on to different productions of a single phoneme. Attention to just that variation that is phonemic in the native language would protect the child from this kind of error and would make the word-learning process that much more efficient.

This explanation can account for the findings reported by Best and colleagues (1988) that sounds completely unlike those used in the native language (Zulu clicks) are not subject to reorganization (and stay discriminable) since they are not perceived as part of the linguistic system. As

noted by Best (1994), they remain discriminable as nonspeech sounds. It can also account for the recent research by Pegg and Werker showing that infants of 10–12 months stop discriminating two allophonic variates that occur in the native language but are not used to distinguish meaning. The Functional Reorganization Hypothesis requires that infants of this age treat such stimuli as the same.

Because the Functional Reorganization Hypothesis does not involve a change in the neural hardware that supports "universal" discrimination in the young infant, the underlying sensitivity to nonnative distinctions is not lost. Adults, whose vocabularies are well established and relatively stable, have the cognitive "distance" and strategic skills to listen for whatever information is required in a particular task. Thus, if the task requires listening to nonnative phonetic distinctions, the adults will—with varying amounts of practice or training—be able to demonstrate such an ability. Similarly, young children moving to a new linguistic environment would have the auditory sensitivity to listen to the relevant phonetic detail to acquire words in their new language. This could account for the finding that young children can move into a new language environment and acquire a second language accent-free.

4.5 Residual Problems for the Functional Recorganization Hypothesis

There are two residual problems for the Functional Reorganization Hypothesis. First, there does seem to be a critical period for language acquisition, including acquisition of accent. The Functional Reorganization Hypothesis would imply that there is continuing flexibility throughout life for acquisition of the phonological structure of a new language. And, indeed, there does seem to be flexibility throughout life for discriminating nonnative phonetic contrasts; but, as noted above, although young children can acquire a second language accent-free, older children and adults cannot. The fact that adults and older children who learn a second language still have an accent (even though they can learn to discriminate the phonetic elements) could be an instance of perceptual flexibility being greater than productive flexibility. There could be a "critical period" for acquisition of the ability to "speak" a new language without an accent, but the ability to perceive the phonological detail in new languages could remain open (see Flege 1992). It is important to note, however, that although adults can improve their performance on nonnative phonetic contrasts, they seldom reach "native" levels of performance (Polka 1992). This remaining difference in perceptual sensitivity may be akin to the "accent" shown by adult second-language learners.

A second problem for the Functional Reorganization Hypothesis is that there appears to be a mismatch between the kinds of perceptual sensitivities shown by infants and those shown by young children. For years there have been reports that children aged 2 to 5 have difficulty discriminating syllables and/or words that differ in only a single phoneme (for a review see Barton 1976). At first blush, it would seem that these results are incompatible with the Functional Reorganization Hypothesis, since it would predict that children should be listening for just that kind of information. Similarly, if children are listening to native language phonemic information, they should be able to use it when learning new words. There is as yet no evidence, though, that infants aged 12–18 months can use fine phonetic detail when learning new words. Indeed, children of this age treat phonetically similar nonsense words as identical and seem unable to pair them with distinct objects (e.g., they treat "dog" and "bog" as variations in the word *dog*). (See Werker and Pegg 1992; Werker and Stager 1995.)

How can the Functional Reorganization Hypothesis be correct if children are not using language-specific phonemic information in word-learning tasks? The answer to this question is not known, but some speculation is provided below. In addition to the influences on consonant and vowel perception, the ambient language exerts profound influences on other aspects of speech processing during the first year of life, with language-specific tuning first evident for the more global properties of the native language (Jusczyk et al. 1992; Mehler et al. 1988). Also, by 10 months of age infants display a variety of cognitive advances that could be important for the way they pick up and use linguistic information. By this age infants are able to detect, remember, and compare separate sources of information. Perhaps most relevant, they are able for the first time to form perceptual categories on the basis of arbitrary correlations of features (Younger and Cohen 1983). One possibility is that these emerging cognitive skills allow the infant to detect the correspondence between certain kinds of speech sounds and regular, systematic use by adults. Indeed, there is evidence that by 10 months of age the infant may for the first time have the cognitive skills in place to selectively attend to just that information in consonants and vowels that is criterial in defining native-language phonemic categories (Lalonde and Werker, in press). Finally, just as prosodic structure might help a child find larger linguistic units such as clauses and phrases (Gleitman et al. 1988), prior tuning to the more global properties of the native language may provide a familiar prosodic context that helps the child find criterial phonetic detail (Christophe, Dupoux, and Mehler 1993; Jusczyk, in press). See also Morgan and Saffran (in press).

Selective perception of native-language phonemic distinctions may not automatically propel the child into using all the details in their perceptual

representations in word-learning tasks. Indeed, just as children overgeneralize in applying the meaning of words in the early stages of language acquisition (Clark 1983), they may overgeneralize in applying the sounds of words in the early stages of language acquisition. The language-specific phonemic perceptual categories established during the latter part of infancy may provide the child with the arsenal of capabilities required to map sound on to meaning, but children may only sample from that arsenal as required to distinguish one word from another in their lexicon. Thus, if they have a vocabulary of only fifty words, very little information might be required to distinguish those words from one another. This is consistent with some of the findings in child phonology, which suggest that, in the early stages of word learning, children represent the phonological structure of words at their most global level (Ferguson and Farwell 1975). As children acquire more words, there is pressure to "fill in" more information. At this time the child may use more of the capacity provided by the language-specific tuning that occurred during the first year of life. Support for this possibility is provided by Barton's work (1976) showing that if children know words very, very well, even 2-year-old children show evidence of easily distinguishing words that differ in only a single, minimal phoneme distinction (for example, "pear" vs. "bear").

According to this line of reasoning, the Functional Reorganization Hypothesis may be maintained in an altered form. Experience and cognitive development allow children to attend selectively to only the phonetic detail that is important for defining the phonemic categories of their native language. Children, however, need use only parts of that available information in word-learning situations, representing no more detail than is required to distinguish items in their lexicon. Finally, once children have a working knowledge of the phonology of the native language, tests of word discrimination are no longer simple perceptual tasks. They are now tasks that require children to use the detail available in their representations of those words. If children represent words only globally, then tests of word discrimination will reveal gaps in detail, unless, as is the case in the Barton study above, there has been sufficient pressure to "fill in" the entire phonetic representation.

4.6 Summary

Research in cross-language speech perception has revealed a number of findings that provide insight into the relations between initial perceptual biases and subsequent language acquisition. To summarize briefly, young infants show a sensitivity to many "universal" properties of human languages, including an ability to discriminate both native and nonnative

phonetic contrasts. During the first year of life infants' performance on speech perception tasks changes to reflect enhanced sensitivity to the properties of the native language. Language-specific discrimination of many aspects of the native language—most clearly consonant contrasts—is first seen at around 9–10 months of age. This change in sensitivity to phonetic contrasts occurs not simply from lack or presence of exposure, but from exposure to phonetic contrasts that are used to contrast meaning in the native language. Furthermore, this age-related change in performance on nonnative phonetic contrasts is not absolute. When tested under more sensitive, or nonlinguistic, testing conditions, older subjects are still able to discriminate nonnative distinctions. Indeed, adults seem to be able to strategically listen for whatever detail in speech sounds is required for successful performance in a task. Children, though, at least up to age 4, are less able to engage in such strategic activity and seem constrained to listen to only native language phonemic differences.

When the original finding of age-related declines in performance on nonnative speech perception tasks was reported, it was viewed as an instance of a Maintenance/loss phenomenon in perceptual development. With the emergence of new data, as summarized above, the Maintenance explanation became less viable. A Functional Reorganization Hypothesis was proposed in which the age-related changes are understood in terms of the functional linguistic tasks the child faces at different stages in development. This hypothesis is helpful in understanding the data collected to date.

The residual problems with the Functional Reorganization Hypothesis point to fascinating new areas of research that will potentially link findings in speech perception and phonological acquisition to the kinds of patterns that have previously been described in the acquisition of both syntax (maturational constraints, bootstrapping) and semantics (overgeneralizations). The unanswered questions also point to areas of possible synthesis between language acquisition, perceptual development, and emerging cognitive capabilities. In conclusion, it appears that in order to fully understand age-related changes in speech perception we have to consider not just the capabilities given by nature and the role the input plays in shaping those capabilities, but also the particular challenges the child faces at each juncture in language acquisition and the potential contribution of developing cognitive skills.

Suggestions for Further Reading

Ferguson, C. A., L. Menn, and C. Stoel-Gammon, eds. (1992). *Phonological development: Models, research, and implications.* Parkton, MD: York Press.

Gottlieb, G. (1981). Roles of early experience in species-specific perceptual development. In

R. N. Aslin, J. R. Alberts, and M. R. Petersen, eds., *Development of Perception*. Vol. 1. New York: Academic Press.

Jusczyk, P. (1993). From general to language-specific capacities: The WRAPSA Model of how speech perception develops. *Journal of Phonetics* 21, 3–28.

Kuhl, P. K. (1993). Innate predispositions and the effects of experience in speech perception: The perceptual magnet theory. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, and J. Morton, eds., *Developmental neurocognition: Speech and face processing in the first year of life*, 275–288. Dordrecht, Netherlands: Kluwer.

Liberman, A. M., F. S. Cooper, D. P. Shankweiler, and M. Studdert-Kennedy (1967). Perception of the speech code. *Psychological Review* 74, 431–461.

MacKain, K. S. (1982). Assessing the role of experience on infants' speech discrimination. *Journal of Child Language* 9, 527–542.

Miller, J. L., and P. D. Eimas (in press). *Handbook of perception and cognition*. Vol. 11, *Speech, language, and communication*. Orlando, FL: Academic Press.

Morgan, J., and Demuth, K. (in press). *Signal to syntax*. Hillsdale, NJ: Erlbaum.

Strange, W. (Ed.) (in press). *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*. Timonium, MD: York Press.

Problems

4.1 Briefly describe the effects of experience on perception of nonnative consonant contrasts.

4.2 Summarize the data that pose a problem for a simple Maintenance/loss explanation of age-related changes in speech perception.

4.3 Point out similarities and differences in the data on acquisition of foreign language production in comparison to foreign language perception.

Questions for Further Thought

4.1 Critically discuss different theoretical frameworks that might be considered for explaining age-related changes in cross-language speech perception.

4.2 What are some of the limitations in the conclusions that can be drawn from data presented in this chapter, and what kinds of experiments might you design to address those limitations?

4.3 How do developmental changes in speech perception parallel other common patterns in language acquisition?

References

Aslin, R. N., D. B. Pisoni, B. L. Hennessy, and A. J. Perey (1981). Discrimination of voice onset time by human infants: New findings and implications for the effect of early experience. *Child Development* 52, 1135–1145.

Barton, D. (1976). The role of perception in the acquisition of phonology. Stanford University Ph.D. dissertation reproduced by the Indiana University Linguistics Club, Bloomington.

Best, C. T. (1994). The emergence of language-specific phonemic influences in infant speech perception. In J. Goodman and H. Nusbaum, eds., *The transition from speech sounds to spoken words: The development of speech perception*. Cambridge, MA: MIT Press.

Best, C. T., R. La Fleur, and G. W. McRoberts (in press). Divergent developmental patterns for infants' perception of two non-native contrasts. *Infant Behavior and Development*.

Best, C. T., G. W. McRoberts, and N. N. Sithole (1988). The phonological basis of perceptual loss for non-native contrasts: Maintenance of discrimination among Zulu clocks by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14, 345–360.

Christophe, A., E. Dupoux, and J. Mehler (1993). How do infants extract words from the speech stream? A discussion of the bootstrapping problem for lexical acquisition. In E. V. Clark, ed., *Proceedings of the 24th Annual Child Language Research Forum*, Stanford, CA.

Clark, E. (1983). Meanings and concepts. In J. Flavell and E. Markman, eds., *Handbook of child psychology*. Vol. 3. New York: Wiley.

Eimas, P. D., E. R. Siqueland, P. Jusczyk, and J. Vigorito (1971). Speech perception in infants. *Science* 171, 303–306.

Ferguson, C. A., and C. B. Farwell (1979). Words and sounds in early language acquisition. *Language* 51, 419–439.

Flege, J. E. (1992). Effects of age of learning on second language (L2) production. In C. A. Ferguson, L. Menn, and C. Stoel-Gammon, eds., *Phonological development: Models, research, and implications*. Parkton, MD: York Press.

Gleitman, L., H. Gleitman, B. Landau, and E. Wanner (1988). Where the learning begins: Initial representations for language learning. In F. Newmeyer, ed., *The Cambridge linguistics survey*. Cambridge: Cambridge University Press.

Gottlieb, G. (1981). Roles of early experience in species-specific perceptual development. In R. N. Aslin, J. R. Alberts, and M. R. Petersen, eds., *Development of perception*. Vol. 1. New York: Academic Press.

Jusczyk, P. W. (in press). Language acquisition: Speech sounds and the beginnings of phonology. In J. L. Miller and P. D. Eimas, eds., *Handbook of perception and cognition*. Vol. 11, *Speech, language, and communication*. Orlando, FL: Academic Press.

Jusczyk, P. W., D. G. Kemler Nelson, K. Hirsh-Pasek, L. Kennedy, A. Woodward, and J. Piwoz (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology* 24, 252–293.

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

Lalonde, C. E., and J. F. Werker (in press). Cognitive influences on cross-language speech perception in infancy. *Infant Behavior and Development*.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Mehler, J., P. W., Jusczyk, G. Lambertz, N. Halstead, J. Bertoncini, and C. Amiel-Tison (1988). A precursor of language acquisition in young infants. *Cognition* 29, 143–178.

Morgan, J. L., and J. R. Saffran (in press). Emerging integration of segmental and suprasegmental information in prelingual speech segmentation. *Child Development*.

Pegg, J. E. (1995). Infant and adult perception of an English phonetic distinction. Ph.D. dissertation. The University of British Columbia.

Pegg, J. E., and J. F. Werker (1994). Age related changes in perception of an allophone distinction in infancy. Presented at the International Conference on Infant Studies, Paris, June 3–5.

Pisoni, D. B., R. N. Aslin, A. J. Perey, and B. L. Hennessy (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance* 8, 298–314.

Polka, L. (1992). Characterizing the influence of native experience on adult speech perception. *Perception and Psychophysics* 52, 37–52.

Polka, L., and J. F. Werker (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance* 20, 421–435.

Stager, C. L., and J. F. Werker (1995). Phonetic similarity influences learning word-object associations in 14-month-old infants. American Psychological Society Meeting, New York, NY. June 29–July 2.

Strange, W., and J. Jenkins (1978). Role of linguistic experience in the perception of speech. In R. D. Walk and H. L. Pick, eds., *Perception and experience*. New York: Plenum Press.

Streeter, L. A. (1976). Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature* 259, 39–41.

Tees, R. C. (1986). Experience and visual development: Behavioral evidence. In W. T. Greenough and J. M. Juraska, eds., *Developmental neuropsychobiology*. New York: Academic Press.

Trehub, S. E. (1973). Infants' sensitivity to vowel and tonal contrasts. *Developmental Psychology* 9, 91–96.

Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development* 47, 466–472.

Tsushima, T., O. Takizawa, M. Sasaki, S. Shiraki, K. Nishi, M. Kohno, P. Menyuk, and C. Best (1994). Discrimination of English /r-l/ and /w-y/ by Japanese infants 6–12 months: Language-specific developmental changes in speech perception abilities. Paper presented at the International Conference on Spoken Language. Yokohama, Japan.

Werker, J. F., J. H. V. Gilbert, K. Humphrey, and R. C. Tees (1981). Developmental aspects of cross-language speech perception. *Child Development* 52, 349–353.

Werker, J. F., and C. E. Lalonde (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology* 24, 672–683.

Werker, J. F., and J. S. Logan (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics* 37, 35–44.

Werker, J. F., and J. F. Pegg (1992). Infant speech perception and phonological acquisition. In C. Ferguson, L. Menn, and C. Stoel-Gammon, eds., *Phonological Development: Models, Research, and Implications*, 285–312. Parkton, MD: York Press.

Werker, J. F., and R. C. Tees (1983). Developmental change across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology* 37, 278–286.

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

Werker, J. F., and R. C. Tees (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America* 75, 1866–1878.

Werker, J. F., and R. C. Tees (1992). The organization and reorganization of human speech perception. In M. Cowân, ed., *Annual Review of Neuroscience* 15, 377–402.

Wiesel, T. N., and D. H. Hubel (1965). *Journal of Neurophysiology* 28, 1060–1072.

Younger, B. A., and L. B. Cohen (1983). Infant perception of correlations among attributes. *Child Development* 54, 858–867.

This excerpt from

An Invitation to Cognitive Science - 2nd Edition: Vol. 1.
Lila R. Gleitman and Mark Liberman, editors.
© 1995 The MIT Press.

is provided in screen-viewable form for personal use only by members of MIT CogNet.

Unauthorized use or dissemination of this information is expressly forbidden.

If you have any questions about this material, please contact cognetadmin@cognet.mit.edu.