PREDISPOSITIONS FOR THE PERCEPTION OF SPEECH BY HUMAN INFANTS

Patricia K. Kuhl, Department of Speech and Hearing Sciences, Child Development and Mental Retardation Center, University of Washington, Seattle, WA. 98195.

The development of speech production and perception in the human infant shares certain themes with the acquisition of communicative repertoires in animal species. Among those themes is the notion that infants of a species demonstrate predispositions for the perception of communicatively relevant acoustic signals. While the animal literature provides examples in which innate predispositions are in evidence, a growing body of literature on the complex role of "normal" experience, and the effects of selective auditory exposure, in maintaining, facilitating, and inducing such behavior is accruing, leading to the hypothesis that infants are predisposed toward fairly simple acoustic features and develop the perception of "configurational" models only with experience. Two approaches to examining the role of experience in the perception of speech by human infants are discussed.

Converging Themes in Developmental Neurobiology

At the end of the first decade of research on the perception of speech by young infants, the list of published experiments is long and the speech features that have been examined is extensive (see Kuhl, In Press, for review). The common theme running through this work is the examination of potential auditory perceptual predispositions that human infants bring to the task of learning language - predispositions that would direct the infant toward the acoustic features that are particularly relevant to the perception of speech, such as those acoustic features which signal the segmental and nonsegmental elements of the language.

The notion that members of a species may be perceptually predisposed to attend to, resolve more precisely, respond to, or to otherwise treat differently, visual and auditory signals that are relevant to their survival is an old theme in the literature on communicative behavior in animals and humans (Lorenz, 1965). Many attribute stimulus prepotencies to species-specific neural mechanisms that have evolved specially for that purpose and perceptual predispositions that are innately determined. The evidence for such mechanisms is both behavioral and physiological and largely stems from work on animal communication (see Schneich, 1977, for a review of neurophysiological data and Gottlieb, 1976a, for a review of behavioral data).

The discovery in behavioral and physiological experiments that communicatively relevant stimuli enjoy special status for the adult perceiver naturally raises questions about the development of these behaviors in infants of the species. While the early theorists (Lorenz, 1965; Tinbergen, 1951) stressed the "instinctiveness" of certain behaviors and underplayed the role of experience, "learning" in the classical sense, or maturation in the development of complex behavior, more recent theorists (Gottlieb, 1976a) have stressed the complex role that experience plays and the variety of different ways experience affects the organism (Gottlieb, 1976b).

Recent physiological evidence suggests that sensory input during early development has an effect on central neural mechanisms, particularly in the visual system; the responsiveness of units in the visual cortex of adults is biased by distorting or denying early "normal" visual experience, or by selective visual exposure. This physiological "plasticity" in the visual system can be species-specific and evidence for "critical periods" exists (see Daniels and Pettigrew, 1976, for review).

The effects of selective auditory exposure are less well known. Silverman and Clopton (1977) and Clopton and Silverman (1977) noted substantial losses in binaural interaction at the inferior colliculus in rat after early monaural deprivation. Clopton and Silverman (1978) demonstrated changes in the latency and duration of neural responses to clicks at the level of the inferior colliculus in rat after early auditory deprivation. Clopton and Winfield (1976) further demonstrated using the rat that exposure during the first four months of life to patterned sound (upward tone sweeps, downward tone sweeps, or noise bursts) increases the response of units in the inferior colliculus to that pattern relative to a similar but inexperienced pattern. No effects of selective exposure were found in an adult population of rats.

Perhaps the best examples from the animal literature on the interactions between innate predispositions and experience are to be found in the growing literature on song learning in the Passerine bird (Marler, 1973). Certain songbirds must hear their songs in order to learn them but there are interesting constraints on learning; the exposure must be to the conspecific song and it must occur during a "critical period." Marler hypothesizes that song vocalization is developed by reference to an "auditory template," a mechanism that is specific enough to detect some of the critical features of the conspecific song and thus direct the bird's attention in its direction, but one which requires exposure to the song to "fill in" the details of its acoustic structure. As Marler (1973) describes, their learning is not left purely to chance, it ". . . takes place within a set of constraints which seem designed to ensure that the learning bird's attention shall be focused on a set of sounds that is biologically relevant. . ." (p.80). To make the songbird parallel even more striking, Nottebohm et al. (1976) have demonstrated functional hemispheric asymmetry for the production of song in these birds. Using ablation techniques, they have demonstrated that the left hemisphere controls song production in the Canary, but if ablation of the left motor area occurs before the bird has passed the critical period for vocal learning, the bird's song develops normally using the subordinate right motor area.

Marler interprets these data as indicating that the innate direction that the infant comes into the world with is simply that - a direction or guideline pointing the infant in the appropriate direction, rather than a complete "schema" of the song. He believes that the predispositions are toward rather simple stimulus features and only with continued exposure to the configuration that is being detected does the infant develop a "schema" of the complex stimulus array.

Predispositions for the Perception of Speech by Human Infants

There are two ways in which the role of experience is currently being examined for the perception of speech by human infants. One approach is to chart the course and examine the nature of perceptual changes that occur as a result of exposure to a particular language. Another approach is to examine the infant's recognition of abstract auditory-phonetic categories rather than simple stimulus features, expecting that the former may reveal developmental trends.

How does linguistic exposure modify the way in which infants perceive speech sounds? While not well understood, the perceptual effects of exposure to one's native language have been documented in adult listeners (Miyawaki et al., 1975; Abramson and Lisker, 1970). Taken together with the existing data on the perception of speech by infants, these data have led to the hypothesis that infants discriminate all of the simple phonetic contrasts at birth regardless of their linguistic environments, but that due to the lack of exposure to certain phonetic units during development the infant somehow loses the ability to distinguish them from contrasting phonetic units.

Attempting to chart developmental changes in an infant's perception that can be attributed to linguistic exposure has received some attention, but we are still without a simple answer to the question. The evidence is fairly convincing that infants being reared in non-English-speaking environments are capable of discriminating at least one phonetic contrast (voiceless-unaspirated /pa/ from voiceless-aspirated /p^{h}a/) that is phonemic in English but not in the infant's native language. Streeter (1976) using the sucking-habituation technique, demonstrated that two-month-old African Kikuyu infants discriminated the English contrast in addition to discriminating a voicing contrast that is phonemic in the Kikuyu language but not in English (prevoiced /ba/ from voiceless-unaspirated /pa/). Lasky, Syrdal-Lasky and Klein (1975) demonstrated similar results for Spanish infants of the same age using a heart-rate technique.

On the other hand, the case for discrimination of the prevoiced /ba/ from the voiceless-unaspirated /pa/ by American infants is not quite as clear. Recent studies (Eilers, Wilson and Moore, 1977; Eimas, 1974) have failed to provide evidence that American infants discriminate pairs of stimuli that are as close on the continuum as those discriminated by the Spanish and Kikuyu infants. However, there are a number of problems with these cross-language comparisons. First, the stimuli are synthesized to manipulate an acoustic cue that is acoustically fragile and is likely to be subject to variation due to the differences in acoustic calibration across laboratories. A more recent set of studies claims to be immune to this criticism. Using the head-turn technique, Eilers, Gavin and Wilson (In Press) tested six-month-old American and Spanish infants in the same laboratory, but in two different studies, and demonstrated that while both groups discriminated the English contrast, only the Spanish infants discriminated the Spanish contrast.

Do infants recognize the configurational properties of phonetic categories? Only recently have researchers attempted to find out whether infants are capable of recognizing the similarity among sounds that have the same phonetic label when the sounds occur in different phonetic contexts, when they occur in different positions in a syllable, or when they are spoken by different talkers.

A conditioned head-turn response for visual reinforcement has been successfully used with six-month-old infants to test the recognition of phonetic categories (Kuhl, 1978). In these tasks, infants are trained to make a head-turn response when one speech token is changed to another speech token (like from /a/ to /i/). During training, vowels produced by a male talker (computer-synthesized) are used; subsequently, infants are tested with computer-synthesized vowels produced by female and child talkers. The ease with which the infant generalizes to new exemplars from the category indicates the degree to which the infant perceives the similarity among the tokens from a given category.

Results to date in these category-formation tasks strongly suggest that vowel categories are readily perceived by the infant listeners. Tasks requiring the infant to recognize a change from the vowel category /a/ to the vowel category /i/ and tasks requiring the infant to recognize a change from the vowel category /a/ to the vowel category /ɔ/ result in near perfect transfer of learning to the new tokens from the categories (Kuhl, 1978). We have also completed studies on the categorization of fricative consonants, such as /f/ vs. /θ/, and /s/ vs. /ʃ/ (Holmberg, Morgan and Kuhl, 1977). In general, our results suggest that the /a-i/ contrast is the easiest in this category-formation task, that the /f-θ/ contrast is the most difficult one, and that the /a-ɔ/ and the /s-ʃ/ contrasts are of intermediate difficulty.

These category-formation experiments (discussed in detail in Kuhl, 1978) have two advantages. First, one can test the infant's recognition of abstract configurational properties of speech-sound categories, and second, one can test how readily or efficiently the infant forms categories based on dimensions that are not phonetically relevant, at least in English, such as pitch contour or stress. These techniques may demonstrate that all infants recognize categories based on certain "focal" auditory dimensions, but that their tendencies to attend to particular acoustic dimensions is modified by exposure to a particular language.

Systematic experiments examining the perception of abstract perceptual categories, rather than simple discriminations, in at least two different populations in which the target acoustic features are chosen such that they are phonemically relevant to one population and not to the other are necessary before the contributions of innate predispositions and experience will be understood in the development of speech perception.

References

Abramson, A. and Lisker, L. (1970): "Discrimination along the voicing continuum: Cross-language tests," in Proceedings of the Sixth International Congress of Phonetic Sciences, Prague, 1967, Academic, 569-573.

Clopton, B.M. and Silverman, M.S. (1977): "Plasticity of binaural interaction. II. Critical period and changes in midline response," J. Neurophysiol. 40, 1275-1280.

Clopton, B.M. and Silverman, M.S. (1978): "Changes in latency and duration of neural responding following developmental auditory deprivation," Exp. Brain Res. 32, 39-47.

Clopton, B.M. and Winfield, J.A. (1976): "Effect of early exposure to patterned sound on unit activity in rat inferior colliculus," J. Neurophysiol. 39, 1081-1089.

Daniels, J.D. and Pettigrew, J.D. (1976): "Development of neuronal responses in the visual system of cats," in Neural and Behavioral Specificity, Vol. III, G. Gottlieb (Ed.), 196-232, New York: Academic Press.

Eilers, R.E., Gavin, W.J. and Wilson, W.R. (In Press): "Linguistic experience and phonemic perception in infancy," Child Develop.

Eilers, R.E., Wilson, W.R. and Moore, J.M. (1977): "Speech discrimination in the language-innocent and the language-wise: A study in the perception of voice-onset time," J. Acoust. Soc. Am. Suppl. 1, 61, S38(A).

Eimas, P.D. (1974): "Linguistic processing of speech by young infants," in Language Perspective-Acquisition, Retardation and Intervention, R.L. Schiefelbush and L.L. Lloyd (Eds.), 55-74, Baltimore: University Park Press.

Gottlieb, G. (1976a): "Early development of species-specific auditory perception in birds," in Neural and Behavioral Specificity, Vol. III, G. Gottlieb (Ed.), 237-281, New York: Academic Press.

Gottlieb, G. (1976b): "The roles of experience in the development of behavior and the nervous system," in Neural and Behavioral Specificity, Vol. III, G. Gottlieb (Ed.), 25-54, New York: Academic Press.

Holmberg, T.L., Morgan, K.A. and Kuhl, P.K. (1977): "Speech perception in early infancy: Discrimination of fricative consonants," J. Acoust. Soc. Am., Suppl. 1, 62, S99(A).

Kuhl, P.K. (1978): "Perceptual constancy for speech-sound categories in early infancy," Chapter presented at the NIH Conference on Child Phonology, May, 1978.

Kuhl, P.K. (In Press): "The perception of speech in early infancy," in Speech and Language: Research and Theory, N.J. Lass (Ed.), New York: Academic Press.

Lasky, R.E., Syrdal-Lasky, A. and Klein, R.E. (1975): "VOT discrimination by four- to six-and-a-half-month-old infants from Spanish environments," J. Exp. Child Psych. 20, 215-225.

Lorenz, K. (1965): Evolution and Modification of Behavior, University of Chicago Press, Chicago.

Marler, P.(1973): The Clarence M. Hicks Memorial Lectures for 1970, University of Toronto Press, Toronto, 69-85.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A., Jenkins J., and Fujimura, O. (1975): "An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English," Percept. Psychophys. 18, 331-340.

Nottebohm, F., Stokes, T.M. and Leonard C.M. (1976): "Central control of song in the Canary," J. of Comp. Neurol. 165, 457-486.

Schneich, H. (1977): "Central processing of complex sounds and feature analysis," in Recognition of Complex Acoustic Signals, T.H. Bullock (Ed.), 161-182, Berlin: Abakon Verlagsgesellschaft.

Silverman, M.S. and Clopton, B. (1977): "Plasticity of binaural interaction, I. Effect of early auditory deprivation," J. Neurophysiol. 40, 1266-1274.

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

Tinbergen, N. (1951): The Study of Instinct, Clarendon Press, Oxford.