

# Discrimination of auditory target dimensions in the presence or absence of variation in a second dimension by infants

PATRICIA K. KUHL

*Department of Speech and Hearing Sciences and Child Development and Mental Retardation Center  
University of Washington, Seattle, Washington 98195*

and

JAMES D. MILLER

*Research Department, Central Institute for the Deaf, St. Louis, Missouri 63110*

Discrimination of two acoustic dimensions of auditory stimuli, vowel identity and pitch contour, was tested with infants between the ages of 4 and 16 weeks using the high-amplitude sucking (HAS) technique. Discrimination of the vowel dimension and the pitch dimension was tested under two conditions: when a change in the target dimension occurred in the *absence* of constant variation in a second dimension, and when a change in the target dimension occurred in the *presence* of constant variation in a second dimension. In addition, discrimination was tested in a combined condition in which one level of the vowel dimension was always combined with one level of the pitch dimension and the stimulus change to be detected was a recombination of the levels of each dimension. Sucking-recovery scores demonstrated that infants always discriminated a change in the target dimension when it occurred without variation in the second dimension, regardless of the dimension that served as the target. However, while variation of the pitch dimension did not alter vowel discrimination, variation in the vowel dimension interfered with discrimination of the pitch dimension. Discrimination was also not evidenced in the combined condition. Analysis of the group time-to-habituation (TH) data revealed that significantly longer TH scores were correlated with a failure to demonstrate discrimination. The data are discussed in terms of the formation of auditory perceptual categories in early infancy as they relate to the acquisition of speech and language and, more generally, to developmental attention and memory for auditory stimuli.

The first decade of research on the perception of speech in early infancy has produced a large number of published experiments demonstrating that infants under 4 months of age are capable of discriminating among many, if not all, of the sounds that encompass the phonetic inventory commonly used in the languages of the world (for reviews, see Eilers, 1980; Eimas & Tartter, 1979; Jusczyk, 1981; Kuhl, 1979a; Morse, 1978). The human infant, therefore, appears to demonstrate sufficient auditory acuity early in life to differentiate among many of the sounds that will eventually carry meaning.

A portion of the data reported here were presented by these authors at the 90th meeting of the Acoustical Society of America, San Francisco, November 1975, and were discussed in a chapter published by the first author (Kuhl, 1976). The research reported here and its preparation for publication were supported by grants from the National Institutes of Health to the Central Institute for the Deaf (NS 03856 and RR00396) and by a grant to P. K. Kuhl from the National Science Foundation (BNS 79-13767). The authors wish to thank A. Maynard Engebretson, who designed the logic device used in the experiment, Stan Garfield, who synthesized the vowel stimuli, and Joyce Shelnett, R.N., who assisted in testing the infants.

Another line of research given considerable attention during the first decade of investigation of infant speech perception has been the infant's tendency to perceive similarities among sounds that are not physically identical. Two types of investigations have specifically aimed at the infant's tendencies to form perceptual groupings that conform to those of adult speakers. First, the well-known studies of categorical perception have compared the infant's tendencies to discriminate between stimuli drawn from an acoustic continuum with discrimination performance of adult listeners. The classic finding is that infants discriminate stimuli that straddle the adult-defined "boundary" between two phonetic categories better than they discriminate two stimuli that represent the same physical difference but fall into a single phonetic category (Eimas, 1974, 1975; Eimas, Siqueland, Jusczyk, & Vigorito, 1971), thus replicating the discrimination data obtained with adult listeners (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). This result has been taken as evidence that infants tend to perceptually group the sounds on either side of a phonetic boundary and that they are

predisposed to hear a discontinuity at the location of the phonetic boundary. In addition, it has been shown that infants respond in this way for contrasts that are not phonemic in the linguistic environment in which the infant is being reared, but are phonemic in some other language environment (Lasky, Syrdal-Lasky, & Klein, 1975; Streeter, 1976), suggesting the possibility that infants are sensitive at a very young age to the universal set of potential phonetic boundaries.

A second type of investigation, again aimed at the perceptual grouping of auditory stimuli by infants, involves investigations of the tendency for a listener to perceive a similarity among stimuli that share their phonetic labels but differ in other significant respects, such as the surrounding phonetic context in which the unit occurs or the talker who produces it (Fodor, Garrett, & Brill, 1975; Kuhl, 1979b). The tendency for adult listeners to perceive phonetic units across context and talker as similar has been termed "perceptual constancy," since it shares many of the characteristics of the examples of perceptual constancy in the literature on vision.

Both of these behaviors, the tendency to perceive differences between sounds drawn from an acoustic continuum categorically, when only one acoustic dimension is varied, and the tendency to perceive a similarity among sounds representing a phonetic category when both their criterial acoustic dimensions and many irrelevant acoustic dimensions vary, tap a listener's proclivity to form auditory categories. The degree to which the young infant manifests auditory predispositions that are similar to those exhibited by adult listeners and that conform to phonetic categories is critical to our understanding of the infant's general biological predispositions to acquire language.

While the data gathered on categorical perception in human infants is extensive, the data on perceptual constancy is not. Fodor et al. (1975) examined the acquisition of a head-turn response for visual reinforcement in 14- to 18-week-old infants under two stimulus conditions. In both conditions, three syllables were randomly presented (/pi/, /ka/, /pu/) but only two of the three were reinforced. In one condition, the stimuli being reinforced were phonetically related (/pi/ and /pu/); in the other condition, the stimuli being reinforced were not phonetically related (/pi/ and /ka/). The authors hypothesized that if infants tend to hear the similarity between two syllables that share the initial consonant in spite of the differences in the acoustic cues for that consonant and in spite of the irrelevant differences between the two syllables, such as their vowels, then their tendencies to learn the association ought to differ in the two conditions. Their hypothesis was supported. While the analyses showed that the proportion of head turns was greater for the reinforced stimuli in

both conditions, a significant interaction followed up by a simple-effects analysis demonstrated that the difference between the proportion of head turns to reinforced and nonreinforced stimuli was significant only for the phonetically similar group.

Kuhl (1979b) took a somewhat different approach with 6-month-old infants, using a transfer-of-learning experimental design. Infants were trained to make a head-turn response when a constantly repeated speech sound (the "background" stimulus) was changed to a second sound (the "comparison" stimulus). If the infant produced a head-turn response during the presentation of the comparison stimulus, the infant was rewarded with a visual stimulus (an animated toy monkey). On an equal number of "control" trials, the stimulus was not changed (i.e., the background stimulus was continuously presented) and the number of "false-positive" head-turn responses was observed (see Kuhl, 1979b, for additional details).

The training stimuli in the Kuhl (1979b) study were the steady-state vowels /a/ and /i/, computer simulated to be appropriate for a male talker and produced with a falling intonation contour. The design of the experiment was such that after the infants met a performance criterion on the discrimination of these two vowels (9 of 10 consecutive correct trials), additional /a/ and /i/ exemplars, computer simulated to be appropriate for female and child talkers with either rising or falling intonation contours, were systematically introduced. In two studies, infants provided evidence which strongly supported the contention that, by 6 months of age, infants readily perceive the similarity among vowel tokens produced by a male, a female, and a child talker, regardless of the intonation contour produced by the talker.

In a second experiment using this same set of vowel stimuli, Kuhl and Hillenbrand (Note 1) demonstrated that another group of 6-month-old infants performed equally well when "sorting" these same stimuli along another dimension, intonation contour. That is, the infants readily learned to produce a head-turn response for visual reinforcement when an /a/ vowel with a rising intonation contour was presented and to inhibit the response when an identical /a/ vowel with a falling intonation contour was presented. Similarly, this learning readily transferred to the /a/ vowels produced by other computer-simulated talkers and to another vowel (the vowel /i/, produced by the set of computer-simulated voices). In other words, the Kuhl (1979b) and Kuhl and Hillenbrand (Note 1) experiments suggested that 6-month-old infants could apply one of two "sorting rules" to this set of stimuli with nearly equal facility.

The experiment reported here was aimed at determining the tendencies of infants under 6 months of age to attend to, and to discriminate changes in, particular acoustic dimensions of multidimensional

stimuli. Since the head-turn technique described above is appropriate only for infants older than 5½ months (see Kuhl, 1980, for further discussion), the high-amplitude-sucking (HAS) technique (Eimas et al., 1971) was modified to address the question. The modification of HAS used in this experiment, similar to one employed by Bornstein, Kessen, and Weiskopf (1976) with visual stimuli, involves the presentation of randomly alternating stimuli during the preshift and postshift phases of certain conditions in the experiment; in these experiments, the alternating stimuli share one of two acoustic dimensions, either their vowel identities or their pitch contours. The investigation was aimed at the infant's ability to discriminate a change in one of these dimensions when variation in the second dimension was present or absent.

## METHOD

### Subjects

Eighty infants, eight in each of 10 groups, completed the experiment. Forty-two of the infants were females and 36 were males. Their ages ranged from 5 to 16 weeks, with a mean of 11.2 weeks. A total of 123 infants were tested to obtain 80 completions, representing a 35% attrition rate. The 43 infants who failed to complete the experiment did so for the following reasons: failure to produce any sucking responses (11), failure to produce any criterion sucking responses for any 2 consecutive minutes prior to the shift point (12), failure to produce at least 15 criterion sucking responses during the 2 habituation minutes (11), and failure to produce at least one criterion sucking response during the first postshift minute (9). Of the 43 excluded infants (who ranged in age from 5 weeks, 1 day to 14 weeks, 4 days, with a mean age of 10.5 weeks), 19 were assigned to an experimental condition and 24 were assigned to a control condition.

The infants were obtained through referral by a group of pediatricians in the greater St. Louis area. The pediatricians personally delivered a letter which explained the study to mothers while they were still in the hospital. The doctors were instructed to give the letter to mothers whose pregnancies were full term, whose deliveries were normal, and whose infants were not considered to be at risk for any reason. Parents were contacted at a later date by telephone to schedule an appointment.

The parents of the infants were paid \$2 for their participation in the study.

### Procedures

Infants were tested individually in a sound-shielded room. The infant was placed in an infant seat in an experimental "bassinet." A rectangular platform was built at the foot of the bassinet to serve two purposes. A 14×7.5 in. rear-projection screen was mounted to the front of the platform such that when the infant was placed in an infant seat at the head of the bassinet, the rear-projection screen (with a slide of a colorful bear projected on it) was approximately 30 in. in front of the infant and at eye level. The slide served to focus the infant's attention directly forward. A loudspeaker (Fisher ST 465) was placed directly on top of the platform.

Once seated, the infant was given a blind nipple to suck on. This was held by an assistant, who was instructed to encourage the infant to suck on the nipple until white noise was presented through earphones. The white noise was turned on by the experimenter after the infant had begun to suck on the nipple; it was presented at a level sufficient to mask the change in the stimulus. From that time on, the assistant was instructed simply to hold

the nipple in the infant's mouth, making no attempt to encourage the infant to suck on it.

The experimenter, located in an adjoining control room, adjusted the amplitude threshold for each infant's sucking responses in order to define those criterion sucking responses which would trigger the sound presentation. The experimenter adjusted this threshold setting on the logic device until the infant produced between 20 and 30 criterion sucking responses per minute. After each threshold adjustment, the experimenter waited 1 min to establish that the baseline criterion was met; if it was, the experiment was begun. During the preshift phase, sound presentation was contingent upon criterion sucking responses. Each criterion response was followed immediately by the presentation of a 500-msec speech sound and 500 msec of silence. If the infant produced another criterion sucking response during this 1-sec interval, another speech token was presented.

The preshift phase continued until the infant's response rate, tallied in 1-min intervals, dropped 20% below the maximum rate achieved in any minute (after the first 2 preshift minutes) for 2 consecutive minutes. This procedure differs slightly from that of Eimas et al. (1971), who used the immediately preceding minute as the referent. After these two "habituation" minutes, defined as the "shift point," the logic automatically triggered a change in the sound presented to infants in the experimental group, while infants in the control group continued to be presented with the same sound. After the shift point, the experiment continued for 4 additional minutes and was then terminated.

### Stimuli

The stimuli were four steady-state vowels (/a/ and /i/) synthesized at the Central Institute for the Deaf in St. Louis on a terminal analog serial synthesizer. Two tokens of the vowel /a/ and two tokens of the vowel /i/, one with a falling intonation contour and one with a monotone intonation contour, were synthesized with formant values and fundamental frequency values in a range appropriate for a male talker. Table 1 lists the formant (F) frequencies and bandwidths for the first three formants for /a/ and /i/. These center frequencies were taken from Peterson and Barney's (1952) averages of the center-frequency measurements of naturally produced /a/ and /i/ in an /h-d/ context. The bandwidths were taken from Dunn (1963). In addition to the first three formants, which varied for the /a/ and /i/ stimuli, each stimulus included constant upper formants (F<sub>4</sub>-F<sub>10</sub>), also shown in Table 1, taken from Rabiner's (1968) estimates for a male talker. The upper formants tend to improve the naturalness of the vowel tokens. The stimuli were synthesized at a 20-kHz sampling rate and formants over 10 kHz were eliminated to avoid aliasing problems. All stimuli were 500 msec in duration.

Table 1  
Center Frequencies (F) and Bandwidths (B) of the Formants  
of the Synthetic Vowels /a/ and /i/

Formant	/a/		/i/	
	F	B	F	B
1	775	54	263	54
2	1064	57	2378	55
3	2614	93	3099	170
4	3500	175	3500	175
5	4500	281	4500	281
6	5500	458	5500	458
7	6500	722	6500	722
8	7500	1250	7500	1250
9	8500	2125	8500	2125
10	9500	4750	9500	4750

Note—Formants 1-3 varied appropriately to produce the two vowels, and Formants 4-10 were held constant. Values are in hertz.

The pitch-contour specifications were as follows. The falling contours, identical for one of the /a/ and one of the /i/ stimuli, were synthesized in a piecewise linear fashion, changing from 112 to 132 Hz in the first 100 msec, remaining there for 40 msec, and then falling to 92 Hz in the remaining 360 msec. The monotone contours were synthesized at a constant fundamental frequency of 112 Hz.

The amplitudes of the /a/ and /i/ stimuli were synthesized to result in tokens that adult listeners perceived to be equally loud. This was achieved when the overall RMS value of the /i/ token was reduced in amplitude by 3 dB. The resulting stimuli were judged by adult listeners to be good exemplars of the vowels /a/ and /i/ spoken with either a falling or a monotone intonation contour.

Each of the four sounds was recorded onto a disk pack (Control Data, Model 846-2-16) of the Random-Access Programmable Recorder of Complex Sounds (RAP), a self-contained digital recorder which provides random access to all stimuli via a number of peripheral controls (Spenner, Engebretson, Miller, & Cox, 1974). The speech tokens were presented at an average level of 68 dB SPL (A-scale) along with a continuously presented speech-shaped (300-4,000 Hz) masking noise at 12 dB SPL. The spectrum of this noise matched the long-term average spectrum of speech and served to mask low-level ambient noise in the room and in the adjacent hallway without differentially affecting the speech sounds themselves.

**Apparatus and Control Equipment**

One of five commercially available blind nipples (Evenflo, Davol, Curity, Nuk, or Hygeia) was used with each infant. Each nipple could be attached to one of two types of adaptors. The Type 1 adaptor fit the first four nipples listed above and was machined from nylon. It had a screw top that was identical to the top of an ordinary baby bottle. The Type 2 adaptor was also machined from nylon and fit the Hygeia nipple. Both adaptors ended in a narrow tube which could be inserted into Tygon (Type R3603) tubing which led to a pressure transducer (National Semi-

conductor Type LX 160 1G). When the nipple was depressed and released, the resulting pressure changes were visually monitored on an oscillographic recording system (Hewlett-Package, 7702) and were electronically monitored by a logic device.

The logic package, constructed at Central Institute for the Deaf, consisted of a central control unit and a printer unit. The central control unit allowed the experimenter to adjust manually the pressure threshold required for a criterion sucking response, monitored the number of criterion sucking responses per minute, and automatically triggered the change from one RAP address to another to effect the change in the sound stimuli at the shift point. To ensure that any change in the sucking rate at the shift point could not be attributed to sounds produced by the logic device or RAP, a RAP-address change was made for control infants as well by assigning the preshift stimuli to two separate RAP addresses.

**Experimental Design**

The design of the experiment required 10 groups of infants: 5 experimental groups, each of which was presented with some kind of sound change at the shift point, and 5 control groups, for which no sound change occurred at the shift point. Table 2 describes the stimuli presented to each group of infants during the preshift and postshift phases of the experiment. The experiment was designed to investigate the infant's ability to detect a change in the vowel identity or in the pitch contour of a stimulus in each of two conditions: first, in the absence of any variation in the second dimension and, second, in the presence of constant variation in the second dimension. The dimension actually changed at the shift point was called the target dimension.

**Experimental groups.** As shown in Table 2, the five experimental groups can be separated by the nature of the target dimension and whether or not constant variation in the second dimension was present. For two of the experimental groups, the target dimension was vowel identity, and for two others, the target dimension was pitch contour. For one of the vowel target groups and one of the pitch target groups, no variation in the second

**Table 2**  
The Stimulus Conditions in Effect During the Preshift and Postshift Phases of All of the Experimental Groups and Their Matching Control Groups

Condition	Experimental Groups		Control Groups	
	Preshift	Postshift	Preshift	Postshift
Vowel Target Dimension				
Without Pitch Variation	[a <sub>m</sub> ]	[i <sub>m</sub> ]	[a <sub>m</sub> ]	[a <sub>m</sub> ]
	[a <sub>f</sub> ]	[i <sub>f</sub> ]	[a <sub>f</sub> ]	[a <sub>f</sub> ]
	[i <sub>m</sub> ]	[a <sub>m</sub> ]	[i <sub>m</sub> ]	[i <sub>m</sub> ]
	[i <sub>f</sub> ]	[a <sub>f</sub> ]	[i <sub>f</sub> ]	[i <sub>f</sub> ]
With Pitch Variation	[a <sub>m</sub> ] + [a <sub>f</sub> ]	[i <sub>m</sub> ] + [i <sub>f</sub> ]	[a <sub>m</sub> ] + [a <sub>f</sub> ]	[a <sub>m</sub> ] + [a <sub>f</sub> ]
	[i <sub>m</sub> ] + [i <sub>f</sub> ]	[a <sub>m</sub> ] + [a <sub>f</sub> ]	[i <sub>m</sub> ] + [i <sub>f</sub> ]	[i <sub>m</sub> ] + [i <sub>f</sub> ]
Pitch Target Dimension				
Without Vowel Variation	[a <sub>m</sub> ]	[a <sub>f</sub> ]	[a <sub>m</sub> ]	[a <sub>m</sub> ]
	[a <sub>f</sub> ]	[a <sub>m</sub> ]	[a <sub>f</sub> ]	[a <sub>f</sub> ]
	[i <sub>m</sub> ]	[i <sub>f</sub> ]	[i <sub>m</sub> ]	[i <sub>m</sub> ]
	[i <sub>f</sub> ]	[i <sub>m</sub> ]	[i <sub>f</sub> ]	[i <sub>f</sub> ]
With Vowel Variation	[a <sub>m</sub> ] + [i <sub>m</sub> ]	[a <sub>f</sub> ] + [i <sub>f</sub> ]	[a <sub>m</sub> ] + [i <sub>m</sub> ]	[a <sub>m</sub> ] + [i <sub>m</sub> ]
	[a <sub>f</sub> ] + [i <sub>f</sub> ]	[a <sub>m</sub> ] + [i <sub>m</sub> ]	[a <sub>f</sub> ] + [i <sub>f</sub> ]	[a <sub>f</sub> ] + [i <sub>f</sub> ]
Combined Target Dimension				
	[a <sub>m</sub> ] + [i <sub>f</sub> ]	[a <sub>f</sub> ] + [i <sub>m</sub> ]	[a <sub>m</sub> ] + [i <sub>f</sub> ]	[a <sub>m</sub> ] + [i <sub>f</sub> ]
	[i <sub>m</sub> ] + [a <sub>f</sub> ]	[i <sub>f</sub> ] + [a <sub>m</sub> ]	[i <sub>m</sub> ] + [a <sub>f</sub> ]	[i <sub>m</sub> ] + [a <sub>f</sub> ]
	[a <sub>f</sub> ] + [i <sub>m</sub> ]	[a <sub>m</sub> ] + [i <sub>f</sub> ]	[a <sub>f</sub> ] + [i <sub>m</sub> ]	[a <sub>f</sub> ] + [i <sub>m</sub> ]
	[i <sub>f</sub> ] + [a <sub>m</sub> ]	[i <sub>m</sub> ] + [a <sub>f</sub> ]	[i <sub>f</sub> ] + [a <sub>m</sub> ]	[i <sub>f</sub> ] + [a <sub>m</sub> ]

Note—The subscripts "m" and "f" denote the monotone and falling fundamental frequency contours, respectively.

dimension occurred and its value was held constant. For example, during the preshift phase, the infants in the vowel target group without pitch variation were presented with repetitions of a single vowel (/a/), with a falling pitch contour, contingent upon their high-amplitude sucking responses. After the shift point, these infants were presented with a single vowel (/i/), also with a falling pitch contour, contingent upon their sucking responses.

In contrast, for one of the vowel target groups and one of the pitch target groups, variation in the second dimension was always present. For example, infants in the vowel target group with pitch variation were presented with a random sequence of two /a/ vowels during the preshift phase, one with a monotone pitch contour and the other with a falling pitch contour. After the shift point, these infants were presented with a random sequence of two /i/ vowels, one with a monotone pitch contour and the other with a falling pitch contour. In other words, these infants were required to detect a change in the identity of the vowel despite the fact that a second, acoustically prominent dimension such as pitch contour served as a potential distractor.

The identical conditions were in effect when pitch contour served as the target dimension. For one group of infants, the pitch-contour change occurred in the absence of variation in the second dimension (vowel identity). For example, during the preshift phase, infants in this condition were presented with the single vowel /a/ with a monotone contour. The single vowel /a/ with a falling contour was presented during the postshift phase. In contrast, the infants tested with variation in the second dimension were presented with a random sequence of two vowels (/a/ and /i/) during the preshift phase whose pitch contours were monotone. During the postshift phase, these infants were presented with a random sequence of the same two vowels whose pitch contours were both falling. In other words, these infants were required to detect a change in pitch contour (from monotone to falling) despite the fact that a second, acoustically prominent dimension such as vowel identity served as a potential distractor.

A fifth experimental group of infants was presented with two of the vowels during the preshift phase and the two other vowels during the postshift phase, making them comparable to the two groups described above, who were tested with variation in the second dimension, except that the vowels did not share either the vowel or the pitch dimension. One level of each of the two target dimensions was combined to form the stimuli presented during the preshift phase, and then the levels were recombined to form new composite stimuli for the postshift phase. For example, infants in this group were presented with a random sequence of the /a/-monotone and the /i/-falling vowels during the preshift phase and a random sequence of the /a/-falling and the /i/-monotone vowels during the postshift phase. This combined group was included to investigate whether discrimination was evidenced when the stimulus presented after the shiftpoint was simply a recombination of the levels already presented during preshift, as opposed to the presentation of a new level of one dimension, as in all other conditions. In the former condition, the infant must detect a novel combination of levels (a novel "compound") because no new "components" (levels of dimensions) are introduced, whereas in the latter, the infant can respond to the introduction of a new component. There is some evidence from studies on visual perception suggesting that infants do not detect novel combinations of components until about 5 months of age (Bower, 1966; Cohen, 1973; Fagan, 1977), though no firm conclusions can yet be drawn (see Cohen, DeLoache, & Strauss, 1979). In infant studies using speech signals, however, 2- to 3-month-olds have been shown to discriminate the recombination of components, suggesting that they do perceive compounds (Eimas & J. L. Miller, in press; J. L. Miller & Eimas, 1979).

As shown in Table 2, the direction of stimulus change was counterbalanced for all of the experimental groups. The level of the second dimension, held constant for one group in each of the pitch target and formant target conditions, was also counterbalanced.

**Control groups.** As shown in Table 2, each of the five experimental groups has a matched control group in which the stimulus conditions for the preshift phase of the experiment are identical but no stimulus change occurs at the shift point in the experiment.

## RESULTS

Two dependent variables, the sucking-recovery scores and the time-to-habituation (TH) scores, were compared for all experimental groups and their respective control groups.

### Sucking-Recovery Scores

The mean number of criterion sucking responses as a function of experimental condition and time is displayed for all five experimental groups and their respective control groups in Figures 1, 2, and 3. In all of the displays, the average number of criterion-sucking responses, shown as a percentage of the rate that occurred during the maximum minute, is shown for the baseline minute, the 2 habituation minutes, and the 4 postshift minutes.

In order to examine whether there were reliable differences among the groups prior to the shiftpoint in the experiment, individual one-way analyses of variance were performed on the following measures: baseline response rates, maximum response rates, and the mean of the response rate during the 2 habituation minutes. In no instance did the differences among groups approach statistical significance ( $p > .10$  in all instances).

### Vowel Target Groups

The data obtained from infants whose target dimension (the dimension that was changed at the shift point) was vowel identity, along with their related control groups, are shown in Figure 1. The left-hand side of the figure shows the data obtained from infants whose target dimension was vowel identity in the absence of random variation in the second dimension. The right-hand side of Figure 1 shows the data for infants whose target dimension was also vowel identity but who were presented with stimuli whose pitch contours were varied randomly (from monotone to falling) throughout the experiment. For both of these groups, the minute-by-minute data reveal substantial differences in sucking rates between the two experimental groups and their related control groups, suggesting that the infants in the experimental groups discriminated the vowel change.

This difference between the experimental and control groups can be readily seen in the plots directly below the minute-by-minute displays, showing the mean change in sucking rate between the 2 habituation minutes and the first 2 postshift minutes. These "difference scores" were obtained by subtracting the mean sucking rate for the 2 habituation minutes from the mean of the first 2 postshift minutes. The

### TARGET DIMENSION: VOWEL

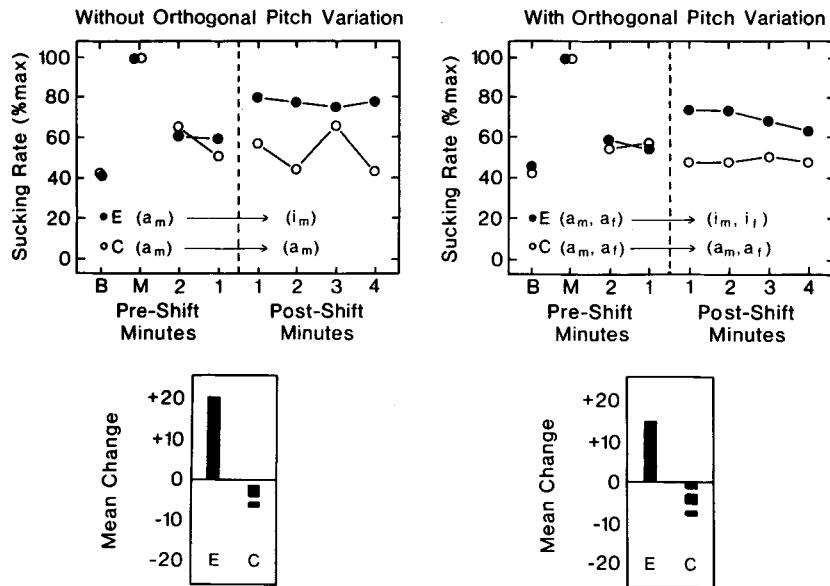


Figure 1. Sucking rates per minute, expressed as a percentage of the maximum sucking rate for baseline, maximum, habituation, and the 4 postshift minutes of the experiment for infants tested in the vowel target groups, both experimental and control. Data for infants tested on the vowel target *without* pitch variation are shown on the left, while data on infants tested on the vowel target *with* pitch variation are shown on the right. The average difference scores for each condition are centered under the minute-by-minute data. They were obtained by subtracting the mean of the two habituation minutes from the mean of the first 2 postshift minutes for each infant.

vowel target group without variation in pitch produced a mean change of +20%, while its control group produced a mean change of -7%. The vowel target group with variation in pitch produced a mean change of +15%, while its control group produced a mean change of -8%.

#### Pitch Target Groups

The data obtained from the two groups of infants whose target dimension was pitch contour, along with their related control groups, are shown in Figure 2. The left-hand side of the figure shows the data obtained from infants whose target dimension was pitch contour, in the absence of random variation in the vowel dimension. The right-hand side of the figure shows the data obtained from infants whose target dimension was also pitch contour, but these infants were presented with stimuli whose vowel identities were varied randomly (from /a/ to /i/) throughout the experiment. The data obtained on infants in the former group, the ones presented with a simple pitch-contour change without random variation in vowel identity, look very much like the two previous experimental groups. That is, infants in the experimental condition demonstrate increased criterion sucking responses after the shift point in the

experiment, when compared with their control group. In contrast, the infants whose target dimension was pitch contour but who were presented with random variation in vowel identity (right-hand side of Figure 2) did not show this characteristic pattern. For these infants, only a small difference in sucking rates separates the experimental and control infants after the shift point in the experiment.

These findings are clearly summarized by the difference-score plots under each of the two minute-by-minute graphs. The pitch target group without variation in the vowel dimension demonstrated a mean change from the 2 habituation minutes to the first 2 postshift minutes of +21%, while its control group demonstrated a mean change of -7%. The pitch target group with variation in the vowel dimension demonstrated a mean change of -5%, and its control group, a mean change of -12.3%.

#### Combined Target Group

In the combined target group, the two stimuli presented during preshift and postshift phases shared neither of the two dimensions, for example, the /a/-monotone stimulus and the /i/-falling stimulus, which changed to the /a/-falling stimulus and the /i/-monotone stimulus.

TARGET DIMENSION: PITCH

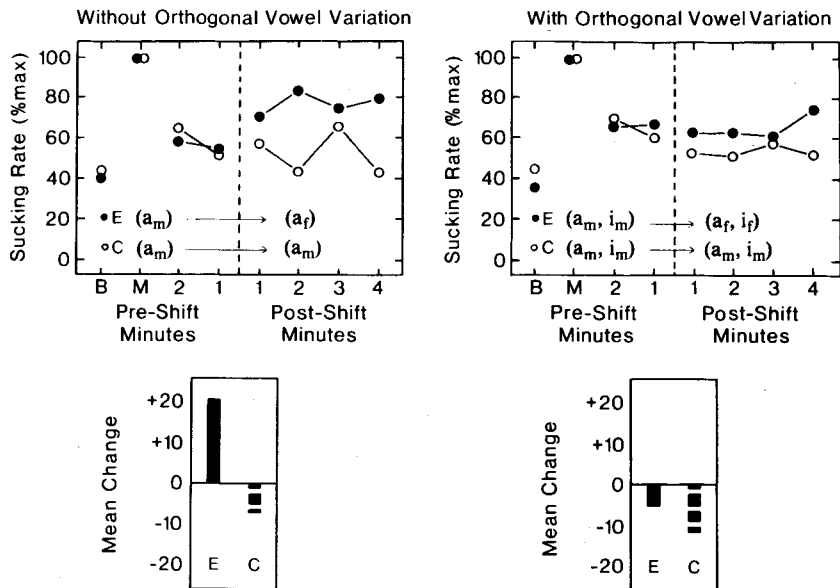


Figure 2. The minute-by-minute and difference-score data for infants tested in the pitch target groups, both experimental and control. Data for infants tested on the pitch target without vowel variation are shown on the left, while data for infants tested on the pitch target with vowel variation are shown on the right.

The results for infants in this group and those in the related control group are shown in Figure 3. As seen in the minute-by-minute plots, the sucking-recovery rates for this experimental group did not differ reliably from those produced by its control group. The mean difference scores were +2.9% and +1.6%, respectively.

Statistical Analyses of the Difference Scores

Statistical analyses of the difference scores obtained from these groups corroborate the results depicted graphically in Figures 1, 2, and 3. The difference scores resulting from conditions in which a change in the target dimensions (either vowel identity or pitch contour) occurred without variation in the second dimension, along with the difference scores from the appropriate control group, were submitted to a one-way analysis of variance. The analysis revealed a significant main effect of groups [ $F(2,29) = 10.34, p < .01$ ]. Newman-Keuls post hoc comparisons (Winer, 1971, p. 217) revealed that the two experimental groups (vowel target and pitch target) did not differ significantly, but that both groups differed significantly from the control group.

Simple t tests were used to compare each of the experimental conditions in which a change in the target dimension occurred in the presence of random variation in the second dimension with the appropriate control group. A t test for independent samples comparing the difference scores obtained by the vowel target with pitch variation group and its

control group revealed a significant difference [ $t(14) = 2.88, p < .01$ ]. In contrast, a t test for independent samples comparing the difference scores obtained by the pitch target group with vowel variation and its control group did not approach statistical significance at the .05 level ( $t = .436$ ). Similarly, a t test

CORRELATED CONDITION

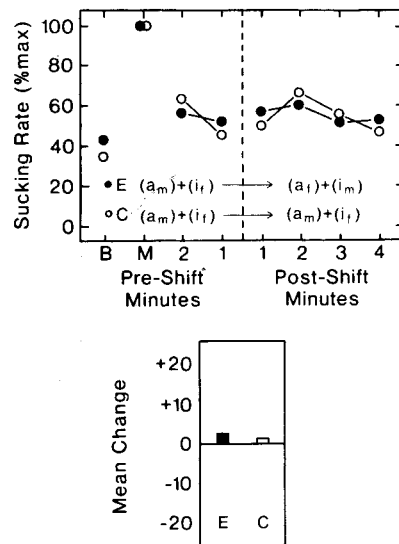


Figure 3. The minute-by-minute and difference-score data for infants in the combined condition, both experimental and control groups.

comparing the combined target group with its control group did not approach statistical significance at the .05 level ( $t = .087$ ).

In summary, the statistical-significance tests parallel the results suggested graphically in Figures 1, 2, and 3. Those results can be summarized as follows: When either a change in the vowel identity or in the pitch contour of a stimulus occurs in the absence of variation in a second dimension, infants demonstrate the ability to discriminate that change. On the other hand, when a change in a target dimension occurs against a background of random change in a second dimension, performance depends upon the specific dimension that serves as the target. When vowel identity is the target dimension and pitch contour serves as the second dimension, infants make the discrimination; however, when pitch contour is the target dimension and vowel identity serves as the second dimension, no evidence of discrimination is obtained. In addition, when the sounds presented during the preshift period involve combinations of one level of each of the two dimensions, and the change at the shift point involves a recombination of the levels of the dimensions to form a new compound, no evidence of discrimination is obtained.

#### Time-to-Habituation Scores

Time-to-habituation (TH) was defined as the number of minutes spent in the preshift phase of the experiment, excluding the baseline minute.

The TH scores for all groups of subjects that were presented with one sound during the preshift

phase are shown in the left-hand side of Figure 4. Subjects in the single-sound conditions were grouped according to the specific stimulus that was presented. There was one group for each of the four stimuli used in the experiment, /a/-monotone, /a/-falling, /i/-monotone, and /i/-falling. While the trend was toward longer TH for the /a/ vowel than for the /i/ vowel and slightly longer for the falling contour than the monotone contour, a two-way ANOVA (vowels  $\times$  pitch contours) did not approach significance at the .05 level for either main effect. The single-stimulus conditions were therefore collapsed for further comparisons.

The TH scores for all groups of subjects that were presented with two sounds during the preshift phase of the experiment are displayed in the right-hand side of Figure 4. The three groups who were presented with two sounds during preshift included: (1) those presented with two sounds that shared the vowel dimension (infants tested on the vowel target with pitch variation, both experimental and control groups); (2) those presented with two sounds that shared the pitch dimension (infants tested on the pitch target with vowel variation, both experimental and control groups); and (3) those presented with two sounds that did not share either the vowel or the pitch dimension (the combined groups, both experimental and control).

A one-way analysis of variance comparing the TH scores for these three groups of infants and the TH scores for the infants in all of the single-stimulus groups revealed a significant effect due to groups [ $F(3,76) = 3.59$ ,  $p < .05$ ]. Newman-Keuls post hoc

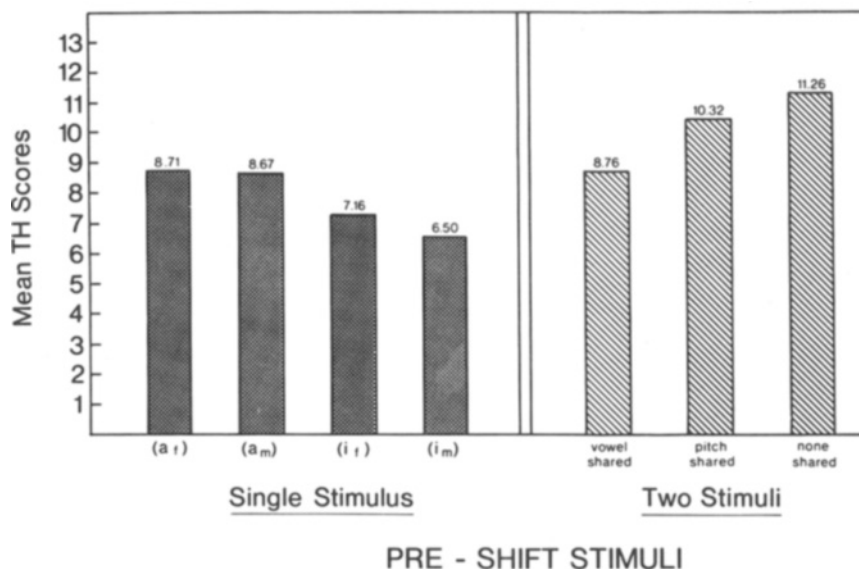


Figure 4. The average time-to-habituation (TH) scores for infants presented with one of the four *single* stimuli and infants presented with one of the three *two-stimuli* pairs during the preshift phase of the experiment.



comparisons revealed that the "pitch-shared" and the "no-dimensions-shared" group differed significantly from both the "vowel-shared" group and the "single-stimulus" group. However, no other significant differences were obtained; neither the former two groups nor the latter two groups differed significantly from one another.

In summary, the time-to-habituation measure revealed that infants took longer to habituate when two stimuli were presented than when only a single stimulus was presented, but that this effect was due to the pitch-shared and no-dimensions-shared groups. The vowel-shared group did not differ significantly from the average time to habituation shown by the single-stimulus groups.

## DISCUSSION

This study of 4- to 16-week-old infants examined and infant's ability to discriminate speech stimuli when a change in one of two dimensions, vowel identity or pitch contour, occurred by itself or against a background of constant variation along the second dimension. Results demonstrated that when either the dimension of vowel identity or of pitch contour occurred independently of any ongoing variation in a second dimension, infants demonstrated the ability to make the discrimination. However, when a change in the target dimension was accompanied by constant variation in the second dimension, discrimination was not always evidenced. More specifically, when a change in the vowel dimension occurred against a background of constant variation in the pitch dimension, discrimination was evidenced. In contrast, when a change in the pitch dimension occurred against a background of constant variation in the vowel dimension, discrimination was not evidenced.

In addition to these results involving the sucking-recovery scores, an analysis of the time-to-habituation data revealed that infants took significantly longer to meet the habituation criterion when two sounds were presented during the preshift period rather than one. However, this result was shown to be asymmetric; it was dependent upon the specific dimension that the two sounds had in common. When the two sounds shared the vowel dimension, time to habituation was not significantly greater than it was in any condition in which a single sound was presented. In contrast, when the two sounds shared the pitch dimension or shared neither of the two dimensions, as in the combined condition, time to habituation was significantly longer than when a single sound was presented or when two sounds that shared the vowel dimension were presented.

The sucking recovery scores and the time-to-

habituation scores complement each other in that evidence of discrimination (based on sucking recovery) was never obtained in cases in which the time-to-habituation scores were significantly greater than those obtained when a single sound was presented during the preshift period. These data are relevant to developmental models of speech perception and to developmental models of attention and memory.

### Habituation and Memory

There are several models of habituation that have been proposed by investigators studying visual perception (Cohen, 1973; Fagan, 1977; Lewis, 1971; McCall, 1971; McCall & McGhee, 1977; Olson, 1976; Sokolov, 1963, 1969). While they differ in some significant respects, each model assumes that in order to demonstrate habituation, the infant must store a composite description of (i.e., "remember") the original stimulus and eventually recognize that the stimulus being presented matches the stored representation. The recognition of the similarity between the stimulus being presented and its stored representation reduces the infant's interest in the stimulus, and therefore results in a decrease in the response being measured. In addition, each model argues that dishabituation depends upon the infant's recognition that the second stimulus is novel, that is, that it does not match the stored representation; this, in turn, increases the infant's interest in the stimulus, and recovery of the response is demonstrated.

Eimas et al.'s (1971) original account of the outcome of the HAS experiments using auditory stimuli was similar. They argued that the infant's decreased willingness to produce high-amplitude sucking responses in the preshift period of the experiment was due to the infant's recognition that the stimulus was not novel and the infant's eventual disinterest in that stimulus. Similarly, the infant's renewed willingness to produce high-amplitude sucking responses in the postshift phase of the experiment was attributed to the infant's recognition that the stimulus was novel and to the infant's interest in that novel stimulus.

This account of the infant's behavior is particularly interesting when the experiment is designed to present more than a single stimulus during the pre- and postshift phases of the experiment. Presumably, the addition of stimuli to the set that are discriminably different should make the task of forming a composite of (i.e., "remembering") the stimuli presented during preshift more difficult, and this should tend to prolong the habituation period. However, if the stimuli presented are similar along one dimension, and the infant is capable of recognizing that similarity, this might serve as an aid to memory, thus decreasing TH scores.

### **The Perception of Similarity and Its Effect on Habituation: Analysis of the TH Scores**

The average TH scores increased when two stimuli were presented during the preshift period rather than one, regardless of the similarity between the two stimuli, suggesting that, in general, it takes longer for the infant to remember two stimuli than to remember one. However, the effect was due to two of the three conditions. TH scores were significantly greater when the two stimuli presented during the preshift period shared the pitch dimension or shared no dimensions, but not significantly greater when the two stimuli shared the vowel dimension. This finding suggests that the shared vowel dimension provided an aid to memory, presumably because the infant was able to recognize the similarity between the two stimuli along the vowel dimension, thus reducing the memory load. A shared dimension did not always make two stimuli easier to remember, though, since, when the two stimuli shared the pitch dimension, a significant increase in TH scores occurred, such that it took almost as long to habituate when the two stimuli shared the pitch dimension as when the two stimuli had nothing in common. One could argue, therefore, that the infants did not recognize the similarity along the pitch dimension, or at least that they could not do so as easily as they could recognize the similarity along the vowel dimension.

These results are similar to those obtained in visual perception studies in infants. For example, Bornstein, Kessen, and Weiskopf (1976) demonstrated differential rates of habituation to visual stimuli varying in wavelength. They contrasted habituation under three conditions: (1) when a single stimulus was presented; (2) when two stimuli varied in wavelength but were both perceived as members of a single color category; and (3) when two stimuli varied in wavelength by the same amount as the second pair but were perceived by adults as members of two different color categories. The rate of habituation was significantly shorter when a single stimulus, or when two stimuli perceived by adults as members of the same color category, were presented than it was when the stimuli did not fall into a single color category, even though the differences in wavelength between the two stimuli were identical in both cases. The authors interpreted these data to indicate that the rate of habituation was dependent upon the infant's perception of similarity, and further, that the perception of similarity by infants served as an aid to memory.

The results from this study are interpreted along similar lines. Infants demonstrated shorter TH scores when two stimuli sharing the vowel dimension were presented, leading to the suggestion that infants recognize the similarity between the stimuli and that doing so serves as an aid in remembering the stimuli. How-

ever, they did not tend to recognize the perceptual similarity among different vowels whose pitch contours were identical. They behaved as though the stimuli had nothing in common, taking equally long to remember the stimuli in both cases.

### **The Discrimination of Changes in Multidimensional Stimuli: Analysis of the Sucking-Recovery Scores**

On the basis of the TH data, we argued that it was easier to remember two stimuli whose vowel identities were similar but whose pitch contours differed than it was to remember two stimuli whose pitch contours were identical but whose vowel identities differed. Nonetheless, the HAS model predicts that, once habituated, the infants "remember" these stimuli equally well and will be equally capable of increasing their sucking responses if a stimulus is presented that is both discriminable to the infants and sufficiently interesting to them.

The sucking recovery data indicate that only three of the five experimental groups demonstrated significant sucking-recovery scores when compared with their respective control groups. Infants discriminated both the vowel change and the pitch change when it occurred in the absence of constant variation in the second dimension, and the vowel change in the presence of constant variation in the pitch contours of the vowels. There are at least two plausible hypotheses that would account for these results, the second one more complicated and interesting than the first. The first explanation argues simply that infants who took a long time to habituate in the preshift phase of the experiment were either too fatigued or too disinterested to dishabituate in the postshift phase. In both conditions in which TH scores were significantly longer, no recovery was evidenced, so the hypothesis that infants are not capable of showing recovery after a long habituation phase cannot be ruled out. On the other hand, individual infants in each of the five experimental groups took longer than 11 min to habituate, and many of these infants showed recovery, while none of the infants in either of the two conditions that failed to show significance overall demonstrated substantial recovery.

A second hypothesis relates the failure to show sucking recovery to the specific choice of the two levels representing each dimension. In the pitch target condition in which variation in the second dimension was presented, one can hypothesize that the failure to discriminate is due to the infants' tendency, given these stimuli, to attend to the vowel dimension rather than the pitch dimension. This explanation holds that in a task such as this, the two dimensions, vowel and pitch, compete for the infant's attention, and when the levels chosen to represent the vowel dimension are values appropriate for the vowels /a/

and /i/ and the levels chosen to represent the pitch dimension are values appropriate to a falling and a monotone pitch contour, the levels of the vowel dimension are more salient, because they are more discriminable; therefore, constant change in the vowel dimension tended to distract the infant's attention from the pitch dimension. This is suggested both by the TH and the sucking-recovery scores. When the vowel dimension served as the target, the sucking recovery scores showed that it was always discriminated, even when pitch contour served as a potential distractor. Furthermore, when the vowel was held constant during the preshift, TH scores were low, indicating the ease with which the infants attended to and remembered stimuli whose vowels were identical. Pitch contour, on the other hand, was discriminated only in isolation, and not when a constant change in the vowel served as a distractor, presumably because of the infant's tendency to attend to the highly discriminable difference between the varying vowels /a/ and /i/ rather than to the pitch contour. The less salient pitch dimension changed at the shift point, but went unnoticed amidst the constant change in the vowel, which the infants were then habituated to. This explanation would correctly predict that the TH scores should increase dramatically whenever the vowel was constantly changed, as they did in the condition in which pitch was the target dimension and vowel identity was the second dimension, presumably because of the infant's tendency to attend to the vowel dimension and the difficulty in remembering two different values of the vowel. This explanation would also account for the longer TH scores and the failure to show sucking recovery in the combined condition. In this condition, the vowel dimension was again constantly changed during the preshift and postshift phases; by our previous argument, this constant change served to increase the TH scores during the preshift, and once habituated to the vowel dimension infants did not show sucking recovery because the same levels of the vowel dimension were again presented. The combination of a specific level of the vowel dimension and a specific level of the pitch dimension changed, but this change went unnoticed (see below for further discussion of the combined condition).

One might be tempted to argue that this tendency to attend to the vowel dimension was due to a *general* precedence effect for the acoustic cues underlying vowel identity over the acoustic cues underlying pitch perception. However, given that the levels chosen to represent the two dimensions were not systematically varied, the data are more reasonably interpreted in terms of a *specific* effect that is attributed to the acoustic difference between the levels chosen to represent each dimension. If the acoustic difference between the two levels is the critical variable dic-

tating the infant's attention, one might adjust the "attentional pull" towards either of the two dimensions by choosing values that make the two levels representing each dimension more or less discriminable. A recent paper by Carrell, Smith, and Pisoni (1981), reporting data on adult listeners using a speeded-classification design and stimuli similar to those employed here, supports this notion. It demonstrated that the latency of response for vowel classification by adults was unappreciably altered by random variation in the pitch of the vowels, but that the latency of response for pitch classification was significantly increased by random variation in the vowel. However, when the acoustic differences related to vowel identity were reduced, making them less discriminable, and the acoustic differences related to the perceived pitch were increased, making them more discriminable, these asymmetric differences disappeared and a pattern of mutual interference, similar to that reported by J. L. Miller (1978) for vowels and pitch, was seen.

We would predict, therefore, that if these experiments on infants were repeated using vowels that were more similar acoustically, the two dimensions would show a pattern of mutual interference. However, if we repeated the experiment with the same two vowels, /a/ and /i/, but with pitch contours typical of those produced by mothers speaking to their infants ("Motherese") (Fernald, Note 2), then the tendency shown in this experiment would be reversed, and pitch contour might become the more salient dimension. Recent experiments in our laboratory (Fernald, Note 3; Fernald and Kuhl, Note 4) demonstrate that 4-month-olds "preferred" auditory signals produced by mothers when speaking to their infants to signals produced by those same mothers when talking to adults. Given a two-choice task in which both right and left head-turn responses were followed by sound presentation, infants turned significantly more often in the direction required to produce infant-directed speech, and continued to do so when both signals were represented only by their fundamental frequencies. This would tend to indicate that fundamental frequency can be an extremely salient acoustic cue for infants at this age. It is possible, therefore, that with further experiments we could demonstrate that the young infant's attention is pulled toward the signals typically produced by caretakers when directing speech to their infants, because caretakers dramatically alter the pitch dimension in ways that make it more salient.

#### **The Perception of Individual Dimensions (Components) vs. Compounds**

Another issue addressed by this study is the infant's tendency to represent multidimensional stimuli as independent dimensions, or components (such

as vowel identity and pitch contour), as opposed to their tendency to represent stimuli as integrated "wholes," or compounds. The combined condition was designed to investigate whether infants detected a change in the stimulus set when that change involved a recombination of the levels of the dimensions to form novel compounds rather than the introduction of a new level of a single dimension. Recall that, in this condition, two stimuli were initially presented, each composed of one level of each of the two dimensions. At the shift point, the levels of each dimension were recombined to form a novel composite without introducing any new components. If infants represent the stimuli as compounds, or "wholes," evidence of discrimination should be obtained; if the stimuli are represented as individual dimensions, evidence of discrimination should not be obtained, since no change in the individual dimensions occurred. Infants in this condition failed to demonstrate recovery and therefore provided no evidence in support of the idea that the infants represented these stimuli as "wholes." However, given the fact that only a single condition in this experiment addressed the issue, further experiments should be completed before asserting that infants at this age either are incapable of representing these stimuli as auditory compounds or have a greater tendency to represent these stimuli in terms of their individual components.

Studies of infant visual perception (Cohen, in Cohen & Gelber, 1975) have shown that when infants were randomly presented with two visual stimuli, a red circle and a green triangle, and then tested with (1) the same stimuli, (2) stimuli formed by a recombination of the color and form dimension that created a novel compound (such as a red triangle and a green circle, or (3) completely new compounds, such as a blue square and a yellow dumbbell, discrimination was evidenced only in the final condition. Infants did not demonstrate discrimination of a recombination of components.

In contrast, however, J. L. Miller and Eimas (1979) demonstrated that 2- to 4-month-olds recognized a recombination of consonants and vowels that formed new syllables. For example, infants were presented with a random sequence of the syllables /ba/ and /dae/ during the preshift phase of the experiment and the syllables /bae/ and /da/ during the postshift phase. The sucking-recovery scores demonstrated that infants were capable of detecting the change when it involved the recombination of consonants and vowels. In a second experiment, Miller and Eimas extended these findings to a situation in which the infants were required to detect a recombination at the distinctive-feature level. Infants were presented with the syllables /ba/ and /ta/ prior to the shift-point and the syllables /da/ and /pa/ after, thus re-

quiring the infants to detect a change in the combination of the levels of the voicing and place features. Again, discrimination was evidenced.

One explanation for the infants' failure to demonstrate discrimination in this experiment while succeeding in J. L. Miller and Eimas's (1979) study is that one segmental and one suprasegmental dimension were employed here while two segmental dimensions were employed by Miller and Eimas (1979). However, in a follow-up study (Eimas & J. L. Miller, in press), these authors present data suggesting that infants also discriminated a recombination of dimensions when one was segmental and the other suprasegmental. In this case, however, the segmental dimension was represented by a consonantal feature rather than a vowel feature, so one might still attribute the difference between the two studies to either the inherent differences between consonants and vowels, experimentally verified in many studies on adult listeners (see, for example, J. L. Miller, 1978) or to the specific choices of the levels for the vowel dimension and the pitch-contour dimension. Further work employing other vowel and pitch-contour stimuli will need to be completed before any resolution to this issue can be reached.

#### **The Perception of Auditory Categories and Developmental Speech Perception**

There are two reasons why investigators of infant audition are interested in experiments requiring infants to perceive the similarity among stimuli. One has to do with the acquisition of language-specific categories. An examination of the adult literature in speech perception demonstrates that the main problems in the field have revolved around the issues of acoustic-cue invariance (Liberman et al., 1967) and acoustic-cue equivalence (Liberman & Pisoni, 1977). These classic issues in the adult literature naturally led developmental speech researchers towards questions concerning the infant's tendency to group stimuli based on their perceived similarity. Studies of categorical perception (Eimas, 1974, 1975; Eimas et al., 1971) and studies of perceptual constancy (Fodor et al., 1975; Kuhl, 1979b, 1980, in press) provide strong evidence in support of the idea that infants between the ages of 4 and 7 months perceptually group stimuli on the basis of their assignment to phonetic categories as defined by adult listeners, and these findings are important to theories of language development.

Quite aside from the importance of studies like these to an infant's language-specific predispositions, studies such as these may reveal more general features of the infant's perception of similarity and of developmental attention and memory. For example, the vowel studies cited earlier (Kuhl, 1979b, in press; Kuhl & Hillenbrand, Note 1) using 6-month-old in-

fants were aimed at comparing the infant's ability to perform in an operant task for a visual reinforcer when one of three perceptual dimensions along which the stimuli varied predicted the occurrence of a visual reinforcer. Data from those investigations demonstrated that 6-month-olds can "sort" a set of vowels that vary along three dimensions, vowel identity (/a/ and /i/ or /a/ and /ɔ/), pitch contour (rise-fall and rising), and talker (male, female, child) using either of two "sorting rules," vowel identity (Kuhl, 1979b, in press) or pitch contour (Kuhl & Hillenbrand, Note 1). A recent study (C. L. Miller, Younger, & Morse, in press) indicates that they can also use the "talker" dimension to sort stimuli. To date, we cannot state which of these sorting rules is easier for the infant to use, but it is clear that the 6-month-old is relatively facile at using a variety of acoustic dimensions as sorting rules.

The experiment reported here places somewhat different constraints on the infant, the most noteworthy of which is that it does not force the infant to attend to one of the dimensions to obtain reinforcement. The reinforcement in this task is simply in listening to the syllables. The design, nonetheless, tends to reveal which of the dimensions the infant most naturally attends to, and, in this experiment, the infant appeared to attend to the vowel dimension, presumably because of its greater discriminability. We have no idea whether an operant paradigm that provides some kind of external reinforcement, such as that used with the older infants, might induce the infant to attend to the pitch dimension (as represented here) in the presence of the distraction provided by vowel variation. One would suspect, however, based on studies of visual perception, that the infant's ability to focus at will on a particular dimension increases developmentally, due both to the increasing ability to direct attention toward only the relevant features of a display (Vurpillot, 1968; Zinchenko, van Chzhi-Tsin, & Tarakonov, 1963) and to the ability to ignore irrelevant features of a display (see Gibson, 1969, for further discussion). The fact that, by 6 months of age, infants in the vowel experiments demonstrated the ability to attend to either the vowel or the pitch dimension could support such a developmental hypothesis, but only weakly, since the paradigms being compared (HAS vs. operant head turn) are substantially different. Designing studies that directly investigate the hypothesis that attention to particular auditory dimensions in multidimensional auditory arrays is maturationally or experientially determined will be challenging. Unlike our colleagues in vision, who simply observe corneal reflections and eye movements, the lack of "ear-lids" and the absence of a need to make a directionally specific movement toward an auditory stimulus in order to hear it mandates that we design

more intricate experiments to learn what aspects of a multidimensional auditory stimulus the infant is listening to. The present study represents an attempt to provide such a paradigm for auditory studies with infants as young as 4 weeks of age.

#### REFERENCE NOTES

1. Kuhl, P. K., & Hillenbrand, J. *Speech perception in early infancy: Perceptual constancy for categories based on pitch contour*. Paper presented at the Society for Research on Child Development, San Francisco, March 1979.
2. Fernald, A. *Rhythm and intonation in mother's speech to newborns*. Paper presented at the Third Annual Boston University Conference on Language Development, Boston, Massachusetts, September 1978.
3. Fernald, A. *Four-month-olds prefer to listen to "Motherese."* Paper presented at the Biennial meeting of the Society for Research in Child Development, Boston, April 1981.
4. Fernald, A., & Kuhl, P. K. *Fundamental frequency as an acoustic determinant of infant preference for Motherese*. Paper presented at the biennial meeting of the Society for Research in Child Development, Boston, April 1981.

#### REFERENCES

- BORNSTEIN, M. H., KESSEN, W., & WEISKOPF, S. The categories of hue in infancy. *Science*, 1976, **191**, 201-202.
- BOWER, T. G. R. Heterogeneous summation in human infants. *Animal Behaviour*, 1966, **14**, 395-398.
- CARREL, T. D., SMITH, L. B., & PISONI, D. B. Some perceptual dependencies in speeded classification of vowel color and pitch. *Perception & Psychophysics*, 1981, **29**, 1-10.
- COHEN, L. B. A two-process model of infant visual attention. *Merrill-Palmer Quarterly*, 1973, **19**, 157-180.
- COHEN, L., DELOACHE, J., & STRAUSS, M. Infant visual perception. In J. Osofsky (Ed.), *Handbook of infancy*. New York: Wiley, 1979.
- COHEN, L. B., & GELBER, E. R. Infant visual memory. In L. Cohen & P. Salapatek (Eds.), *Infant perception: From sensation to cognition* (Vol. 1): *Basic visual processes*. New York: Academic Press, 1975.
- DUNN, H. K. Acoustic characteristics of vowels. In *Engineering Summer Conference on Automatic Recognition*. Ann Arbor: University of Michigan Press, 1963.
- EILERS, R. E. Infant speech perception: History and mystery. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology* (Vol. 2): *Perception*. New York: Academic Press, 1980.
- EIMAS, P. D. Auditory and linguistic processing of cues for place of articulation by infants. *Perception & Psychophysics*, 1974, **16**, 513-521.
- EIMAS, P. D. Auditory and phonetic coding of the cues for speech: Discrimination of the /r-/l/ distinction by young infants. *Perception & Psychophysics*, 1975, **18**, 341-347.
- EIMAS, P. D., & MILLER, J. L. Organization in the perception of segmental and suprasegmental information by infants. *Infant Behavior and Development*, in press.
- EIMAS, P. D., SIQUELAND, E. R., JUSCZYK, P., & VIGORITO, J. Speech perception in infants. *Science*, 1971, **171**, 303-306.
- EIMAS, P. D., & TARTTER, V. C. On the development of speech perception: Mechanisms and analysis. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behavior* (Vol. 13). New York: Academic Press, 1979.
- FAGAN, J. F. An attentional model of infant recognition. *Child Development*, 1977, **48**, 345-359.

- FODOR, J. A., GARRETT, M. F., & BRILL, S. L. Pi ka pu: The perception of speech sounds by prelinguistic infants. *Perception & Psychophysics*, 1975, **18**, 74-78.
- GIBSON, E. J. *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts, 1969.
- HILLENBRAND, J. *Perceptual organization of speech in early infancy*. Unpublished doctoral dissertation, University of Washington, 1980.
- JUSCZYK, P. W. Infant speech perception: A critical appraisal. In P. D. Eimas & J. L. Miller (Eds.), *Perspectives on the study of speech*. Hillsdale, N.J.: Erlbaum, 1981.
- KUHL, P. K. Speech perception in early infancy: The acquisition of speech-sound categories. In S. K. Hirsh, D. H. Eldredge, I. J. Hirsh, & S. R. Silverman (Eds.), *Hearing and Davis: Essays honoring Hallowell Davis*. St. Louis: Washington University Press, 1976.
- KUHL, P. K. The perception of speech in early infancy. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 1). New York: Academic Press, 1979. (a)
- KUHL, P. K. Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America*, 1979, **66**, 1668-1679. (b)
- KUHL, P. K. Perceptual constancy for speech-sound categories. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology* (Vol. 2): *Perception*. New York: Academic Press, 1980.
- KUHL, P. Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development*, in press.
- LASKY, R. E., SYRDAL-LASKY, A., & KLEIN, R. E. VOT discrimination by four-to-six-and-a-half-month-old infants from Spanish environments. *Journal of Experimental Psychology*, 1975, **20**, 215-225.
- LEWIS, M. Individual differences in the measurement of early cognitive growth. In J. Hellmuth (Ed.), *Exceptional infant* (Vol. 2): *Studies in abnormalities*. New York: Brunner/Mazel, 1971.
- LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P., & STUDDERT-KENNEDY, M. Perception of the speech code. *Psychological Review*, 1967, **74**, 431-461.
- LIBERMAN, A. M., & PISONI, D. B. Evidence for a special speech-perceiving subsystem in the human. In T. H. Bullock (Ed.), *Recognition of complex acoustic signals*. Berlin: Abakon Verlagsgesellschaft, 1977.
- MCCALL, R. B. Attention in the infant: Avenue to the study of cognitive development. In D. Walchur & D. Peters (Eds.), *Early childhood: The development of self-regulatory mechanisms*. New York: Academic Press, 1971.
- MCCALL, R. B., & MCGHEE, P. E. The discrepancy hypothesis of attention and affect in infants. In F. Weizman & I. C. Uzgiris (Eds.), *The structuring of experience*. New York: Plenum, 1977.
- MILLER, C. L., YOUNGER, B. A., & MORSE, P. A. The categorization of male and female voices in infancy. *Infant Behavior and Development*, in press.
- MILLER, J. L. Interactions in processing segmental and supra-segmental features of speech. *Perception & Psychophysics*, 1978, **24**, 175-180.
- MILLER, J. L., & EIMAS, P. D. Organization in infant speech perception. *Canadian Journal of Psychology*, 1979, **33**, 353-367.
- MORSE, P. A. Infant speech perception: Origins, processes, and alpha centauri. In F. D. Minifie & L. L. Lloyd (Eds.), *Communicative and cognitive abilities—Early behavior assessment*. Baltimore: University Park Press, 1978.
- OLSON, G. M. An information-processing analysis of visual memory and attention in infants. In T. J. Tighe & R. N. Leaton (Eds.), *Habituation: Perspectives from child development, animal behavior, and neurophysiology*. Hillsdale, N.J.: Erlbaum, 1976.
- PETERSON, G. E., & BARNEY, H. L. Control methods used in a study of the vowels. *Journal of the American Speech Association*, 1952, **24**, 175-184.
- RABINER, L. R. Digital-formant synthesizer for speech studies. *Journal of the Acoustical Society of America*, 1968, **24**, 175-184.
- SOKOLOV, E. N. *Perception and the conditioned reflex*. New York: Macmillan, 1963.
- SOKOLOV, E. N. The modeling properties of the nervous system. In M. Coles & I. Maltzman (Eds.), *A handbook of contemporary Soviet psychology*. New York: Basic Books, 1969.
- SPENNER, B. F., ENGBRETSON, A. M., MILLER, J. D., & COX, J. R. Random-access programmable recorder of complex sounds (RAP): A digital instrument for auditory research. *Journal of the Acoustical Society of America*, 1974, **55**, 427.
- STREETER, L. Language perception shows effects of both innate mechanisms and experience. *Nature*, 1976, **259**, 39-41.
- VURPILLOT, E. The development of scanning strategies and their relation to visual differentiation. *Journal of Experimental Child Psychology*, 1968, **6**, 622-650.
- WINER, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1971.
- ZINCHENKO, V. P., VAN CHZHI-TSIN, & TARAKONOV, V. V. The formation and development of perceptual activity. *Soviet Psychology and Psychiatry*, 1963, **2**, 3-12.

## NOTE

1. The term "habituation" is used only because it facilitates a comparison of the results obtained here with those obtained in visual experiments and with the habituation models proposed to account for the results of visual experiments. The effect could more accurately be called "satiation," since that terminology would appropriately focus on stimulus effects rather than any physiological correlates of the change in response. While we choose to use the term "habituation," no physiological metaphor is intended.

(Manuscript received June 17, 1980;  
revision accepted for publication November 4, 1981.)