

# Syllabic pitch perception in 2- to 3-month-old infants

ROANNE G. KARZON

*Central Institute for the Deaf, St. Louis, Missouri*  
*and Washington University School of Medicine, St. Louis, Missouri*

and

JOHANNA G. NICHOLAS

*Central Institute for the Deaf, St. Louis, Missouri*  
*and Washington University, St. Louis, Missouri*

The pitch patterns present in speech addressed to infants may play an important role in perceptual processing by infants. In this study, the high-amplitude sucking procedure was used to assess discrimination by 2- to 3-month-old infants of rising versus falling pitch patterns in 400-msec synthetic [ra] and [la] tokens. The syllables' intonation contour was modeled on infant-directed speech, and covered a range characteristic of an adult female speaker (180-300 Hz). Group data indicated that the 2- to 3-month-old infants discriminated the pitch contour for both stimuli. Results are discussed with reference to previous studies of syllabic pitch perception.

It has been hypothesized that the pitch patterns present in speech addressed to infants may serve to facilitate perceptual processing by highlighting important distinctions within the sound sequence (Fernald, 1982; Garnica, 1977; Karzon, 1985), attaining or maintaining infant attention (Blount & Padgug, 1976; Chapman, 1981; Fernald, 1978; Garnica, 1977; Sachs, 1977; Sachs, Brown, & Salerno, 1976; Snow, 1977), and/or segmenting or organizing the speech stream for the infant (Fernald, 1978). The ability of infants to process pitch information has been studied from several perspectives. Pitch experiments may be grouped into two classes on the basis of stimulus characteristics: (1) studies that use tonal stimuli to assess sound frequency resolution (e.g., Berg, 1972; Olsho, 1984, 1985; Olsho, Schoon, Sakai, Turpin, & Sperduto, 1982; Wormith, Pankhurst, & Moffitt, 1975) or to test discrimination of tone sequences (e.g., Bundy, Colombo, & Singer, 1982; Trehub, Bull, & Thorpe, 1984; Trehub, Thorpe, & Morrongiello, 1985), and (2) studies that employ speech stimuli to assess perception of syllabic pitch (Eimas & Miller, 1981; Kuhl & Hillenbrand, 1979; Kuhl

& Miller, 1982; Morse, 1972; Sullivan & Horowitz, 1983) or intonation contour (Kaplan, 1970a, 1970b). Our study follows from the experiments in the second class.

At least three previous investigators have reported results related to infants' perception of rising versus falling intonation patterns. Kaplan (1970a, 1970b) measured heart rate (HR) and other behaviors and reported that 8-month-old infants, but not 4-month-old infants, demonstrated discrimination of rising versus falling intonation contours in the naturally spoken sentence "see the cat." Based on results of a high-amplitude sucking (HAS) experiment, Morse (1972) reported that 2-month-old infants could discriminate rising versus falling [ba] intonation contours. Sullivan and Horowitz (1983) used an infant-control auditory preference paradigm and found evidence that infants attended more to a rising intonation contour for naturally produced [ba] tokens, but reversed and attended more to the falling contours for the synthetic [ba] tokens. It should be noted that the differential attention to one token over another directly implies that the infants discriminated the two stimuli.

In the three studies mentioned above, acoustic variables varied in several important dimensions. Kaplan's (1970a, 1970b) stimuli were three syllables, whereas Morse (1972) and Sullivan and Horowitz (1983) used the single syllable [ba]. Acoustic specifications were not reported for Kaplan (1970a, 1970b). The 500-msec synthetic [ba] tokens of Morse (1972) rose from 120 to 194 Hz ( $\Delta F = +74$  Hz) and fell from 120 to 70 Hz ( $\Delta F = -50$  Hz). The synthetic [ba] stimuli of Sullivan and Horowitz rose from 140 to 225 Hz ( $\Delta F = +85$  Hz) and fell from 150 to 125 Hz ( $\Delta F = -25$  Hz).<sup>1</sup> Thus, it is possible that infants discriminated the rising and falling contours on the basis of differences in absolute frequency range or extent of frequency excursion. Another variable that was

---

This work was supported in part by Grant NS03856 from the National Institute of Neurological and Communicative Disorders and Stroke to the Central Institute for the Deaf. The Department of Otolaryngology of Washington University School of Medicine and the Department of Psychology of Washington University are acknowledged for support provided during data analysis and manuscript preparation. The authors wish to express their appreciation to the hospitals, mothers, and infants who participated. The nursery staff and administration of St. John's Mercy Medical Center were especially helpful throughout the course of the study. Special thanks are extended to James D. Miller, Judith Lauter, Margaret Skinner, Ira J. Hirsh, and Nelson Cowan for their assistance in the preparation of the manuscript. Address reprint requests to Roanne Karzon, Department of Otolaryngology, Audiology Division, Washington University School of Medicine, 4911 Barnes Hospital Plaza, St. Louis, MO 63110.

**Table 1**  
Center Frequencies and Bandwidths of Formants

Speech Sound	F1	F2	F3	F4	B1	B2	B3	B4
[l]	380	1200	3400	4200	50	100	280	250
[r]	380	1200	1580	4200	70	100	120	250
[a]	900	1350	2850	4200	190	70	160	250

Note—F2, F2, F3, and F4 are the center frequencies and B1, B2, B3, and B4 are the bandwidths associated with formants 1, 2, 3, and 4, respectively. Values represent steady-state portions of the phonemes noted.

not controlled was the decibel equivalence of the tokens. Given the lack of control for several acoustic dimensions, infants could have used any of several acoustic cues to make the discrimination in any of these three previous experiments.

The present study was designed to extend the findings reported by Morse (1972) and Sullivan and Horowitz (1983) on single-syllable intonation to new phonemes ([ra] and [la]), combined with methodological controls on stimulus characteristics, such as the decibel equivalence of the rising versus falling intonation contours, the extent and range of frequency excursion, and the fundamental frequency configuration. The test syllables were produced synthetically, according to characteristics of the female voice and incorporating the intonation range of infant-directed speech (Karzon, 1982, 1985).

## METHOD

### Stimuli

All sounds for the experiments were synthesized using the Klatt Synthesizer and were digitally stored in disks for playback in a digital record and playback device described elsewhere (Spencer, Engebretson, Miller, & Cox, 1974). Stimuli were processed and presented in the same manner as in Karzon (1985).

The [ra] and [la] tokens were synthesized using Klatt's (1980) model in conjunction with available subroutines written at the Central Institute for the Deaf. Synthesis values were selected to be appropriate for a female voice. All tokens were 400 msec in duration, and began with formants in a steady state for the consonant and ended with formants in a steady state for the vowel /a/. The target values for center frequency and bandwidth of the formants for the steady-state portion of each experimental speech sound are shown in Ta-

ble 1. The pitch of the rising-contour syllables began at 180 Hz for 100 msec and rose linearly to 300 Hz over the final 300 msec. The pitch for the falling-contour syllables began at 300 Hz for 100 msec and fell linearly to 180 Hz over the final 300 msec. The pitch change began during the transition from consonant to vowel.

The rising and falling tokens, checked for peak decibel equivalence using a Daven VU meter, were within .4 dB of each other for both sets of phonemes. In addition, informal listening by three adults indicated that the sounds within a set were perceived as equally loud. The amplitude of each syllable peaked at approximately 73 dB SPL at the site of the infant's head.

### Subjects

Infants were selected from well-baby nurseries in the St. Louis region according to the following five criteria: (1) Apgar score of at least 7 at 5 min, (2) birth weight between 2,500 and 4,000 g, (3) gestation between 38 and 42 weeks, (4) permission granted by the attending pediatrician, and (5) no medical history of problems known to be correlated with hearing impairment. Parents came from a wide variety of ethnic and socioeconomic backgrounds. Mothers wishing to participate brought their infants to the laboratory for one session, arranging an appointment time at which the infant would be "awake and happy."

Of the 176 infants tested, 64 provided usable data. Infants were excluded for the following reasons: crying ( $n = 70$ ), sleeping ( $n = 18$ ), low sucking rate ( $n = 8$ ),<sup>2</sup> failure of the sucking rate to rise above baseline during the preshift portion of the experimental session ( $n = 5$ ), refusal to suck the pacifier ( $n = 7$ ), equipment failure ( $n = 2$ ), and experimenter error ( $n = 2$ ). Within each experiment, infants were assigned at random to either the control or the experimental condition. Sex and age data for infants included in each condition are shown in Table 2.

### Apparatus

The subjects were tested in a single-wall sound booth with ambient noise of approximately 26 dBA at the site of the infant's head. An adjacent control room contained the instrumentation for controlling stimulus presentation and response measurement appropriate to a HAS test paradigm. The apparatus employed was identical to that of Karzon (1985), in which a detailed description may be found for the following: blind nipple and transducer, logic device, graphic level recorder, crib, two-way talk and masking system, and television monitor.

### Procedure

The experimental protocol employed was similar to that of Kuhl (1976). After a brief introduction to test procedures, the mother

**Table 2**  
Infant Data

	N	Age (in Days)		Range	Sex		Sucking Rate			
		M	SD		M	F	Preshift Covariate		Postshift Dependent Variable	
							M	SD	M	SD
[la] Token										
Experimental Group	16	76.3	12.2	58-96	9	7	33.8	17.6	42.3	21.7
Control Group	16	75.9	14.4	56-99	3	13	27.6	10.2	25.4	13.5
[ra] Token										
Experimental Group	16	73.4	11.5	61-96	8	8	38.3	24.3	41.3	26.0
Control Group	16	69.6	9.4	58-96	6	10	30.1	12.2	29.8	12.6

Note—The covariate value is the average preshift sucking rate based on the 2 min prior to the stimulus shift point, that is, the point at which >20% decrease in sucking rate was observed. The dependent variable is the average postshift sucking rate based on the 2 min following the stimulus shift point.

placed the infant in the seat within the wooden crib. An assistant inserted and held the nipple in the infant's mouth. Once the infant began sucking, the experimenter adjusted the threshold setting of the logic device to establish the criterion for a high-amplitude suck. After adjustment, the sucking rate was monitored for 1 min to ensure that a baseline of 20 to 35 high-amplitude sucks per minute had been met. A 750-msec "stimulus time window" was selected so that once a high-amplitude suck triggered a stimulus presentation of 400 msec, another sound could not be presented for the remaining duration of the selected window (i.e., for 350 msec). Stimulus presentation of rising and falling contours was counterbalanced within both the experimental and control groups for both [ra] and [la] tokens.

The initial preshift stimulus was presented until the rate of high-amplitude sucking dropped for 2 consecutive minutes to a value 20% below the maximum rate achieved in any of the previous minutes of sound presentation. After this habituation criterion, provided the infant had been sucking for at least 4 min, the logic device activated the postshift stimulus. Experimental subjects received a novel stimulus during postshift minutes; control subjects were presented the same stimulus as in the preshift period. After 4 min of postshift monitoring, the session was terminated.

The experiment had two experimental conditions: (1) [la] with a rising intonation contour versus [la] with a falling intonation contour, and (2) [ra] with a rising intonation contour versus [ra] with a falling intonation contour. Each subject listened to either [la] or [ra]. The presentation of rising or falling contours for the initial preshift stimulus was counterbalanced across subjects. For each experimental condition, there was a corresponding control condition in which the initial preshift stimulus did not change.

## RESULTS

The interpretation that the dependent variable of sucking rate provides an index for discrimination of the stimuli presumes that the performance of the experimental and control groups does not differ significantly in the preshift portion of the experiment. Two response measures for the preshift portion of the experiment were also analyzed with a one-way analysis of variance: (1) the sucking rate during the baseline minute prior to sound presentation and (2) the mean of the response rate during the last 2 preshift minutes. No significant differences were observed on either of these measures. Since both control and experimental subjects received the same treatment during the preshift portion of the experiment, no group differences were expected in these two preshift response measures.

In analyzing whether or not infants responded to the rising versus falling stimuli, the dependent variable was the number of sucks per minute, averaged across the 2 min immediately following the stimulus shift point. A 2 (experimental/control)  $\times$  2 (stimulus type) analysis of covariance (ANCOVA) was used to analyze the scores. Number of sucks per minute, averaged across the 2 min immediately preceding the stimulus shift point, was used as the covariate. The mean sucking rates in the preshift and postshift periods for the experimental and control groups are shown in Table 2.

The analysis revealed an overall  $F(4,59) = 47.09$  ( $p < .0001$ ). A main effect was found for experimental versus control group assignment [ $F(1,59) = 6.49$ ,

$p < .0135$ ]; that is, infants who heard the stimulus change had significantly higher postshift sucking rates than infants who heard the same stimulus in the postshift period. However, no main effect was found for stimulus type (i.e., [ra] versus [la]). The effect of the interaction between stimulus type and experimental/control group assignment was not significant.<sup>3</sup>

## DISCUSSION

These results, combined with those reported by Morse (1972) and Sullivan and Horowitz (1983), provide evidence that by 2 to 3 months of age, infants can discriminate rising and falling syllabic pitch contours in some phonemic contexts. It should be noted that the infants discriminated the rising and falling contours in the [la] and [ra] tokens on the basis of frequency contour alone. Presumably, this is a more difficult task than that presented by the stimuli used by Morse and Sullivan and Horowitz, which varied with respect to extent of frequency excursion, absolute frequency range, and possibly amplitude.

There is at least one example in the literature on infant speech perception in which successful discrimination performance has been demonstrated with HAS but not with HR. A number of authors using HAS have shown that 1- to 4-month-old infants can discriminate [ba] versus [pa] (Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Streeter, 1976; Trehub & Rabinovitch, 1972); however, HR procedures failed to provide evidence of the same phonemic discrimination (Miller & Hankes-Ruzicka, 1978; Roth & Morse, 1975).

It is possible that Kaplan (1970a, 1970b) failed to observe discrimination in the 4-month-old infants due to methodological limitations. In the heart rate procedure she employed, experimental infants heard a set number of stimuli with a fixed interstimulus interval prior to presentation of the novel stimulus. Thus, a certain percentage of infants may not have received a sufficient number of tokens or had enough time to habituate prior to the stimulus shift. With the HAS procedure of Morse (1972) and the present study, the sucking response that results in a stimulus presentation requires that the infant play an active role in the experiment (Miller & Hankes-Ruzicka, 1978). The amount of time and number of tokens presented is tailored to the individual infant. A criterion of decreased sucking rate must be observed prior to a switch to the novel stimulus. This requirement of active participation, as well as the habituation criterion of decreased sucking rate, may make the HAS paradigm a more sensitive index of infant discrimination than the HR technique. The infant-control auditory preference paradigm of Sullivan and Horowitz (1983) does rely on a fixed amount of observation time, as does the HR technique. However, it is similar to the HAS paradigm with respect to the infant's control of the stimulus presentation. Thus, the finding of discrimination of intonation at younger ages with the HAS experiments of Morse (1972) and the present study, as well as with the infant-control

auditory preference paradigm, compared with the HR and behavioral observation technique of Kaplan (1970a, 1970b) may be due to methodological factors.

In summary, the present study was designed to investigate infant discrimination of syllabic pitch contrasts under more controlled stimulus conditions (e.g., absolute frequency, extent of excursion, decibel equivalence) than have been used previously. Different phonemes were employed to test the generalizability of previous results. The findings of the present study, along with those of Morse (1972) and Sullivan and Horowitz (1983), suggest that by the age of 2 months, human infants can discriminate falling from rising voice pitch under some conditions in some phonemic contexts.

The advent of natural-sounding synthetic speech and the technology of digital manipulation of natural speech tokens affords the possibility of carrying out parametric studies of intonation discrimination with infants. The present study employed a pitch contrast that to some extent reflected the increased inflection used in speech addressed to infants. To examine the issue of whether infants require a large pitch discrepancy to process intonation and develop speech and language, it would be of interest to successively decrease the extent of the frequency excursion until the infants failed to demonstrate discrimination. Our experience with HAS and reports in the literature suggest that it may be more fruitful to address these issues with older infants using a head turn procedure. In general, the head turn procedure requires fewer subjects per stimulus pair to ascertain discrimination ability. The head turn procedure has the added advantage that infants from a wide age range (i.e., 6 to 18 months of age), as well as adults, may be tested and compared (Olsho, 1984; Olsho et al., 1982; Trehub, Schneider, & Endman, 1980).

#### REFERENCES

- BERG, W. K. (1972). Habituation and dishabituation of cardiac responses in 4-month-old, alert infants. *Journal of Experimental Psychology*, **14**, 92-107.
- BLOUNT, B. G., & PADGUG, E. J. (1976). Mother and father speech: Distribution of parental speech features in English and Spanish. *Papers & Reports on Child Language Development* (Stanford University), **12**, 47-59.
- BUNDY, R. S., COLOMBO, J., & SINGER, J. (1982). Pitch perception in young infants. *Developmental Psychology*, **18**, 10-14.
- CHAPMAN, R. S. (1981). Mother-child interaction in the second year of life. In R. S. Schiefelbusch & D. D. Bricker (Eds.), *Early language: Acquisition and intervention* (pp. 201-250). Baltimore: University Park Press.
- EIMAS, P. D., & MILLER, J. L. (1981). Organization in the perception of segmental and suprasegmental information by infants. *Infant Behavior & Development*, **4**, 395-399.
- EIMAS, P. D., SIQUELAND, E. R., JUSCZYK, P., & VIGORITO, J. (1971). Speech perception in infants. *Science*, **171**, 303-306.
- FERNALD, A. (1978, September). *Rhythm and intonation in mother's speech to newborns*. Paper presented at the 3rd Annual Boston University Conference on Language Development, Boston, MA.
- FERNALD, A. (1982). *Acoustic determinants of infant preference for motherese*. Unpublished doctoral dissertation, University of Oregon, Eugene.
- GARNICA, O. K. (1977). Some prosodic and paralinguistic features of speech to young children. In C. Snow & C. Ferguson (Eds.), *Talking to children: Language input and language acquisition* (pp. 63-88). Cambridge, England: Cambridge University Press.
- KAPLAN, E. L. (1970a). Intonation and language acquisition. *Papers & Reports on Child Language Development* (Stanford University), **1**, 1-21.
- KAPLAN, E. L. (1970b). The role of intonation in the acquisition of language (Doctoral dissertation, Cornell University, 1969). *Dissertation Abstracts International*, **30**, 3407B.
- KARZON, R. G. (1982). Speech discrimination by infants (Doctoral dissertation, Washington University, 1982). *Dissertation Abstracts International*, **43**, 1644B.
- KARZON, R. G. (1985). Discrimination of polysyllabic sequences by one- to four-month-old infants. *Journal of Experimental Child Psychology*, **39**, 326-342.
- KLATT, D. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, **67**, 971-995.
- KUHL, P. (1976). Speech perception in early infancy: The acquisition of speech sound categories. In S. Hirsh, D. E. Eldredge, I. J. Hirsh, & R. S. Silverman (Eds.), *Hearing and Davis* (pp. 265-280). St. Louis, MO: Washington University Press.
- KUHL, P. K., & HILLENBRAND, J. (1979, March). *Speech perception by young infants: Perceptual constancy for categories based on pitch contour*. Paper presented to the Society for Research on Child Development, San Francisco.
- KUHL, P. K., & MILLER, J. D. (1982). Discrimination of auditory target dimensions in the presence or absence of variation in a second dimension by infants. *Perception & Psychophysics*, **31**, 279-292.
- MILLER, C. L., & HANKES-RUZICKA, E. (1978). A parametric investigation of the cardiac no-delay discrimination paradigm and voice-onset-time discrimination in infants. *Infant Development Laboratory: Status Report No. 2* (pp. 275-314). Madison: University of Wisconsin.
- MORSE, P. A. (1972). The discrimination of speech and non-speech stimuli in early infancy. *Journal of Experimental Psychology*, **14**, 477-492.
- OLSHO, L. W. (1984). Infant frequency discrimination. *Infant Behavior & Development*, **7**, 27-35.
- OLSHO, L. W. (1985). Infant auditory perception: Tonal masking. *Infant Behavior & Development*, **8**, 371-384.
- OLSHO, L. W., SCHOON, C., SAKAI, R., TURPIN, R., & SPERDUTO, V. (1982). Auditory frequency discrimination in infancy. *Developmental Psychology*, **18**, 721-726.
- ROTH, P. L., & MORSE, P. A. (1975). An investigation of infant VOT discrimination using the cardiac OR. *Infant Development Laboratory: Status Report No. 1*. Madison: University of Wisconsin.
- SACHS, J. (1977). The adaptive significance of linguistic input to pre-linguistic infants. In C. Snow & C. Ferguson (Eds.), *Talking to children: Language input and language acquisition* (pp. 51-61). Cambridge, England: Cambridge University Press.
- SACHS, J., BROWN, R., & SALERNO, R. A. (1976). Adults' speech to children. In W. von Raffler-Engel & Y. Lebrun (Eds.), *Baby talk and infant speech* (pp. 240-245). Lisse, The Netherlands: Swets and Zeitlinger.
- SNOW, C. (1977). The development of conversation between mothers and babies. *Journal of Child Language*, **4**, 1-22.
- SPENNER, B. F., ENGBRETSON, A. M., MILLER, J. D., & COX, J. R. (1974). Random-access programmable recorder of complex sounds (RAP): A digital instrument for auditory research. *Journal of the Acoustical Society of America*, **55**(Suppl), 427A.
- SPRING, D. R., & DALE, P. S. (1977). Discrimination of linguistic stress in early infancy. *Journal of Speech & Hearing Research*, **20**, 224-232.
- STREETER, L. A. (1976). Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature*, **259**, 39-41.
- SULLIVAN, J. W., & HOROWITZ, F. G. (1983). The effects of intonation on infant attention: The role of the rising intonation contour. *Journal of Child Language*, **10**, 521-534.
- TREHUB, S. E., BULL, D., & THORPE, L. (1984). Infants' perception

- of melodies: The role of melodic contour. *Child Development*, *55*, 821-830.
- TREHUB, S. E., & RABINOVITCH, S. R. (1972). Auditory-linguistic sensitivity in early infancy. *Developmental Psychology*, *6*, 74-77.
- TREHUB, S. E., SCHNEIDER, B. A., & ENDMAN, M. (1980). Developmental changes in infants' sensitivity to octave-band noises. *Journal of Experimental Child Psychology*, *29*, 282-293.
- TREHUB, S. E., THORPE, L. A., & MORRONGIELLO, B. A. (1985). Infants' perception of melodies: Changes in a single tone. *Infant Behavior & Development*, *8*, 213-223.
- WORMITH, S. J., PANKHURST, D., & MOFFITT, A. R. (1975). Frequency discrimination by young infants. *Child Development*, *46*, 272-275.

#### NOTES

1. Sullivan and Horowitz (1983) used both synthetic and natural [ba] tokens. Values reported here are for the synthetic tokens and are based on reading the graphs from the published report.

2. If fewer than 10 high-amplitude sucks were recorded in each of 2 consecutive minutes or fewer than 8 for 1 min following the peak preshift rate, the infant was excluded for low preshift sucking rate. If the infant registered 3 or fewer high-amplitude sucks in the first postshift minute, the infant was excluded for low postshift sucking rate. Extremely low sucking rates near the shift point of the procedure can result in long interstimulus intervals, thereby introducing a significant memory factor into the discrimination task.

3. We also analyzed the data in terms of percent of maximum for each individual infant (Karzon, 1985; Kuhl & Miller, 1982; Spring & Dale, 1977), with the same result. Using the percent scores, the overall  $F(4,59) = 5.78$  ( $p < .0005$ ), and there was a main effect for group assignment [ $F(1,59) = 4.38$ ,  $p < .0408$ ]. The effect of stimulus type and the interaction were not significant.

(Manuscript received November 2, 1987;  
revision accepted for publication June 17, 1988.)

### Second International Facet Theory Conference University of Surrey, Guildford, England July 4-6, 1989

#### CALL FOR PAPERS

The Second International Facet Theory Conference is to be held at the University of Surrey, England, on July 4-6, 1989. Papers on any application of facet theory are welcome. While the diverse application of facet theory, or developments in the facet approach, will be reflected in the content of the conference, it is expected that the major substantive focus of the conference will reflect the principal areas of activity in psychology and the social sciences. In addition, there will be sessions exploring methodological and statistical developments related to facet theory. Selected papers will be published following the conference.

Format of submissions: An abstract and summary should be submitted by March 31, 1989, to David Canter or Ian Donald, from whom further details are also available. The summary should be a maximum of four pages, and contain sufficient detail to allow an assessment of its appropriateness to be made. All papers will be refereed. A full version of the proposed paper may be submitted along with the summary and abstract.

Address for submissions and further details: Department of Psychology, University of Surrey, Guildford, Surrey GU2 5XH, U.K.