J. Child Lang. **19** (1992), 1–17. Printed in Great Britain

Canonical and non-canonical syllable discrimination by two-day-old infants*

CHRISTINE MOON
Pacific Lutheran University

THOMAS G. BEVER
University of Rochester

AND

WILLIAM P. FIFER
Columbia University and *New York State Psychiatric Institute*

(*Received* 22 *August* 1989. *Revised* 9 *May* 1991)

ABSTRACT

Canonical syllables may be important units in early speech perception as well as production. Twenty infants (mean age 51 hours) (and twenty controls) were tested for their ability to discriminate between members of syllable pairs which were either canonical (pæt and tæp) or non-canonical (*pst* and *tsp*). A discrimination learning method was used in which syllables signalled the availability of either a recording of the mother's voice or silence – one of which was presented if the infant began a sucking burst. Infants in the canonical condition changed sucking patterns during signals over an 18-minute experimental session and activated their mother's voice more than silence, consistent with previous experiments using mother's voice as a reinforcer. In the non-canonical condition, infants also changed sucking patterns but sucked more during the signal for quiet than mother's voice, contrary to previous findings. Differential sucking during the syllables indicated discrimination in both conditions, but infants responded differently depending upon whether the syllables were canonical or non-canonical. The activation of silence in the non-canonical condition may be the result of a preference for quiet, but it is better explained as a failure to

[*] This research was supported by NICHD grant RO1 HD22817 to the first author and RO1 HD20102 to the third. Stimuli were prepared at Haskins Laboratory with support of NICHD contract NO1 HD52910. We would like to thank the Departments of Neonatology and newborn nursery personnel at North Central Bronx Hospital and Babies Hospital, Columbia-Presbyterian Medical Center, for their co-operation and support.
Address for correspondence: Christine Moon, Department of Psychology, Pacific Lutheran University, Tacoma, WA 98447, USA.

progress to a level of differential responding that was reached by the canonical group.

INTRODUCTION

Speech and speechlike sounds appear to be more salient to newborn infants than other sounds. Newborns alter sucking patterns to activate recordings of singing voices but not instrumental music (Cairns & Butterfield, 1975). Vowels elicit larger cardiac responses in two-day-olds than either sine or square waves (Clarkson & Berg, 1983). In fact, sine waves have been reported to elicit no observable behavioural (Turkewitz, Birch & Cooper, 1971) or cardiac response in two-day-olds (Graham & Jackson, 1970; Turkewitz, Birch & Cooper, 1971).

Speech and nonspeech are lateralized in the neonatal brain, as evidenced by both electrophysical and behavioural data. Auditory-evoked potentials of one-day-old infants demonstrate a unique left-hemisphere response for second formant transitions (Molfese & Molfese, 1979). Hemisphere differences to speech and non-speech stimuli have been recorded in premature infants two weeks after birth (Molfese & Molfese, 1980). Preterms have also provided behavioural evidence for lateralization by a reduction in limb tremors on the right side during exposure to speech (Segalowitz & Chapman, 1980). Two-day-olds respond behaviourally to non-speech stimuli presented in the left ear but not the right (Prescott, 1985; Davis & DeCasper, 1989).

Which characteristics of the speech signal segregate it from other auditory events and make it salient to the newborn? Electrophysical measures of newborn response to speech and non-speech stimuli have left the question open (Segalowitz, 1983). One likely characteristic is syllabicity. The canonical syllable, which is an important unit of organization in the development of speech production (Oller, 1980), is likely to be important in early perception as well. Evidence is emerging from deaf infants that auditory perception plays a role in early canonical syllable production (Oller & Eilers, 1988; Lynch, Oller & Steffens, 1989). Canonical syllables are between 100 and 500 msec in duration; have nuclei (vowels) and margins (consonants) which differ from each other in intensity, resonance and vocal tract openness; and have specified margin transition duration (Oller, 1986).

There are some indications in the literature that newborns may be sensitive to the parameters that define canonical syllables. Very young infants have been found to produce heart-rate decelerations – theoretically a sign of attention (Graham & Clifton, 1966) – to acoustic stimuli which are pulsed rather than continuous (Clarkson & Berg, 1983). Temporal transitions, such as those present at syllable margins, may play an important role in infant auditory perception (Berg & Berg, 1979). Recent experiments with four-day- and two-month-olds suggest that very early perception of the speech signal may be organized around the syllable nucleus, with perceptual representation of margins becoming more elaborated with maturation and/or experience (Bertoncini, Bijeljac-Babic, Blumstein & Mehler, 1987; Bertoncini, Bijeljac-Babic, Jusczyk, Kennedy & Mehler, 1988).

To date, only one study of a direct comparison of newborn perception of canonical versus non-canonical syllables has been published. Using the dishabituation of sucking response, Bertoncini & Mehler (1981) investigated one-month-olds' ability to discriminate a change in syllable margin in the context of canonical and non-canonical syllables. They concluded that discrimination was better when the syllable was canonical, but methodological problems cloud their interpretation, especially since infants discriminated the change in non-canonical as well as canonical syllables (Aslin, Pisoni & Jusczyk, 1983).

A new method for testing auditory discrimination as early as 24 hours after birth shows promise for investigating early speech perception (Moon & Fifer, 1990*a*). It is an adaptation of a method in which the presence or absence of a tone signals availability of a recording of mother's voice if the infant begins to suck on a non-nutritive nipple (DeCasper & Fifer, 1980; Fifer, 1981; Spence & DeCasper, 1987). In the present design, one of two syllables is the signal for availability of mother's voice. Differential responding to the signal for the voice is an indication of syllable discrimination. The operant discrimination sucking procedure compares favourably in two important ways with the high-amplitude sucking dishabituation procedure which is most commonly used with older infants. More subtle differences in responding to the two syllables across the session may emerge, thus providing a more qualitative evaluation than is possible with the single data-point provided by dishabituation. Moreover, the presence of a potent reinforcer (mother's voice) is likely to motivate newborns to respond, thus providing a more sensitive measure than is possible by presenting syllables alone (Cairns & Butterfield, 1975). The following experiment had two aims: (1) to test the usefulness of the neonatal operant discrimination sucking procedure with syllables, and (2) to compare the responses of infants with very little postnatal auditory experience to canonical versus non-canonical syllables.

METHOD

Subjects

The subjects were four groups of 10 infants between the ages of 27 and 81 hours since birth who met the following criteria: 2500–4000 grammes birthweight, Apgar scores of no less than 8 at 1 and 5 minutes, normal spontaneous vaginal delivery with no pre- or perinatal complications. Infants in the *pat/tap* group had a mean age of 51 hours and 5 were girls.

Twenty-seven infants were tested in order to obtain data from 10. Seventeen were disqualified for the following reasons: (i) cessation of sucking for two 45-second periods during the session due to crying (7) and drowsiness (7); (ii) equipment problem (2); (iii) experimenter error (1). In the *pst/tsp* group the mean age was 50 hours and 4 of the subjects were girls. Thirty-seven were tested to obtain data from 10. Disqualifications were for the following reasons: (i) cessation of sucking due to crying (11) and drowsiness (9); (ii) equipment problem (4); (iii) experimenter error (3).

In addition to the subjects in the two experimental groups, two additional groups of 10 subjects were added in a control condition. The 10 infants in the single syllable *pat* group averaged 36 hours of age and included 4 girls. Twenty were tested in order to obtain data from 10 infants. Disqualifications were for the following reasons: (i) crying (4) and drowsiness (4); (ii) equipment problem (2). The 10 infants in the single-syllable *pst* group had a mean age of 44 hours and 6 were girls. Nineteen were tested. Disqualifications were for the following reasons: crying (6) and drowsiness (3).

The two experimental groups were run concurrently. The two control groups were concurrent with each other but were added to the study after completion of the experimental groups, subsequent to relocation of the laboratory in a new hospital. Recruitment was from a nursery ward where infants were eligible for discharge at 40 hours rather than 72 hours of age, as was the case for the experimental groups. As a result, the *pat* control group (mean 36 hours) was significantly younger than the *pat/tap* group (mean 51 hours) (t (d.f. 18) $= 3.24$, $p < 0.01$). The *pst* control group (mean 44 hours) did not differ significantly in age from the *pst/tsp* group (mean 50 hours) (t (d.f. 18) $= 1.01$, $p < 0.10$). In a previous, similar study, there was no difference in response to tone/no tone signals between one- and three-day-old subjects (Fifer, 1981). However, in other studies we have found significant age differences in proportion of responses to syllable signals. Younger infants (between one and two days of age) show GREATER differential responding to syllable signals than older infants (Moon & Fifer, 1990*b*). The younger *pat* group therefore should not have been at a disadvantage relative to the older *pat/tap* group in differential responding if the signal strings were discriminable.

Also related to the addition of the two control groups subsequent to the experimental groups is the lower attrition rate for the later groups. This is most likely due to the increased experience of the experimenters in recognizing an appropriate infant behavioural state to begin a session, thus increasing the likelihood that infants would remain alert for the entire 18 minutes of the session. Over time in our laboratory we have seen a steady decrease in attrition rates in similar experiments with newborns.

Apparatus

All infants heard the speech stimuli through Telephonics TDH-39 headphones used with standard rubber ear-pieces. Infants sucked on a standard hospital feeding nipple which was connected via surgical tubing to a pressure-transducer attached to a pressure-monitor. The measurements in millimetres of mercury (mmHg) were sent to both a chart-drive event-recorder for visual on-line inspection of data and to an analogue-to-digital converter. The digital reading served as input to a microcomputer which recorded the subjects' sucking patterns for future analysis and also controlled stimulus presentation.

Stimuli

Syllable signals. For the experimental groups there were two pairs of syllables: (1) /pæt/ and /tæp/, (2) /pst/ and /tsp/. For the control groups only one syllable was presented. For one of the groups it was /pæt/ and for the other group it was /pst/. For the experimental groups, the right channel of the tape-recorder contained strings of one syllable, and the left channel contained the other. The channels were routed through two separate computer-driven relays which controlled syllable presentation by blocking the transmission of the continuously running tape at appropriate times. For the control groups, the output of only one channel of the tape-recorder was utilized (e.g. the left channel with /pæt/ recorded on it). The signal from this one channel was routed through the two relays which resulted in identical output from both relays into the headphones.

All syllables were 350 msec in duration with 450 msec of interstimulus silence. They were originally natural tokens spoken in a monotone by a male voice and edited for duration at a sampling rate of 20 kHz. Natural tokens were used in order to preserve all possible coarticulation information, particularly in the case of the unusual non-canonical syllables. Necessary adjustments in duration were made by extracting from or duplicating the stable centre portion of the syllable nuclei. The plosive margins and their transitions were unedited. All codas had audible release bursts. The syllables were equated for loudness by four adult listeners using the method of limits. The syllables /pæt/ and /tæp/ were presented at 71 dB SPL (General Radio Sound Level Meter, Type 1551 C, Scale A). The syllables /pst/ and /tsp/ were presented at 63 dB (SPL, A).[1]

[1] It was not possible to present the canonical and non-canonical stimuli at equal dB levels. When the canonical syllables were presented at appropriate levels of intensity for adults (approximately 70 dB, Scale A) the non-canonical syllables at the same level were

Mother's voice

The sound of the voice of each infant's mother was recorded in an interview in her hospital room on the day prior to the experimental session. From about 5 minutes of conversational speech, approximately 30 seconds of uninterrupted mother's voice were spliced together. This tape-segment was repeated and re-recorded until there were 25 continuous minutes of the mother's voice with no pauses longer than 1·5 seconds. The infant heard this recording through the headphones at conversational sound level.

Procedure

In order to maximize the probability of obtaining an alert and calm infant, sessions were conducted about an hour prior to an anticipated feed. Newborns were moved in their own bassinets to a quiet, dimly lighted room adjoining the nursery. The lights were kept low to maximize infant eye-opening. Infants were brought to a quiet, alert state with open eyes and adequate muscle tone by talking to them, gently massaging their limbs, or holding them upright. The head end of the bassinet was elevated by about 20 degrees to help maintain alertness. Infants who were not judged to be calm or alert were returned to the nursery for later testing if possible. Infants who did not reach the point of stimulus presentation were not counted in calculation of the attrition rate. Subjects began the session swaddled in their blanket unless they appeared likely to become drowsy, in which case the blanket was loosely wrapped around them. For the infants who began a session, the earphones were placed securely over both ears, and the nipple was placed in the baby's mouth. A two-minute period of adjustment followed for the infant to begin sucking in discrete bursts. Newborns typically suck in a train of individual sucks which lasts for a few seconds alternating with an approximately equal period of non-sucking. After observing that sucking was occurring in discrete bursts, stimulus presentation began. While one experimenter monitored the recording equipment, the other experimenter, blind to the experimental condition and unable to hear the stimuli, monitored the baby and held the nipple in the baby's mouth. In the first 12 minutes of the session if the infant's eyes closed and muscle tone became flaccid, the

uncomfortably loud to adults and could not ethically be presented through headphones to newborns. Behavioural measures of threshold and equiloudness contours for complex auditory stimuli do not exist for newborn infants. Data from Werner & Gillenwater (1990) indicate that for 4000 Hz pure tones, 63 dB is at or slightly above threshold. However, it is likely that infant sensitivity would be greater in a procedure using reinforced responding and speech stimuli (Werner & Gillenwater, 1990). Response bias appears to play a significant role in infant response to auditory events (Olsho, Koch & Carter, 1988; Werner & Gillenwater, 1990).

experimenter monitoring the infant loosened the baby's blanket to try to restore alertness. If the baby began to become active and irritable, the experimenter attempted to calm the infant by restraining the hands over the chest. These manoeuvres were not permitted during the final 6 minutes of the session. Sessions lasted 18 minutes unless terminated prematurely. Sessions were terminated prematurely if the baby failed to produce measurable sucks due primarily to fussiness, drowsiness, or failure of sucking to be recorded as discrete bursts. Session termination occurred the second time that a 45-second period of no sucking or unmeasurable bursts occurred.

The sounds were presented contingent upon sucking pressure. A sucking burst was defined as -20 mmHg for a minimum of 2 seconds. Syllable signals were presented only when the computer detected the absence of a sucking burst for at least 2 seconds. The syllables were in strings such that 4 seconds of one syllable alternated with 4 seconds of the other syllable. If the infant initiated a sucking burst during either string, the syllable string was terminated (unless the string had been activated for less than a second, in which case it ceased after 1 second), and either the mother's voice or quiet was presented for as long as the sucking burst continued. For example, in the canonical syllable experimental group, if *pat* was the signal for the mother's voice, and the infant sucked while hearing the *pat* string, that string ceased and the baby then heard the mother's voice for the duration of the sucking burst (usually several seconds). The signal for quiet in this example would be *tap*, and if the baby sucked during that signal string, silence would be presented for the duration of the sucking burst. At the end of a burst, the 4-second syllable strings resumed alternation. For five of the infants, *pat* signalled the mother's voice and *tap* signalled silence, while for the remaining infants the signal/voice pairing was reversed. Each of the two strings had a random probability of beginning the sequence, with the prohibition of more than three occurrences of the same string following successive bursts.

As described above, two groups of infants were presented with a pair of syllables (either canonical or non-canonical). In the other two groups, infants were presented with only one syllable (either canonical or non-canonical). For example, subjects in one of these latter two groups heard only strings of a single canonical syllable, *pat*, during the periods between sucking bursts. Initiation of a burst resulted in a reinforcer (mother's voice or quiet). There was a 0·50 probability of hearing the mother's voice, the same for quiet. The second control group heard only the non-canonical syllable *pst* during interburst intervals.

RESULTS

The measure of differential responding to the two-syllable signals was the relative frequency with which the infant initiated sucking bursts during the signal for the mother's voice versus the signal for quiet. Results from other experiments indicate that when the signal for the mother's voice is differentiable from the signal for an alternative, infants change sucking patterns over the course of the session: there is no difference in the number of bursts during each signal at the beginning of the session, and there is a greater frequency during the signal for the mother's voice by the end. Previous analyses have been based upon a division of the experimental session into thirds (DeCasper & Fifer, 1980; Fifer, 1981) or 6-minute periods of an 18-minute session (Moon & Fifer, 1990*a*). It was anticipated that where differences in responding did occur, the differences would be in favour of the mother's voice rather than quiet due to previous evidence of infant responsiveness to female voices (DeCasper & Carstens, 1981) and to the maternal voice in particular (Mills & Melhuish, 1974; Mehler, Bertoncini, Barrière & Jassik-Gerschenfeld, 1978; DeCasper & Fifer, 1980; Fifer 1981; Moon & Fifer, 1990*a*).

It was assumed, as in DeCasper & Fifer (1980), that during the earlier minutes of the session infants would not yet have learned the sucking/voice contingency and would not be as likely to show consistent responding to the signal for voice as they would during the final third of the session. An increase was expected in the proportion of times a response was made to the signal for voice versus the signal for quiet across the session, given signal discriminability.

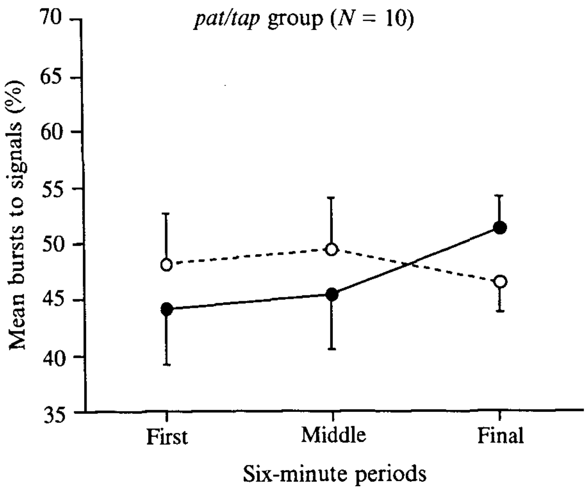
Analyses were conducted on experimental and control groups separately. In experiments using the habituation of high-amplitude sucking (HAS), comparisons are made between experimental and control groups. In the HAS procedure, control groups are necessary due to spontaneous sucking-rate changes. These rate changes could be erroneously interpreted as habituation or dishabituation to the auditory stimuli (Jusczyk, 1985). In the present study using a new operant two-signal discrimination procedure, the control groups were added for a different purpose. They were added to rule out the possibility that infants in this procedure discriminate unintended signals instead of the syllables. Examples of such signals include experimenter cues, clicking of relays, or changes in ambient light from the computer screen, etc. Finding that there is no difference in response during the two signals, which are intended to be indiscriminable, helps to validate the new method.

Response to signals was indicated by a ratio of the number of times the subject initiated a sucking burst during a signal string divided by the number of signal strings presented during the period. (In the following, 'signals' will be used interchangeably with 'signal strings'.) For example, during the final 6 minutes an infant who initiated sucking bursts during 14 of 20 signals for voice and 11 of 21 signals for quiet would have rates of 70% and 55%, respectively. (Note that the sucking burst/signal rates need not sum to 100%.) The sucking burst/signal proportions for voice and quiet were then combined to form a PREFERENCE RATIO for each infant by dividing the

TABLE I. *Individual subjects' proportionate responding to two-syllable signals: experimental groups*

				Six-minute periods								
				First			Middle			Final		
S.	Sex	Age (hr)	Signal	M's voice	Quiet	PR[a]	M's voice	Quiet	PR	M's voice	Quiet	PR
				A. Canonical syllable pair: /pæt/ and /tæp/								
1	F	31	pæt	0·35	0·56	0·39	0·59	0·68	0·46	0·63	0·55	0·53
2	F	58	pæt	0·46	0·54	0·46	0·29	0·32	0·48	0·55	0·48	0·53
3	M	69	pæt	0·52	0·58	0·47	0·74	0·59	0·56	0·58	0·50	0·54
4	M	51	pæt	0·34	0·27	0·56	0·41	0·42	0·50	0·42	0·40	0·51
5	F	44	pæt	0·32	0·33	0·49	0·33	0·43	0·44	0·41	0·37	0·53
6	F	67	tæp	0·59	0·44	0·58	0·42	0·55	0·44	0·48	0·54	0·47
7	M	54	tæp	0·30	0·52	0·37	0·28	0·54	0·34	0·50	0·46	0·52
8	M	54	tæp	0·64	0·69	0·48	0·64	0·72	0·47	0·65	0·58	0·53
9	F	39	tæp	0·22	0·30	0·43	0·42	0·31	0·58	0·46	0·44	0·51
10	M	43	tæp	0·65	0·59	0·53	0·42	0·38	0·53	0·46	0·34	0·58
Mean		51·1		0·44	0·48	0·47	0·45	0·49	0·48	0·51	0·47	0·53
S.D.		12.0		0·15	0·14	0·07	0·15	0·15	0·07	0·08	0·08	0·03
				B. Non-canonical syllable pair: /pst/ and /tsp/								
1	F	29	*pst*	0·82	0·89	0·48	0·71	0·79	0·47	0·58	0·57	0·51
2	F	60	*pst*	0·79	0·73	0·52	0·77	0·71	0·52	0·32	0·41	0·44
3	M	32	*pst*	0·36	0·44	0·45	0·30	0·39	0·43	0·46	0·59	0·44
4	M	71	*pst*	0·68	0·66	0·51	0·50	0·70	0·42	0·52	0·74	0·41
5	M	38	*pst*	0·51	0·35	0·59	0·48	0·57	0·46	0·52	0·71	0·42
6	M	48	*tsp*	0·42	0·64	0·39	0·55	0·87	0·39	0·43	0·37	0·54
7	F	38	*tsp*	0·54	0·72	0·43	0·24	0·31	0·44	0·38	0·52	0·42
8	M	45	*tsp*	0·37	0·44	0·46	0·39	0·59	0·40	0·47	0·31	0·60
9	F	81	*tsp*	0·44	0·27	0·62	0·75	0·92	0·45	0·71	0·73	0·49
10	M	59	*tsp*	0·36	0·29	0·55	0·28	0·31	0·48	0·33	0·63	0·35
Mean		50·2		0·53	0·54	0·50	0·50	0·62	0·45	0·47	0·56	0·46
S.D.		17·2		0·18	0·21	0·07	0·20	0·22	0·04	0·12	0·16	0·07

[a] Preference ratio is the proportion of responses to mother's voice divided by the sum of proportions to mother's voice plus quiet.



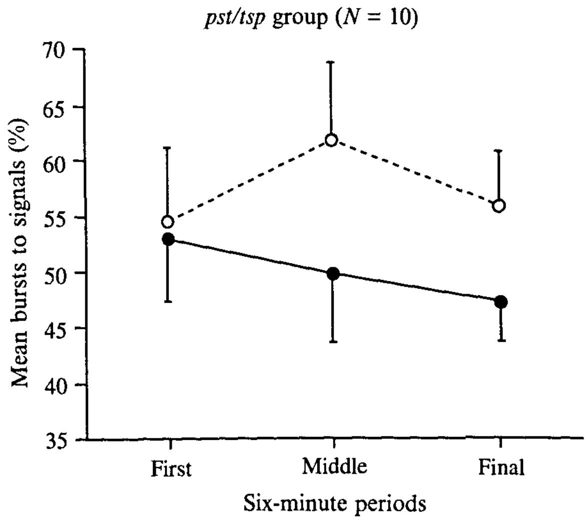


Fig. 1. Mean responding to signals for mother's voice (●——●) and for quiet (○---○) for 6-minute periods for two experimental groups. (Percentage responses are the number of sucking bursts divided by the number of signal presentations during a period.)

percentage for the mother's voice by the sum of the percentage for the mother plus the percentage for quiet (Spence & DeCasper, 1987). A preference ratio greater than 0·50 indicates proportionately more frequent responding during the signal for the mother's voice, and less than 0·50 indicates more frequent responding during the signal for quiet.

Comparisons were conducted for the three 6-minute periods. In the *pat/tap* condition, in the final 6-minute period, nine of ten subjects had preference ratios greater than 0·50 ($p = 0{\cdot}022$, binomial test, two-tailed): see Table 1. The mean preference ratio (0·53) was significantly greater than 0·50, t (d.f. 9) = 3·01, $p = 0{\cdot}015$, two-tailed test. During the first 6-minute period, four of the ten infants sucked relatively more frequently during the signal for the voice (mean preference ratio = 0·47), and the difference from 0·50 was not significant, t (d.f. 9) = 1·19, $p = 0{\cdot}26$. Eight of ten subjects increased preference ratios from first to final 6-minute periods, a difference which approached significance, t (d.f. 9) = 2·04, $p = 0{\cdot}07$ (two-tailed). A test was conducted for the *pat/tap* group in the middle period (mean preference ratio = 0·48). There was no significant difference from 0·50 preference ratio for this group, t (d.f. 9) = 1·10, $p = 0{\cdot}32$: see Figure 1.

For the *pst/tsp* group during the final 6-minute period, three of ten infants had preference ratios greater than 0·50. The mean (0·46) was not significantly less than 0·50, t (d.f. 9) = 1·65, $p = 0{\cdot}13$ (two-tailed test). Mean response during the first 6-minute period was 0·50. The decline in mean preference ratio from first to final period was not significant, t (d.f. 9) = 0·99, $p = 0{\cdot}35$. A test of the middle third of the session revealed that nine of ten infants sucked relatively more frequently during the signal for quiet than for the mother's voice, a mean preference ratio difference that was significant, t = 4·40, $p = 0{\cdot}002$, two-tailed test.

For the two control groups, there were no significant differences in response to the signal for the mother's voice compared to the signal for quiet, for any comparisons. For both the *pat* and the *pst* groups, mean preference ratios did not differ in any of the three 6-minute periods from 0·50. Moreover, mean preference ratios of the two groups did not differ from each other for any of the periods: see Table 2.

Response to signals was the dependent variable of primary interest in the study. However, there was a second experimental contingency to which infants could respond. Duration of sucking burst and the presentation of voice or quiet were synchronous, i.e. the longer the burst, the longer the presentation of voice (or quiet). Infants have shown the ability to respond on this measure with longer sucking bursts during their mother's voice as compared to the voice of a female stranger (Fifer, 1981) and as compared to quiet (Moon & Fifer, 1990*a*). However, there have also been results of no differences in burst duration during intra-uterine sounds versus quiet (DeCasper & Sigafoos, 1983) and during mother's voice versus quiet (Moon & Fifer, 1990). In the current study there were no differences in mean burst duration during the mother's voice versus quiet for any of the four groups of infants during a 6-minute period.

TABLE 2. *Individual subjects' proportionate responding to two-syllable signals: control groups*

				Six-minute periods								
				First			Middle			Final		
S.	Sex	Age (hr)	Signal	M's voice	Quiet	PR[a]	M's voice	Quiet	PR	M's voice	Quiet	PR
				A. Single canonical syllable: /pæt/								
1	F	27	pæt	0·55	0·77	0·41	0·65	0·54	0·55	0·46	0·32	0·59
2	M	48	pæt	0·44	0·46	0·49	0·58	0·69	0·46	0·54	0·61	0·47
3	F	38	pæt	0·72	0·71	0·50	0·52	0·50	0·51	0·64	0·46	0·58
4	M	31	pæt	0·54	0·37	0·60	0·44	0·28	0·61	0·75	0·62	0·55
5	M	37	pæt	0·38	0·45	0·46	0·42	0·57	0·43	0·42	0·51	0·45
6	M	29	pæt	0·55	0·38	0·60	0·52	0·42	0·55	0·19	0·23	0·45
7	M	28	pæt	0·56	0·34	0·62	0·64	0·42	0·60	0·63	0·52	0·55
8	F	48	pæt	0·28	0·49	0·37	0·32	0·50	0·39	0·39	0·36	0·52
9	F	33	pæt	0·38	0·31	0·55	0·48	0·61	0·44	0·18	0·39	0·32
10	M	44	pæt	0·60	0·36	0·63	0·65	0·55	0·55	0·70	0·63	0·53
Mean		36·3		0·50	0·46	0·52	0·52	0·51	0·51	0·49	0·47	0·50
S.D.		8.1		0·13	0·16	0·09	0·11	0·11	0·08	0·20	0·14	0·08
				B. Single non-canonical syllable: /pst/								
1	M	37	*pst*	0·29	0·33	0·47	0·44	0·52	0·46	0·29	0·57	0·34
2	F	48	*pst*	0·42	0·15	0·74	0·45	0·59	0·43	0·56	0·55	0·51
3	F	33	*pst*	0·63	0·50	0·56	0·46	0·60	0·43	0·48	0·63	0·43
4	M	52	*pst*	0·33	0·29	0·54	0·69	0·79	0·47	0·61	0·62	0·50
5	F	30	*pst*	0·88	0·88	0·50	0·79	0·64	0·55	0·87	0·83	0·51
6	F	41	*pst*	0·38	0·52	0·42	0·59	0·41	0·59	0·43	0·44	0·49
7	F	53	*pst*	0·47	0·28	0·63	0·45	0·39	0·53	0·21	0·33	0·39
8	M	49	*pst*	0·69	0·67	0·51	0·70	0·58	0·55	0·75	0·76	0·50
9	F	60	*pst*	0·39	0·29	0·58	0·50	0·40	0·56	0·32	0·41	0·44
10	M	35	*pst*	0·63	0·52	0·55	0·63	0·77	0·45	0·71	0·42	0·63
Mean		43·8		0·51	0·44	0·55	0·57	0·57	0·50	0·52	0·56	0·47
S.D.		10·0		0·19	0·22	0·09	0·13	0·14	0·06	0·22	0·16	0·08

[a] Preference ratio is the proportion of responses to mother's voice divided by the sum of proportions to mother's voice plus quiet.

DISCUSSION

Infants who heard canonical syllables *pat* and *tap* as signals for their mother's voice and for quiet activated a recording of their mother's voice relatively more frequently than quiet during the final 6 minutes of the session. There was no difference in response to the two signals in the beginning of the session, and the change over time suggests that the infants learned the experimental contingency. These predicted results replicate experiments showing that newborns discriminate canonical syllables (Bertoncini & Mehler, 1979; Bertoncini *et al.* 1987) and that they are capable of auditory-discrimination learning when their mother's voice is a reinforcer (DeCasper & Fifer, 1980; Moon & Fifer, 1990*a*).

Subjects in the non-canonical syllable condition also evinced discrimination of the syllables by sucking differentially during the two signals *pst* and *tsp*, but they unexpectedly sucked more frequently during the signal for quiet than for the mother's voice. Thus they demonstrated discrimination of the syllables but in a manner not consistent with the canonical group nor with previous studies in which young infants have responded preferentially to the sound of the mother's voice. No published study to date has reported preferential responding for an alternative to the mother's voice.

Unlike the groups of infants who heard two different syllables as signals preceding voice and quiet, the infants who heard a single syllable showed no difference in sucking patterns across the session during syllables which preceded their mother's voice versus those which preceded quiet. The undifferentiated response of the single-syllable subjects demonstrates that artifacts of the procedure such as clicking of relays or experimenter cues were not being used by the infants as predictors of the availability of voice, and therefore that differential responding in the two-syllable groups was based upon unique characteristics of the two syllables.

Burst duration was not a primary variable in this study, but it is of some interest. There were no significant differences in burst durations during presentations of the mother's voice and quiet in any group. There is evidence that, under similar conditions, newborns can respond differentially on the burst duration contingency (Fifer, 1981; Moon & Fifer, 1990*a*). One might expect that auditory stimulation during sucking, with a salient stimulus such as the mother's voice, might affect burst duration through some form of orientating arousal. The current study and others (DeCasper & Sigafoos, 1983; Moon & Fifer, 1990*a*) suggest that burst-duration changes are not obligatory newborn responses to contingent auditory stimulation. The results contribute to the plausibility that in experiments in which burst durations do change, it is as a result of learning.

The main prediction of the current study was confirmed: that infants would respond differentially in the canonical syllable condition than in the non-canonical condition. The results are compatible with the observation (Aslin *et al.* 1983) that in Bertoncini & Mehler (1979), infants showed dishabituation and therefore discrimination both in the canonical and non-canonical conditions.

In the current experiment the unexpected result of greater responding to the signal for quiet than for the mother's voice by the non-canonical group requires a *post hoc* explanation. Two alternatives will be offered: (1) in the non-canonical condition, the infants preferred quiet to the mother's voice and rapidly learned to activate quiet more than voice; and (2) in the non-canonical condition, greater responding during the signal for quiet was due to a failure to progress in acquiring the response which was acquired by infants in the canonical condition.

The canonical and non-canonical conditions may have differed in the levels of arousing stimulation presented to the infants. Such sounds as *pst* and *tsp* are used by adults in many cultures as attention-getting signals. Repetition of these sounds through headphones may have resulted in a change of infant behavioural state. In both the two-syllable and single-syllable non-canonical groups, infants tended to respond to signals at higher rates than infants in the canonical groups – evidence perhaps for greater arousal in the non-canonical conditions. One way in which infants respond to over-stimulation or discrepant information is by shutting out the stimulation, typically by averting the head or eyes (Stern, 1977; Murray & Trevarthen, 1985). Perhaps over-stimulation in the non-canonical condition created an exceptional preference for quiet over the sound of the mother's voice, and so the infants rapidly learned to activate more quiet than voice. There is some evidence for this in the comparison of responses to voice and silence across the session. There was a relatively stable level of response to the signal for voice over the three periods in contrast to an increase in response to the signal for quiet in the middle period which persisted into the final period. Evidence against the over-stimulation hypothesis is that attrition rates did not differ for the canonical and non-canonical groups which would have been expected if infants were overstimulated in the non-canonical groups. It is also implausible, based upon evidence from similar experiments, that the subjects could have learned the sound/sucking contingency in as few as 7–11 minutes.

The second hypothesis takes levels of stimulation into account but offers an explanation which does not require unprecedented, rapid acquisition of the sucking response in the non-canonical condition. Learning to activate the voice recording includes learning to time the onset of sucking bursts during the appropriate 4-second syllable presentation period. There may be a phase during normal acquisition of the response in which burst onset is mistimed. If the infant, for example, anticipates the onset of the signal period for the mother's voice toward the end of the 4-second period of the signal for quiet, and begins to suck prematurely, more presentations of quiet would occur.

Such timed anticipatory responding occurs in pigeons (Gibbon & Balsam, 1981). Or, if the infant orients, with consequent somatic quieting (Graham & Clifton, 1966), to the signal for the mother's voice during its 4-second presentation, the onset of sucking may be sufficiently delayed for the response to occur during the subsequent signal for quiet. In support of this hypothesis was the (non-significant) tendency in the acquisition period for infants in the canonical as well as the non-canonical group to respond more to the signal for quiet. The canonical group then, if the response timing hypothesis is correct, learned to time sucking-burst onset appropriately by the final third of the session, whereas the non-canonical group did not. The non-canonical group's failure to acquire adequate timing could be due to their arousal state.

Admittedly, these explanations for the unexpected response of the non-canonical group are speculative. They do provide, however, fertile ground for experimentation. In future studies using more fine-grained analyses than were available here, it should be possible to measure latencies for response to signals and thus test the hypothesis that infants distribute burst-onset differently across time and conditions. Moreover, autonomic measures such as heart rate or heart-rate variability may indicate differential responding to canonical and non-canonical syllables or response differences to signals prior to behavioural change. Latency to respond and autonomic indicators could provide clues to the course of acquisition of discrimination learning which could help explain the results of the canonical and non-canonical conditions.

In addition to the process by which infants learned to respond differentially to the signals in the two conditions, future research should investigate the characteristics of the canonical and non-canonical syllables which produced divergent results in newborns. Perhaps the effect is unique to the presence of sustained high-frequency, aperiodic noise, in which case canonical syllables with fricative margins (e.g. /sa/ versus /fa/) should replicate the effect found here. It may be, however, that the critical dimension is violation of canonical form, and that parameters of margin transition, nucleus resonance, and duration affect infant response.

This study makes two contributions to the investigation of the development of speech perception. One is methodological. The operant discrimination learning procedure holds promise for being a sensitive measure of discrimination at the youngest ages. Furthermore, the process of infant response acquisition may tell us more than simply whether the infant has discriminated or not.

The second contribution is the evidence that two-day-olds respond differently to canonical than they do to non-canonical syllables, and the difference is not simply a failure to discriminate the non-canonical pair. The non-canonical syllables, although produced by a vocal tract, may not be processed as speech. This would suggest that canonical syllabicity is a requirement for the processing of an auditory event as speech by very young infants.

REFERENCES

Aslin, R. N., Pisoni, D. B. & Jusczyk, P. W. (1983). Auditory development and speech perception in infancy. In M. M. Haith & J. J. Campos (eds), *Infancy and developmental psychobiology*. Vol. 2. Carmichael's handbook of child psychology (4th edition). New York: Wiley.

Berg, W. K. & Berg, K. M. (1979). Psychophysiological development in infancy: state, sensory function, and attention. In J. Osofsky (ed.) *Handbook of infant development*. New York: Wiley.

Bertoncini, J., Bijeljac-Babic, R. V., Blumstein, S. E. & Mehler, J. (1987). Discrimination in neonates of very short CVs. *Journal of the Acoustic Society of America* **82**, 31–7.

Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P. W., Kennedy, L. J., & Mehler, J. (1988). An investigation of young infant's perceptual representations of speech sounds. *Journal of Experimental Psychology* **117**, 21–33.

Bertoncini, J. & Mehler, J. (1979). Syllables as units in infant speech perception. *Infant Behavior and Development* **4**, 247–60.

Cairns, G. F. & Butterfield, E. C. (1975). Assessing infants' auditory functioning. In B. Z. Friedlander, G. M. Sterritt, & J. Kirk (eds) *The exceptional infant: assessment and intervention*. vol. 3. New York: Brunner/Mazel.

Clarkson, M. G. & Berg, W. K. (1983). Cardiac orienting and vowel discrimination in newborns: crucial stimulus parameters. *Child Development* **54**, 162–71.

Davis, F. A. & DeCasper, A. J. (1989). Intrauterine heartbeat sounds are reinforcing for newborns because of active right-lateralized processes. Paper presented at Society for Research in Child Development, Kansas City, MO.

DeCasper, A. J. & Carstens, A. A. (1981). Contingencies of stimulation: effects on learning and emotion in neonates. *Infant Behavior and Development* **4**, 19–35.

DeCasper, A. J. & Fifer, W. P. (1980). Of human bonding: newborns prefer their mother's voices. *Science* **208**, 174–6.

DeCasper, A. J. & Sigafoos, A. D. (1983). The intrauterine heartbeat: a potent reinforcer for newborns. *Infant Behavior and Development* **6**, 19–25.

Fifer, W. P. (1981). Early attachment: maternal voice preference in 1- and 3-day-old infants. Unpublished doctoral dissertation, University of North Carolina, Greensboro.

Gibbon, J. & Balsam, P. (1981). Spreading association in time. In C. M. Locurto, H. S. Terrance & J. Gibbon (eds), *Autoshaping and conditioning theory*. New York: Academic Press.

Graham, F. K. & Clifton, R. K. (1966). Heart-rate change as a component of the orienting response. *Psychological Bulletin* **65**, 305–20.

Graham, F. K. & Jackson, J. C. (1970). Arousal systems and infant heart rate responses. In H. W. Reese & L. P. Lipsitt (eds), *Advances in child development and behavior*. Vol. 5. New York: Academic Press.

Jusczyk, P. W. (1985). The high-amplitude sucking technique as a methodological tool in speech perception research. In G. Gottlieb & N. A. Krasnegor (eds), *Measurement of audition and vision in the first year of postnatal life: a methodological overview*. Norwood, NJ: Ablex.

Lynch, M. P., Oller, D. K. & Steffens, M. (1989). Development of speech-like vocalizations in a child with congenital absence of cochleas: the case of total deafness. *Applied Psycholinguistics* **10**, 315–33.

Mehler, J., Bertoncini, J., Barrière, M. & Jassik-Gerschenfeld, D. (1978). Infant recognition of mother's voice. *Perception* **7**, 491–7.

Mills, M. & Melhuish, E. (1974). Recognition of mother's voice in early infancy. *Nature* **252**, 123–4.

Molfese, D. L. & Molfese, V. J. (1979). Hemisphere and stimulus differences as reflected in the cortical responses of newborn infants to speech stimuli. *Developmental Psychology* **15**, 505–11.

—— &—— (1980). Cortical responses of preterm infants to phonetic and nonphonetic speech stimuli. *Developmental Psychology* **16**, 574–81.

Moon, C. & Fifer, W. P. (1990*a*). Syllables as signals for 2-day-old infants. *Infant Behavior and Development* **13**, 377–90.

—— &—— (1990*b*). Newborns prefer a prenatal version of mother's voice. Poster presented at biannual meeting of International Society of Infant Studies, Montreal.

Murray, L. & Trevarthen C. (1985). Emotional regulation of interactions between 2-month-olds and their mothers. In T. M. Field & N. A. Fox (eds), *Social perception in infants.* Norwood, NJ: Ablex.

Oller, R. K. (1980). The emergence of the sounds of speech in infancy. In Yeni-Komshian, G. H., Kavanagh, J. F. & Ferguson, C. A. (eds), *Child phonology*: Vol 1, *Production.* New York: Academic Press.

—— (1986). Metaphonology and infant vocalizations. In B. Lindblom & R. Zetterstrom (eds), *Precursors of early speech.* New York: Stockton Press.

Oller, R. K. & Eilers, R. E. (1988). The role of audition in infant babbling. *Child Development* **59**, 441–9.

Olsho, L. W., Koch, E. G. & Carter, E. A. (1988). Nonsensory factors in infant frequency discrimination. *Infant Behavior and Development* **11**, 205–22.

Prescott, P. A. (1985). Differential reinforcing value of speech and heartbeats: a measure of functional lateralization in the neonate. *Dissertation Abstracts International* **48**, 286-B.

Segalowitz, S. J. (1983). Cerebral asymmetries for speech in infancy. In S. J. Segalowitz (ed.), *Language functions and brain organization : perspectives in neurolinguistics, neuropsychology and psycholinguistics.* New York: Academic Press.

Segalowitz, S. J. & Chapman, J. S. (1980). Cerebral asymmetry for speech in neonates: a behavioral measure. *Brain and Language* **9**, 281–8.

Spence, M. J. & DeCasper, A. J. (1987). Prenatal experience with low-frequency maternal voice-sounds influence neonatal perception of maternal voice samples. *Infant Behavior and Development* **10**, 133–142.

Stern, D. (1977). *The first relationship*, Cambridge, MA: Harvard University Press.

Turkewitz, G., Birch, H. G. & Cooper, K. K. (1971). Responsiveness to simple and complex auditory stimuli in the human newborn. *Developmental Psychobiology* **5**, 7–19.

Werner, L. A. & Gillenwater, J. M. (1990). Pure-tone sensitivity of 2- to 5-week-old infants. *Infant Behavior and Development* **13**, 355–75.