Prenatal Maternal Speech Influences
Newborns’ Perception of Speech Sounds*

ANTHONY J. DECASPER AND MELANIE J. SPENCE
University of North Carolina at Greensboro

Pregnant women recited a particular speech passage aloud each day during their last 6 weeks of pregnancy. Their newborns were tested with an operant-choice procedure to determine whether the sounds of the recited passage were more reinforcing than the sounds of a novel passage. The previously recited passage was more reinforcing. The reinforcing value of the two passages did not differ for a matched group of control subjects. Thus, third-trimester fetuses experience their mothers’ speech sounds and that prenatal auditory experience can influence postnatal auditory preferences.

Human newborns do not act like passive and neutral listeners. They prefer their own mothers’ voices to those of other females, female voices to male voices, and intrauterine heartbeat sounds to male voices, but they do not prefer their fathers’ voices to those of other males (Brazelton, 1978; DeCasper & Fifer, 1980; DeCasper & Prescott, 1984; Fifer, 1980; Panneton & DeCasper, 1984; Wolff, 1963). Why should newborns prefer some sounds over others? One hypothesis is that their auditory preferences are influenced by prenatal experience with their mothers’ speech and heartbeats (DeCasper & Prescott, 1984). Several considerations suggest this hypothesis is plausible.

Third-trimester fetuses hear, or are behaviorally responsive to, sound (e.g., Bernard & Sontag, 1947; Birnholz & Benacerraf, 1983; Grimwade, Walker, Bartlett, Gordon, & Wood, 1971; Johansson, Wedenberg, & Westin, 1964; Sontag & Wallace, 1935). Intrauterine recordings taken near term indicate that maternal speech and heartbeats are audible in utero (Querleu & Renard, 1981; Querleu, Renard, & Crepin, 1981; Walker, Grimwade, & Wood, 1971). Nonmaternal speech, for example male speech, is less audible because of attenuation by maternal tissue and/or masking by intrauterine sounds (Querleu & Renard, 1981; Querleu et al., 1981).

* This research was supported by a Research Council Grant from the University of North Carolina at Greensboro and a generous equipment loan by Professor Michael D. Zeiler. We wish to thank the medical and administrative staff of Moses H. Cone Hospital, Greensboro, NC and, especially, the mothers and their infants for making this research possible. Thanks also to G. Gottlieb, K. Harter, R. Hunt, R. Panneton, K. Smith, and, especially, W. Salinger for their helpful comments on drafts of the manuscript. Portions of this paper were presented at the Third Biennial International Conference on Infant Studies, March 1982, Austin, TX.

Correspondence and requests for reprints should be addressed to Anthony J. DeCasper, Department of Psychology, University of North Carolina at Greensboro, Greensboro, NC 27412.
The newborns' preference for their own mothers' voices requires that they had some prior experience with her voice, but there is no evidence that the necessary experience occurred after birth. Fifer (1980) failed to find any relation between maternal-voice preference and postnatal age, whether the newborns roomed with their mother or in a nursery, or whether they were breast fed or bottle fed. Since the maternal voice is audible in utero, and since third-trimester fetuses can hear, perhaps the necessary experience occurred before birth. In contrast, newborns show no preference for their own fathers' voices, even if they had explicit postnatal experience with his voice. Since male voices are not very audible in utero, perhaps the absence of a paternal-voice preference indicates the absence of prenatal experience with his voice (DeCasper & Prescott, 1984). The correlations between the presence or absence of specific-voice sounds before birth, and the presence or absence of specific-voice preferences after birth suggest that prenatal auditory experiences influence the earliest voice preferences.

Consider that complex auditory stimuli can function as positive reinforcers, neutral stimuli, or negative reinforcers of newborn behavior. Known reinforcers include vocal-group singing, solo female singing, prose spoken by a female, synthetic speech sounds, and intrauterine heartbeat sounds (Butterfield & Cairns, 1974; Butterfield & Siperstein, 1972; DeCasper, Butterfield, & Cairns, 1976; DeCasper & Carstens, 1981; DeCasper & Sigafoos, 1983). On the other hand, male speech and instrumental music lack reinforcing value, while white noise and faster-than-normal heartbeat sounds are aversive (Butterfield & Siperstein, 1972; DeCasper & Prescott, 1984; Salk, 1962). The differential reinforcing effectiveness of these sounds seems to covary more with their similarity to sounds that were present in utero than with any general acoustic characteristic(s), which further suggests that prenatal auditory experience influences postnatal auditory perception.

Finally, prenatal auditory experience has been shown to cause postnatal auditory preferences in a variety of infrahuman species (e.g., Gottlieb, 1981; Vince, 1979; Vince, Armitage, Walser, & Reader, 1982).

The hypothesis implies that prenatal experience with maternal speech sounds causes some property of the sounds to be differentially reinforcing after birth. Speech sounds enable at least two kinds of discriminations; some speech cues allow discrimination of language-relevant sounds, per se, or what is said, and some allow discrimination of the speaker or source of the speech sounds (Bricker & Pruzansky, 1976; Studdert-Kennedy, 1982). Thus, the prenatal experience hypothesis implies that newborns prefer their own mothers' voices, regardless of what she says, because of prenatal experience with her voice-specific cues. This implication, however, cannot be directly tested for obvious ethical and practical reasons. The hypothesis also implies that newborns will prefer the acoustic properties of a particular speech passage if their mothers repeatedly recite that passage while they are pregnant.

We directly tested the latter implication in the following way. First, pregnant women tape-recorded three separate prose passages. Then, they recited
one of the passages, their target passage, aloud each day during the last 6 weeks of pregnancy. After birth their infants were observed in an operant learning task where recordings of the target passage and a novel passage, one their mothers had recorded but had not recited, were both available as reinforcers. Then their relative reinforcing effectiveness was evaluated. If the prenatal experience with the target passage increases its reinforcing value then: (a) the acoustic properties of the target passage will be more reinforcing than those of a novel passage; (b) the differential reinforcing value of the target passage should be carried by its language-relevant cues and, thus, should not require the presence of the infant's own mother's voice cues; and (c) the reinforcing values of the target and novel passages should not differ for control newborns who had never been exposed to either passage.

**METHOD**

**Prenatal Phase**

*Pregnant Subjects.* Thirty-three healthy women approximately 7½ months pregnant were recruited from childbirth preparation classes after being informed about the project. All were experiencing uncomplicated pregnancies.

*Prenatal Procedures.* After becoming familiar with three short children's stories they tape-recorded all three. Recordings were made in a quiet room on an Akai 4000 stereophonic tape recorder. The tapes would be used as reinforcers in a postnatal learning task. Each woman was then assigned one of the stories as her target story. Assignment was made after all three had been recorded to prevent them from biasing the recording of their target, for example, by exaggerated intonation.

The women were instructed to read their target story aloud "two times through each day when you feel that your baby (fetus) is awake" and to "read the story in a quiet place so that your voice is the only sound that your baby can hear." They maintained a log of their daily recitations and were occasionally checked by the researchers.

*Story Materials.* The stories were *The King, the Mice, and the Cheese* (Gurney & Gurney, 1965), the first 28 paragraphs of *The Cat in the Hat* (Seuss, 1957), and a story we called *The Dog in the Fog*, which was the last 28 paragraphs of *The Cat in the Hat* with salient nouns changed. The three stories were about equally long, they contained 579, 611, and 642 words, respectively. Each could be comfortably recited in about 3 min. Each was also composed from equal size vocabularies of 152, 142, and 154 words, respectively. Salient, high-frequency nouns common to at least two stories were changed. For example, cat and hat in *The Cat* became dog and fog in *The Dog*, and cat and dog from those stories became turtle and zebra in *The King*. *The Cat* contained 46 unique words (i.e., words that appeared only in *The Cat*), which accounted for 22% of the total word count; *The Dog* contained 57 unique words, which accounted for 22% of the total word count; and *The King* contained 85 unique
words, which accounted for 44% of the total word count. All three stories contained common high-frequency words. For example, *a, all, and, did, do, he, I, in, like, not, now, of, said, that, the, to, with, and you* occurred at least three times in each. The common high-frequency words accounted for 43% of *The Cat*, 38% of the *The Dog*, and 36% of the *The King*. The remaining words occurred at least once in at least two of the stories. The stories also differed in prosodic qualities, such as patterns of syllabic beats. Thus, they differed in the acoustic properties of individual words as well as in prosody. *The Cat* and *The Dog* sounded more similar to each other than either did to *The King*, but we could readily identify the origin of short (several seconds) segments from all three.

Postnatal Phase

*Experienced Newborns.* Sixteen of the 33 fetal subjects completed testing as newborns. The 16 had been prenatally exposed to their target story an average of 67 times or for about 3.5 hours in all. They were tested at an average age of 55.8 hours (SD = 10). Each had to have had an uncomplicated full-term gestation and delivery, a birth weight between 3500-3900 grams, and APGAR scores of 8, 9, or 10 at 1 and 5 min after birth. If a subject was circumcised, he was not tested until at least 12 hours afterward. Parents gave informed consent for the testing and were invited to observe.

Seventeen infants were not tested or did not complete a test session: 5 because their mothers failed to return their logs, 4 because they encountered intrapartum or postpartum difficulties, 5 failed to meet state criteria at the time of testing or cried, and 3 subjects’ sessions were unavoidably interrupted.

*Apparatus.* Sessions occurred in a quiet, dimly lit room adjacent to the nursery. The infants lay supine in their bassinets and wore TDH-39 earphones, which were suspended from a flexible rod. They sucked on a regular feeding nipple with the hole enlarged to 1 mm. Rubber tubing connected the nipple to a Statham P23AA pressure transducer that was connected to a Grass polygraph and solid state programming and recording components. Each infant heard a tape recording of his/her target story and a tape recording of a novel story, one of the others their mother had recorded but not recited. Both stories were recorded by the same woman, and each was played on separate channels of the stereo recorder. The tape ran continuously, and sound was electronically gated to the earphones by the automated programming equipment. Intensities averaged 70 dB SPL at the earphones.

*Testing Procedures.* Sessions began about 2.5 hours after a scheduled feeding in order to maximize the chance of obtaining an awake, alert, and cooperative infant (Cairns & Butterfield, 1974). Each infant was brought to a quiet-alert state before testing could begin (Wolff, 1966) and had to visually fixate and follow an experimenter’s face when he/she spoke to the infant. (If
the infant was not alert and did not fixate or follow, he/she was returned to
the nursery, and another attempt was made after a later feeding.)

The infant was then placed supine in the bassinet and the earphones were
locked in place. One researcher, who could not be seen by the infant and who
was blind to the exact experimental condition in effect, held the nonnutritive
nipple loosely in the infant's mouth. Another monitored the equipment. The
infant was then allowed 2 min to adjust to the situation and had to emit sucks
having negative pressures of at least 20-mm Hg, a pressure normally exceeded
by healthy infants. (If the infant failed to suck adequately, he/she was returned
to the nursery, and another attempt was made after a later feeding.)

Testing began with 5 min of baseline sucking during which no voices were
presented over the earphones. Unconstrained nonnutritive sucking occurs as
groups or bursts of individual sucks separated by interburst intervals of several
seconds. A sucking burst was defined as a series of individual sucks separated
from one another by less than 2 s; when 2 s elapsed without a suck the equip-
ment registered the end of the burst. Thus, interburst intervals (IBIs) began 2 s
after the last suck of one burst and ended with the onset of the the first suck of
the next burst. This criterion accurately captures the burst-pause pattern of
newborns' nonnutritive sucking (see Figure 1). IBIs tend to be unimodally dis-
tributed for individual infants, and modal values vary between infants. The
baseline was used to estimate the distribution and median value of each infant's
IBIs just before reinforcement began. Differential reinforcement of IBIs began
after baseline had been established. (Hereafter, if the infant stopped sucking

Figure 1. Polygraph record of a newborn's nonnutritive sucking. Wide horizontal
marks indicate the onset and offset of a sucking burst. The time between the end
of one burst and the beginning of the next denotes an interburst interval. Onset
of the narrow event mark denotes that the time criterion, t seconds, has elapsed
since the end of the last burst. Vertical lines indicate time in seconds.
for two 1-min periods for any reason, he/she was returned to the nursery and not tested again.)

Reinforcement Contingencies. For eight randomly selected infants, sucking bursts that terminated IBIs equal to or greater than the infants' baseline medians \((t)\) produced the recording of a woman's voice reciting the target story. Bursts terminating IBIs less than the baseline median were reinforced with the same woman's recording of a novel story. Thus, only one of the two stories was presented binaurally with the first suck of a burst and remained on until the burst ended. Reinforcement contingencies were completely controlled by the solid-state equipment. Reinforcement contingencies were reversed for the other eight newborns, to control for the effects of any response bias that might arise from either of the contingencies or from changes in the behavioral dispositions of the infants, for example, arousal or fatigue. Differential reinforcement lasted about 20 min.

The same differential reinforcement procedures were used in earlier voice-preference studies (DeCasper & Fifer, 1980; DeCasper & Prescott, 1984). The rationale is based on well established reinforcement procedures that differentiate the temporal properties of behavior: Differentially reinforcing a range of IBIs causes the shorter differentially reinforced IBIs to increase in frequency (see newborn studies by DeCasper & Fifer, 1980; DeCasper & Sigafoos, 1983; as well as animal studies by Anger, 1956; Catania, 1970; DeCasper & Zeiler, 1977; Malott & Cumming, 1964).

Subject Controls. Twelve control newborns matched to a prenatally experienced counterpart on sex, race, and median interburst interval of baseline were also tested. They met the same selection criteria and were tested under exactly the same conditions as their counterparts, but their mothers had never recited any of the three stories.

Other Experimental Controls. The influence of mother-specific voice cues on the reinforcing effects of the target stories was controlled by having nine newborns reinforced with recordings made by their own mother and seven with recordings made by some other infant's mother. Both stories heard by an infant were recorded by the same woman to insure that their reinforcing value could not be unequally influenced by the speaker's voice characteristics. The acoustic properties of any one story could not systematically influence the reinforcing value of the target because each of the three stories had served as the target at least four times. No particular combination of target/novel pairings could systematically influence the reinforcing value of the target because five of the six possible target/novel pairings occurred at least twice. Unpredictable subject loss prevented precise counterbalancing of voices and target/novel pairings (see Table 1).
TABLE 1
Conditional Probability of Responding with \((0.0t \leq IBI < 0.4t)\) and \((1.0t \leq IBI < 1.4t)\)
During Baseline and During Reinforcement for Experienced and Control Subjects

<table>
<thead>
<tr>
<th>Sex</th>
<th>Target/Novel</th>
<th>Criteria for Target</th>
<th>Voice</th>
<th>Experienced Subjects</th>
<th>Control Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>Cat/Dog</td>
<td>&lt;5</td>
<td>M</td>
<td>.11</td>
<td>.06</td>
</tr>
<tr>
<td>F</td>
<td>Dog/Cat</td>
<td>&lt;6</td>
<td>M</td>
<td>.06</td>
<td>.26</td>
</tr>
<tr>
<td>F</td>
<td>Dog/King</td>
<td>&lt;3</td>
<td>M</td>
<td>.12</td>
<td>.10</td>
</tr>
<tr>
<td>M</td>
<td>King/Dog</td>
<td>&lt;6</td>
<td>M</td>
<td>.10</td>
<td>.16</td>
</tr>
<tr>
<td>M</td>
<td>Cat/King</td>
<td>&lt;3</td>
<td>O</td>
<td>.14</td>
<td>.16</td>
</tr>
<tr>
<td>F</td>
<td>Dog/King</td>
<td>&lt;4</td>
<td>O</td>
<td>.02</td>
<td>.15</td>
</tr>
<tr>
<td>M</td>
<td>Dog/King</td>
<td>&lt;4</td>
<td>O</td>
<td>.09</td>
<td>.12</td>
</tr>
<tr>
<td>M</td>
<td>King/Dog</td>
<td>&lt;6</td>
<td>O</td>
<td>.03</td>
<td>.09</td>
</tr>
<tr>
<td>M</td>
<td>Cat/Dog</td>
<td>≥3</td>
<td>M</td>
<td>.29</td>
<td>.37</td>
</tr>
<tr>
<td>M</td>
<td>Dog/Cat</td>
<td>≥5</td>
<td>M</td>
<td>.23</td>
<td>.21</td>
</tr>
<tr>
<td>F</td>
<td>Dog/King</td>
<td>≥3</td>
<td>M</td>
<td>.16</td>
<td>.17</td>
</tr>
<tr>
<td>M</td>
<td>King/Dog</td>
<td>≥3</td>
<td>M</td>
<td>.26</td>
<td>.32</td>
</tr>
<tr>
<td>M</td>
<td>Cat/King</td>
<td>≥5</td>
<td>M</td>
<td>.17</td>
<td>.31</td>
</tr>
<tr>
<td>F</td>
<td>King/Dog</td>
<td>≥3</td>
<td>O</td>
<td>.06</td>
<td>.08</td>
</tr>
<tr>
<td>F</td>
<td>Dog/King</td>
<td>≥3</td>
<td>O</td>
<td>.03</td>
<td>.20</td>
</tr>
<tr>
<td>F</td>
<td>Cat/King</td>
<td>≥6</td>
<td>O</td>
<td>.23</td>
<td>.28</td>
</tr>
</tbody>
</table>
Data Analysis. Interburst intervals were read off the polygraph records. Times between the event marks signalling the end of one burst and the beginning of the next burst were measured and rounded down to the nearest whole second (see Figure 1). Thus, the scorers (AJD and MJS), who were highly practiced, did not have to make detailed judgments about IBI values that might bias the data. Interscorder reliability approached 100%.

Next, each subject’s IBIs from the baseline and reinforcement phases were converted to a proportion of their time criterion ($t$). For example, if $t = 4$ s then all 2-s IBIs had the value 0.5$t$, and if $t = 6$ s then 2-s IBIs had the value of 0.33$t$. Converted IBIs were grouped into bins that were 0.2$t$ s wide; Bin 1 contained IBIs between 0.0$t$ and 0.2$t$ s, Bin 2 contained IBIs between 0.2$t$ and 0.4$t$ s, ..., and Bin 10 contained IBIs between 1.8$t$ and 2.0$t$ s. Bin 11 contained all IBIs greater than 2.0$t$ s. IBIs were assumed to be equally distributed within a bin. The conversion equates the relative size of IBIs across subjects and allows averaging over subjects.

RESULTS

Experienced Newborns

The hypothesis asserts that in utero exposure to the acoustic properties of the target story will make it more reinforcing than the novel story. If so, the relative frequency of short IBIs should increase over baseline when reinforced by the target stories in the $IBI < t$ condition and the relative frequency of IBIs slightly longer than the baseline median should increase when reinforced by target stories in the $IBI > t$ condition.

Baseline IBI distributions were examined first in order to determine whether they differed between reinforcement contingencies. They did not differ: A mixed ANOVA of the relative frequencies of baseline IBIs, with Contingencies ($< t$ vs. $> t$) and Bin (1-10) as factors, indicated a significant effect of Bin, $F(9, 126) = 13.3$, $p < .001$. The effect merely confirms that the IBIs were unimodally distributed. Most important, there was no Contingency effect, $F(1,14) p > 1.0$, and no Contingency $\times$ Bin interaction, $F(9, 126) = 1.67$, $p > .10$.

The predictions of the hypothesis were first assessed by examining the differences between the relative frequencies of IBIs that occurred during baseline and those that occurred during reinforcement. Difference scores were entered into a mixed ANOVA with Contingencies ($< t$ vs. $> t$) and Bin (1-10) as factors. There was no effect of Contingency, $F(1,14) < 1.0$, and a significant effect of Bin, $F(9,126) = 5.48$, $p < .025$. Most important, there was a significant Contingency $\times$ Bin interaction, $F(1,126) = 2.07$, $p < .05$. Planned tests of simple effects confirmed that the interaction occurred because with the $IBI < t$ contingency the relative frequency of short IBIs increased over baseline levels, while those of all other IBIs either decreased or did not change. With the $IBI > t$ contingency the relative frequency of IBIs slightly greater than $t$ seconds increased, while those of the others decreased or did not change. Any IBI between 0 and $t$ seconds would have produced the target story under the
IBI < t contingency, and any IBI ≥ t seconds would have produced it under the IBI > t contingency. But only the relative frequencies of the shorter IBIs reinforced by the targets systematically increased.

The differential reinforcement effects are more clearly revealed in the analysis of IBIs between 0.0t and 0.4t (the shorter IBIs) and those between 1.0t and 1.4t (IBIs slightly longer than t seconds). Conditional probabilities of baseline and reinforced IBIs in these classes were obtained by dividing the relative frequency of IBIs in each class by the relative frequency of that class and all longer IBIs (see Table 1). This is a sensitive measure of temporally differentiated responding because: (a) it adjusts the inherently unequal opportunity for infants to emit equal numbers of short and long IBIs in a limited period of time; (b) it measures the probability that an infant will emit a particular class of IBIs given the opportunity to do so (cf. Anger, 1956; DeCasper & Fifer, 1980); and (c) it renders the conditional probabilities of IBIs between 0.0t and 0.4t, and those between 1.0t and 1.4t, arithmetically independent of one another. The dependent variables for the target story and for the novel story were their reinforcement ratios: (conditional probabilities of IBIS during reinforcement) divided by (conditional probability of IBIS during reinforcement) plus (conditional probability of IBIS during baseline).

The average values of baseline conditional probabilities of target-story IBIs and novel-story IBIs did not differ, t(15) = 1.37, p > .10. However, their reinforcement ratios differed as expected. A mixed ANOVA with Contingency (< t vs. > t) and Interval (0.0t–0.4t vs. 1.0t–1.4t) as factors, revealed no effect of Contingency, F(1,14) < 1.0, and no effect of Interval, F(1,14) = 1.59, p > .20, but a significant Contingency × Interval interaction, F(1,14) = 6.65, p < .025 (Figure 2). Target-story reinforcement ratios were larger than novel-story reinforcement ratios, independent of the contingency and of the interval. The fact that 13 of the 16 infants had larger target ratios than novel ratios (p = .011 by the binomial test) and 13 of the 16 had target-story ratios greater than .50 indicates this result was typical. The individual-subject consistency implies that maternal voice cues were not necessary for producing the differential reinforcement effect. Neither the target-story reinforcement ratios nor the difference between the target ratios and novel ratios differed between the 9 infants who heard their own mothers' voices and the 7 who heard unfamiliar voices, p-values of both t-tests > .10.

Control Subjects

The following analysis of control-subject performances parallels that of the experienced subjects. The relative frequency distributions of baseline IBIs did not differ between reinforcement contingency conditions. A mixed ANOVA with Contingency (< t vs. > t) and Bin (1–10) as factors revealed a marginal effect of Bin, F(9,90) = 1.90, .10 < p > .05, but no effect of Contingency, F(1,10) < 1.0, or of the Contingency × Bin interaction, F(9,90) = 1.38, p > .10. The subsequent mixed ANOVA on the difference scores of IBIs that occurred
during the baseline and reinforcement phases revealed no effect of Contingency, $F(1,10) < 1.0$, a significant effect of Bin, $F(9,90) = 5.19$, $p < .001$, and a significant Contingency $\times$ Bin interaction, $F(9,90) = 3.48$, $p < .005$. However, none of the follow-up tests of simple effects were statistically reliable; the interaction seemed to result from unsystematic variation in the difference scores of the two contingency conditions in Bins 1-5.

Subsequent analysis of conditional probabilities confirmed that the preceding interaction did not result from systematic effects of target-story reinforcement. The baseline conditional probabilities of target and novel stories did not differ, $t(11) < 1.0$; neither did their reinforcement ratios computed for the intervals $0.0t-0.4t$ and $1.0t-1.4t$. The mixed ANOVA with Contingency and Interval as factors revealed no reliable effects whatever, $p$ values of all $F$ statistics $> .10$ (Figure 3).

A comparison of the reinforcement ratios of matched-subject pairs revealed that experienced newborns had larger target-story ratios than their matched naive counterparts, $t(11) = 2.68$, $p < .05$, but that their novel-story ratios did not differ, $t(11) < 1.0$.

**DISCUSSION**

Three implications of the prenatal-experience hypothesis were confirmed: (1) For experienced subjects the target story was more reinforcing than the novel
The conclusion implies that the fetuses had learned and remembered something about the acoustic cues which specified their particular target passage (e.g., prosodic cues such as syllabic beat, the voice-onset-time of consonants, the harmonic structure of sustained vowel sounds, and/or the temporal order of these sounds). Recall also that newborns prefer their mothers' voices over
that of another female, when both speak the same novel material (DeCasper &
Fifer, 1980; Fifer, 1980). The present results add to the evidence indicating
that the maternal-voice preference also originated in utero. If so, then fetuses
also register some specific information about their mothers’ voices (e.g., spectra
of nasals and vowels, glottal frequency and spectrum, and/or the temporal
characteristics of pitch, intensity, and formants) (Bricker & Pruzansky, 1976).
The specific acoustic cues that register in utero and which influence subsequent
perception of speech and voice sounds are not known at present. However,
whether language-relevant cues or voice-specific cues play an active role in
newborns’ perception has now been shown to depend upon: (a) which class of
cues are differentially available, (b) the infants’ prenatal experience with the
cues, and (c) the circumstances attending postnatal perception (e.g., behavioral
contingencies, infant state, or the presence or absence of other sounds).

The present study suggests noninvasive, ethically acceptable methods to
further study the effects of prenatal auditory stimulation on postnatal auditory
function and development, especially the development of speech perception.
Such research might also benefit clinical treatment of the perinate, for exam-
ple, by aiding in the diagnosis of fetal condition and by providing information
for designing environments of preterm infants.

Some Post Hoc Considerations

Learning is generally and most satisfactorily inferred from a change in perfor-
manee rather than from absolute measures of performance. However, change
scores—the difference scores and reinforcement ratios used in this study—are
almost always inversely related to prelearning performance, the baseline proba-
focuses on the extent to which the preceding inferences about differential-
reinforcement effects were influenced by the relation between baseline levels of
performance and the difference scores and reinforcement ratios. The issue is
salient here because the hypothesis asserts that reinforcement would differential-
ly affect specific IBIs whose baseline probabilities varied considerably.

The abscissa of Figure 4 shows the mean conditional probabilities of base-
line IBIs for each of the eight subconditions represented in Table 1. The mean
baseline conditional probabilities of IBIs between 0.0t and 0.4t (subconditions
1–4 in Figure 4) are lower than the mean baseline conditional probabilities of
IBIs between 1.0t and 1.4t (subconditions 5–8). They differ simply because
baseline IBIs between 0.0t and 0.4t come from the left of a unimodal distribu-
tion and IBIs between 1.0t and 1.4t come from near the median of the distribu-
tion. The primary means of experimentally controlling for the influence of
these baseline differences was to counterbalance the reinforcers associated
with the IBI < t and IBI > t contingencies: As Figure 4 suggests, and as reported
earlier, when the values of baseline probabilities are pooled over IBI < t and the
IBI > t contingencies (1 with 6; 2 with 5; 3 with 8; 4 with 7) the average baseline
probabilities do not differ.
Figure 4. Mean conditional probability that subjects in the eight subconditions would emit IBIs between 0.01-0.41 and between 1.01-1.41 during reinforcement as a function of the mean conditional probability that they would do so during baseline. Open circles refer to Experienced subjects reinforced by the target story with IBI < t (2) and with IBI > t (5). Filled circles refer to Experienced subjects reinforced by the novel story with IBI < t (4) and with IBI > t (7); to Control subjects reinforced by the target story with IBI < t (3) and with IBI > t (8); and to Control subjects reinforced by the novel story with IBI < t (1) and with IBI > t (6). The solid line represents the regression equation (.07 + .3B [baseline probability]) for the six control subconditions (filled circles) and the dashed lines represent the 95% confidence interval around the regression line.

Figure 4 also shows the empirical relation between the mean baseline probabilities and the mean probabilities occurring with reinforcement for each subcondition. The solid line represents the regression equation relating the baseline and reinforcement probabilities for the six subconditions in which no differential-reinforcer effect was expected (filled symbols), \( r = .89, p < .02 \). For these six subconditions the probability of responding during reinforcement is almost completely determined by the prior baseline probability. Their reinforcement probabilities do not increase over baseline probabilities, but instead become increasingly smaller than baseline as the baseline probability increases. Figure 5 shows that when reinforcement ratio (a change score) is substituted for reinforcement probability (the absolute score), the strong linear relation between baseline performance and reinforced performance is preserved, but for statistical and mathematical reasons, the correlation is negative, \( r = -.93, p < .01 \).
Since the means of the subcondition baselines were not equal these, correlations raise an important question. Might the reinforcement probabilities and reinforcement ratios that resulted when Experienced newborns were reinforced with their target story be determined simply by their baseline probabilities? That is, do the differences in the subgroups' terminal performances, as measured by reinforcement probabilities or difference scores and reinforcement ratios, reflect differential reinforcement effects or just the fact that the subgroups began with different baseline probabilities?

Figure 4 shows that the mean baseline probabilities of the two conditions where Experienced subjects were reinforced with their target story (open symbols) fall within the range of baseline probabilities entailed in the correlation. Significantly, however, the mean probabilities that occur with reinforcement by the target story are both above their baseline levels and above the 95% confidence interval of the regression line ($p < .0006$). Similarly, both reinforcement ratios are well above .50 and also above the 95% confidence interval of the regression line of Figure 5. Thus, the possibility that the reinforcement probabilities and reinforcement ratios occurring when Experienced subjects were reinforced with their target story were determined by or could be predicted by

![Figure 5](image-url)

**Figure 5.** Reinforcement ratio as a function of baseline conditional probability for the eight subconditions described in Figure 4. The regression equation is $.57 - .63 \times \text{baseline probability}$ for the control subconditions and the dashed lines represent the 95% confidence interval.
their baseline probabilities can be rejected. The favored alternative hypothesis, of course, is that prenatal experience increased the reinforcing effectiveness of their target stories: The effect of prenatal experience with the reinforcer was to increase the conditional probability of reinforced responding by 40% over the level predicted by baseline in the IBI < t condition and by 76% in the IBI > t condition. Reinforcement ratios were increased by 20% and 26% over the levels predicted by baseline performance.

It may still be argued, however, that the preceding analysis was based on subgroup means and that the pattern of individual-subject baseline probabilities within the subgroups was biased toward producing difference scores and reinforcement ratios that supported the prenatal hypothesis. That is, if the baseline probability of each Experienced subject reinforced with the target story had been the same as the baseline of a control subject, then their reinforcement probabilities and reinforcement ratios might not differ.

The following analyses addressed this possibility by comparing selected groups of subjects after matching individual infants on baseline probabilities. Subject matching was accomplished by applying the following three rules: (1) baseline probabilities had to be within ±.02 of each other; (2) if possible, the subjects were to have the same reinforcement contingency; and (3) if more than one match was possible, pairs were matched so as to minimize the difference between conditional probabilities that occurred with reinforcement. No other factors were considered.

In the first comparison, 10 of the 16 baseline probabilities produced by Experienced infants in the conditions where they were reinforced with the target story were matched to the baseline probabilities of 10 of the 16 Experienced subjects in the conditions where they were reinforced with the novel story. The 10 baseline pairs were: (.06/.06), (.06/.06), (.09/.09), (.10/.10), (.12/.12), (.14/.13), (.16/.17), (.17/.17), (.23/.24), and (.26/.26). The mean baseline probability for each group was .14. The mean probability occurring with reinforcement by the target story (.20) was greater than that occurring with reinforcement by the novel story (.15), t(9) = 2.61, p < .01 (1-tail t test) T = 4.5, p < .025 (Wilcoxon test). the mean probability of responding with target-story reinforcement was greater than the baseline mean, t(9) = 2.99, p < .005; T = 3, p < .01, but the mean probability of responding with novel-story reinforcement did not differ from baseline, t(9) = .58; T = 22.5. In addition, the mean occurring with target-story reinforcement was well above the 95% confidence interval of Figure 4, but the mean occurring with novel-story reinforcement was well within the interval. The reinforcement ratio of the target story was well above the 95% confidence interval of Figure 5, but the reinforcement ratio for the novel-story was within the interval.

Six infants from each reinforcement condition could not be matched. The mean of the unmatched baselines for the target-story condition was .12; the mean occurring with reinforcement was .18. For the novel-story infants
these means were .22 and .17, respectively. The reinforcement mean and reinforcement ratio occurring with target-story reinforcement were above the 95% confidence intervals of Figures 4 and 5, but the analogous measures resulting from reinforcement with the novel story were well within the confidence intervals.

Next, baselines of Experienced subjects who were reinforced with the target story were matched to baselines of Control subjects reinforced with the target story. Nine pairs could be formed: (.03/.04), (.06/.06), (.06/.06), (.09/.09), (.10/.10), (.14/.13), (.16/.16), (.26/.26), and (.29/.31). The baseline mean of each group was .13. The mean reinforcement probability for the Experienced subjects (.19) was greater than that of Control infants (.10), \( t(8) = 2.58, p < .01; T = 1, p < .005 \). The Experienced subjects’ reinforcement probabilities were larger than baseline probabilities, \( t(8) = 2.94, p < .005; T = 0, p < .005 \), but the Control subjects’ were not, \( t(8) = 1.19, p > .10; T = 15 \). Here, too, the Experienced subject’s mean reinforcement probability and reinforcement ratio were both well above the 95% confidence intervals of Figures 4 and 5. The reinforcement probability of the Control group was within the interval of Figure 4. Their reinforcement ratio, however, fell below the 95% confidence interval of Figure 5, even though the members of the group had exactly the same baseline probabilities as their Experienced counterparts.

Seven Experienced subjects’ and three Control subjects’ baselines could not be matched. The baseline means of these subjects are .13 and .40, respectively. Their respective means occurring with reinforcement by the target story were .19 and .20. The reinforcement mean and reinforcement ratio of the Experienced subjects both lay well above the 95% confidence intervals of Figures 4 and 5. Analogous scores for Control subjects were within the intervals.

Finally, Experienced subjects reinforced with the novel story were matched to Control subjects who were also reinforced with the novel story. The nine pairs of baseline probabilities were: (.06/.06), (.09/.09), (.10/.10), (.12/.12), (.13/.13), (.19/.20), (.19/.19), (.24/.24), and (.33/.32). The mean for each group was .16. Neither the between-group difference nor the changes from baseline were statistically reliable (all \( t \) values < 1.0; all \( T \) values > 16). All reinforcement means and reinforcement ratios fell within the confidence intervals of Figures 4 and 5. Seven Experienced infants and three Control infants could not be matched. Their respective baseline means were .19 and .16, and their respective reinforcement means were .18 and .13. All reinforcement means and reinforcement ratios were within the confidence intervals of Figures 4 and 5.

After equating the baseline probability of IBIs of individual infants in specific conditions, the only consistent finding was that the target story was the more effective reinforcer for Experienced infants. In sum, the results of this study cannot be accounted for by differences in the baseline values of subconditions or individual subjects. The previous conclusion can be retained: The postnatal reinforcing value of a speech passage is increased by prenatal experience with the passage.
REFERENCES


1 January 1985; Revised 20 December 1985