Echoic Storage in Infant Perception

Nelson Cowan, Karen Suomi, and Philip A. Morse
University of Wisconsin–Madison

Cowan, Nelson; Suomi, Karen; and Morse, Philip A. *Echoic Storage in Infant Perception.* Child Development, 1982, **53,** 984–990. Preperceptual auditory or "echoic" storage was investigated in 8–9-week-old infants using a modification of an adult masking paradigm and a nonnutritive sucking discrimination procedure. Experiment 1 provided validation of a new version of the nonnutritive sucking procedure usng the standard stimulus contrast [ba] versus [pa]. In experiment 2, infants were presented with repeating pairs of brief vowels with a stimulus onset asynchrony (SOA) of 50 msec for each pair. Within each series, the first vowel in a pair changed (backward masking), the second vowel changed (forward masking), or neither vowel changed (control). Discrimination of the change occurred in the forward- but not in the backward-masking condition. In experiment 3, discrimination occurred in a backward-masking condition with an SOA of 400 msec, but not with an SOA of 250 msec or in a control condition. In conjunction with the adult literature, these results suggest that echoic storage contributes to auditory perception in infancy, as in adulthood, but that the useful lifetime of an echoic trace may be longer in infancy.

The present paper demonstrates that one of the earliest stages of information processing, the preperceptual storage of auditory information in an unanalyzed "echoic" form, can be observed in infancy by modifying the forward- and backward-masking procedures used with older subjects. In the standard masking procedure (e.g., Massaro 1972, 1973) two brief sounds are presented in rapid succession and subjects must identify the first sound (in backward masking) or the second sound (in forward masking) in a forced choice. Because the second sound in a pair interferes with the echoic storage of the first sound at relatively short stimulus onset asynchronies (SOA), performance generally is better with forward than with backward masking. Moreover, interference with performance in backward masking decreases to an asymptotic level at an SOA of about 250 msec. This may be the period for which echoic storage is useful in the auditory recognition process.

There has been little or no research on the maturation of echoic storage. However, the properties of echoic storage undoubtedly are related to the organism's auditory- and speech-perception abilities. For example, a longer-lasting echoic trace in infancy might help the infant to compensate for a slower processing rate. Recently, Lasky and Spiro (1980) have reported work in the visual modality supporting the suggestion that preperceptual processing in infancy outlasts the 250-msec duration that has been observed in masking studies with adult subjects.

In order to study echoic storage in infancy, a masking procedure was used in which repeating pairs of brief sounds were presented. Vowel sounds were employed, because infants have been shown capable of discriminating brief vowels in isolation (Swoboda, Kass, Morse, & Leavitt 1978). Within each series of vowel pairs in the present study, there was a change in the first vowel (backward masking), the second vowel (forward masking), or neither vowel (control).

Discrimination Paradigm

To conduct testing in very young infants with a procedure that was as sensitive as possible, the nonnutritive sucking discrimination paradigm used by other investigators (e.g., Eimas, Siqueland, Jusczyk, & Vigorito 1971) was modified. The first experiment to be reported was not a masking experiment, but a validation of the modified infant paradigm with a standard [ba/pa] stimulus contrast.

We wish to thank Robert Engelke, Ken Kaiser, and Mike Epp for technical assistance, the Waisman Center computing staff at the University of Wisconsin, and the parents who brought their infants to listen to such unusual stimuli. This project was supported by NICHD grants HD08240 and HD03352. Address reprint requests to Philip A. Morse, Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.

[*Child Development,* 1982, 53, 984–990. 0009-3920/82/5304-0008$01.00]

Instead of the baseline-acquisition-satiation-recovery sequence that comprises a trial within the standard infant paradigm, the present procedure included only baseline and acquisition phases. Within the acquisition phase, infants in an experimental group received 30-sec blocks of one stimulus (in the masking experiments, a pair of stimuli) alternating with 30-sec blocks of a second stimulus (or second pair of stimuli) with presentations contingent upon high-amplitude sucking. In the control group, however, infants received contingent repetitions of the same stimulus (or pair) throughout the acquisition phase. Unlike the traditional sucking discrimination procedure, this new paradigm provides infants in experimental conditions with multiple stimulus shifts from the first to the second stimulus (pair) and vice versa, at 30-sec intervals. Discrimination is indexed by a greater rate of sucking across twenty 30-sec periods (relative to the last 30 sec of baseline) in an experimental group of infants than in the control group. The rationale for this expected pattern of results is that the stimulus shifts, if perceived by the infant, should result in less habituation to the stimulus tape.

Experiment 1

Method

Subjects

Infants from primarily upper-middle-class families in the Madison, Wisconsin, area were located through birth announcements, and the parents contacted by mail and follow-up phone calls. Each infant's data were included in the study only if the infant remained in a quiet, alert state for at least 30 sec within each successive minute of testing. Of 58 8–9-week-old normal infants tested, 28 (48%) remained in an acceptable state throughout the experiment (14 subjects each in the experimental and control groups).

Stimuli and Apparatus

The stimuli in the first study were not masking pairs, but synthetic 400-msec tokens of the consonant-vowel (CV) syllables [ba] (+20-msec voice-onset time) versus [pa] (+40-msec voice-onset time) used by Eimas et al. (1971). The stimuli were digitally stored and recorded on high-fidelity audiotape at the rate of one sound per second. Thirty-token blocks of [ba] alternated with 30-token blocks of [pa] on the experimental tape. The control tape contained only tokens of [ba] at 1-sec intervals.

Testing was conducted with the infant in an adjustable infant seat in an Audio-Suttle sound attenuated chamber. A closed-circuit television system allowed the infant to be viewed by the parents, the experimenter, and an assistant within the infant chamber who was hidden from the infant by a cloth partition. The assistant listened over headphones to music that masked the changes in speech stimuli, and indicated unacceptable infant state with a foot pedal that was connected to a polygraph channel outside of the chamber. Infants sucked on a nonnutritive nipple mounted on a plastic base connected to a Statham P23-BC pressure transducer by plastic tubing. A Grass model 7B polygraph amplified the output, and a potentiometer circuit permitted high-amplitude sucks (the peak 20%–50% of all sucks) to be selected and counted by an Automated Data Systems minicomputer (ADS 1800-E).

The stimuli were presented at 67 dB(A), measured with a General Radio 1551-C sound level meter against an ambient noise level of 45 dB(A). The stimulus tape was mounted on a TEAC 3300s tape deck, which was connected to an audiogate, a Crown D60 amplifier, and an ADS L710 speaker mounted in front of and above the infant. The minicomputer controlled the audiogate and permitted the presentation of each speech token to the infant only if it began within one second following a high-amplitude suck.

Procedure

Infants were not fed in the hour prior to testing. Once the infant was in an acceptable state within the test chamber, the rater placed the nipple in the infant's mouth and the experimenter adjusted the potentiometer to determine the high-amplitude sucking threshold. Criteria for entry into the acquisition phase were 9–20 high-amplitude sucks in 30 sec and 21–50 total high-amplitude sucks in 1 min within the silent baseline period. Infants generally met these criteria within the first 3 min of testing. When the computer display indicated that the acquisition period was about to begin, the experimenter activated the audiotape at the beginning of the first 30-sec stimulus block. Based on pilot data, sessions were terminated after the tenth postbaseline minute.

Results and Discussion

In the last 30-sec period of baseline, infants averaged 17.8 high-amplitude sucks. There were no significant differences between the experimental and control groups either in this silent baseline period or in the first 30-sec acquisition period, within which both groups heard only [ba] stimuli (t tests, p values $> .05$). However, a group × period ANOVA across 20 postbaseline periods indicated that infants in the experimental group sucked at a substantially higher rate than control infants, $F(1,26) = 7.50$, $p < .025$. Mean high-amplitude sucks relative to baseline equalled +7.1 per minute for infants in the experimental group versus −8.4 per minute for infants in the control group. There was also a group × period interaction, $F(19,494) = 2.56$, $p < .005$. An analysis of trends on the two groups indicated that this interaction was due to a difference in linear trend across 30-sec periods in the experimental group versus the control group, $F(1,26) = 11.16$, $p < .005$. These results are illustrated in figure 1. Thus, this modified nonnutritive sucking paradigm did provide strong evidence of discrimination of a standard stimulus contrast.[1]

The second experiment was conducted to confirm that echoic storage can be demonstrated in infancy using this new discrimination procedure. Instead of the single stimuli employed in experiment 1, stimuli were paired in experiment 2, with an SOA of 50 msec within each pair. At this brief SOA value, adult subjects generally demonstrate relatively little forward masking (interference of the first member of the pair with recognition of the second member), but the second stimulus in a pair strongly interferes with the echoic storage of the first stimulus, resulting in backward masking. It was expected, therefore, that at this SOA infants and adults would exhibit release from forward masking, but no release from backward masking. This should yield superior discrimination of a stimulus change in the forward-masking condition, relative to performance in the backward-masking and no-change control groups.

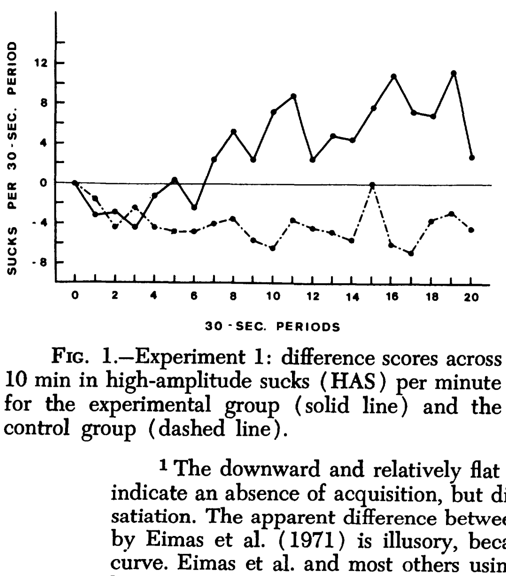

FIG. 1.—Experiment 1: difference scores across 10 min in high-amplitude sucks (HAS) per minute for the experimental group (solid line) and the control group (dashed line).

Experiment 2

Method

Subjects

Of 83 8–9-week-old infants tested, 54 (65%) met the behavioral state criteria (18 each in the forward, backward, and control groups). Additionally, 10 adult volunteers who had no prior exposure to the stimuli or knowledge of the experimental hypotheses participated in a comparison experiment.

Stimuli

Two 50-msec vowels were constructed with a software series synthesizer (Klatt 1980): the vowel [a] with F1 = 750 Hz, F2 = 1150 Hz, F3 = 2400 Hz; and the vowel [ε] with F1 = 275 Hz, F2 = 2250 Hz, F3 = 3000 Hz. The fourth and fifth formants were fixed at 3300 Hz and 3750 Hz, respectively, for both vowels. Band widths for formants 1–5 were 50, 70, 110, 250, and 200 Hz, respectively. Both vowels had 10-msec onset and offset ramps and a fundamental frequency that fell linearly from 200 Hz to 40 Hz across 50 msec. Stimulus pairs were recorded on audiotape with no silent time between stimuli in a pair (interstimulus interval = 0, SOA = 50 msec) and with a 900-msec silent interval between pairs.

Six tapes were constructed for the infant experiment: two tapes each for the forward, backward, and control conditions. One of the tapes for each condition began with [a-a] and the other began with [ε-ε]. On the forward-masking tapes, the second vowel in each pair changed once every 30 sec (e.g., [a-a] → [a-ε]).

[1] The downward and relatively flat trend observed in the control group probably does not indicate an absence of acquisition, but differences among infants in the rates of acquisition and satiation. The apparent difference between this downward trend and the upward trend observed by Eimas et al. (1971) is illusory, because the present result is based on a forward learning curve. Eimas et al. and most others using their paradigm have analyzed results with backward learning curves (however, cf. Trehub & Chang 1977). Backward curves are inappropriate for the present paradigm due to the lack of a reference point. Swoboda, Morse, & Leavitt (1976) obtained a similar initial downward trend during the early postbaseline period when the data were Vincentized. Note that this initial downward trend was obtained in all three of the present infant experiments.

On the backward-masking tapes, however, the first vowel in each pair changed (e.g., [a-a] → [ɛ-a]). On the control tapes, subjects heard a single repeated pair ([a-a] or [ɛ-ɛ]) throughout the session. Half of the infants in each group (forward, backward, and control) listened to the tape that began with [a-a], and half listened to the tape that began with [ɛ-ɛ]. Each vowel pair was presented at 70 dB(A) SPL.

Procedure

Adult experiment.—The purpose of this experiment was to ensure that the vowel pairs employed in the infant experiment yielded superior discrimination by adults in the forward-masking condition relative to the backward-masking condition. On each trial, adults received one vowel pair, a 900-msec delay interval, and then a second vowel pair, and were required to label the two stimulus pairs as "same" or "different." Every subject received a randomized presentation that included seven trials of each of the contrasts used in the infant experiment.

Infant testing.—The infant procedure was identical to that of the first experiment except for the changes in stimuli discussed above.

Results and Discussion

Adults

Adult listeners responded correctly to 97% of all control presentations (two [a-a] pairs or two [ɛ-ɛ] pairs) and 98% of all forward-masking contrasts ([a-a] vs. [a-ɛ], or [ɛ-ɛ] vs. [ɛ-a]) but only 75% of all backward-masking contrasts ([a-a] vs. [ɛ-a], or [ɛ-ɛ] vs. [a-ɛ]). Interference with echoic storage was indicated by significantly poorer performance in the backward condition than in the forward ($p < .005$) or control ($p < .01$) conditions (randomization tests for matched pairs). The levels of performance in the latter two conditions did not differ reliably from one another, $p > .05$.

Infants

The mean number of sucks in the last 30 sec of baseline was 18.28. A comparison with *t* tests revealed no group differences approaching significance in this baseline period. However, within the first 30-sec period of acquisition there was significantly more sucking in the forward- than in the backward-masking group, $t(34) = 3.08, p < .01$, even though both groups received the same stimulus presentations during that period. This difference probably reflects a sampling bias between groups. However, the magnitude and curvilinear trends of group differences across 10 min, shown in figure 2, and the validation of the discrimination paradigm by experiments 1 and 3 (to be discussed), strongly suggest that there was an effect of the between-group stimulus manipulation in the present experiment that did not result from this initial difference.[2]

A group (3) × period (20) × starting stimulus ([a-a] vs. [ɛ-ɛ]) analysis of sucking during the acquisition period revealed a significant main effect for groups, $F(2,48) = 7.21$, $p < .005$. Mean rates of high-amplitude sucking across 10 min of acquisition, expressed as difference scores, were: forward masking, $\bar{X} = +7.94$ per minute; backward masking, $\bar{X} = -9.12$ per minute; and control, $\bar{X} = -6.48$ per minute. Post hoc Scheffé tests indicated that the forward-masking group differed reliably from the backward-masking group ($p < .01$) and the control group ($p < .05$) but that the

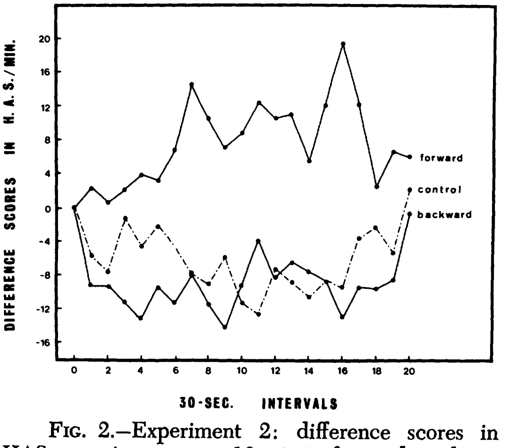

Fig. 2.—Experiment 2: difference scores in HAS per minute across 10 min in forward-masking, backward-masking, and control groups.

[2] The results of experiments 1 and 3 suggest that responding in the first 30-sec postbaseline period is quite variable and in no way systematically related to the pattern of responding across periods. In experiment 1, group differences in discrimination obtained when there was no evidence of a difference in the first 30-sec postbaseline period. Moreover, in experiment 3 a difference from the control group within the first 30-sec postbaseline period obtained for the 250-msec group (but not the 400-msec group), whereas evidence of reliable discrimination obtained for the 400-msec group (but not the 250-msec group). This evidence, as well as the orderly pattern of masking results in experiments 2–3 combined, suggests that the group differences in experiment 2 did not result from differential responding in the first 30-sec acquisition period.

latter two groups did not differ reliably from one another.

In addition, there was a small but significant group $\times$ period interaction, $F(38,912) = 1.53$, $p < .05$. As shown in figure 2, the mean rate of high-amplitude sucking in the forward-masking group increased to well above baseline, but generally remained below baseline in the other groups. Trend analyses on two groups at a time yielded a significant group $\times$ quadratic trend over periods in the forward-masking versus control comparison, $F(1,34) = 7.69$, $p < .01$, and also a significant group $\times$ cubic trend over periods, $F(1,34) = 5.72$, $p < .025$. These nonlinear trends reflect a return to baseline in both groups, perhaps due to habituation in the forward-masking group and a relatively delayed pattern of strong acquisition in the control group (which also resembled the backward-masking group). No other effects in the trend analysis or the main ANOVA were significant. In conjunction with the adult masking literature, the finding that infants were able to recognize a change in vowel stimuli in a forward- but not a backward-masking situation suggests that the second vowel in a pair was capable of interfering with the infant's use of preperceptual, echoic storage of the first vowel.

The third experiment addressed the issue of the temporal properties of echoic storage in infancy as compared with adulthood. In backward-masking paradigms, adults are able to extract information from an uninterrupted echoic trace for about 250 msec, and further silent processing time between a target and mask beyond this point of asymptote does not improve performance (Massaro 1972, 1973). However, given the neurological immaturity of the infant, the duration of echoic processing in infancy might exceed 250 msec. Experiment 3 examined this possibility by administering to one experimental group a backward-masking series with an SOA of 250 msec, and to a second experimental group a backward-masking series with a 400-msec SOA.

To present all infants with identically timed stimuli, thereby controlling behavioral state across conditions, a minor change in the paradigm was necessary. For each of the three groups, the tape contained a pseudorandom mixture of 250-msec and 400-msec SOA vowel pairs. However, on the tape for the 400-msec group, only the 400-msec vowel pairs changed. Similarly, on the tape for the 250-msec group, only the 250-msec pairs changed. On the control tape, no changes occurred. Finally, the same stimulus tapes were played to adult subjects in order to assess the comparability of performance in this procedure to performance in standard adult paradigms.

Experiment 3

Method

Subjects

Of 46 8–9-week-old infants tested, 30 (65%) remained in an acceptable state throughout the experiment. Additionally, 10 adult subjects with no prior exposure to the stimuli or knowledge of the experimental hypotheses listened to portions of the infant tapes.

Stimuli

Three tapes were constructed with the 50-msec [a] and [ɛ] vowels used in experiment 2. The results of the previous experiment revealed no significant differences between stimulus tapes beginning with [a-a] versus tapes beginning with [ɛ-ɛ]. Thus, in the present experiment all three tapes began with [a-a]. The control tape contained only [a-a]. On this tape, pairs with SOAs of 250 msec and 400 msec were randomized with the restriction that each 30-sec period contained 15 pairs with each of these two SOAs. The silent intervals between pairs were adjusted so that vowel pairs began at equal intervals, 1 sec apart.

The experimental tapes used the same randomization of SOAs as the control tape. However, in alternate 30-sec periods within the 250-msec condition tape, the pairs with a 250-msec SOA changed from [a-a] to [ɛ-a]. Similarly, in the 400-msec condition tape, the pairs with a 400-msec SOA changed from [a-a] to [ɛ-a].

Procedure

Adult experiment.—Half of the 10 adult subjects listened to the 250-msec condition tape, and half listened to the 400-msec condition tape. Each subject was tested individually. For the first minute, subjects were simply instructed to listen, following which they were presented with 2 min of the tape and instructed to raise their hands each time the sound changed in quality. The experimenter recorded the subjects' responses.

Infant testing.—The infant procedure was identical to that of experiment 2 except for the changes in stimuli discussed above.

Results and General Discussion

Adults

Out of a total of 300 opportunities for a correct hand raise, subjects "missed" only twice.

Moreover, out of 900 possible false alarms, only three such errors were made. These findings suggest that both SOAs resulted consistently in a release from backward masking in adults.

Infants

The mean number of sucks in the last 30-sec baseline period was 20.3, and t tests revealed no significant differences between groups in this period. In the first 30-sec period of acquisition, a group difference obtained that was dissimilar to the discrimination pattern obtained across 10 min of acquisition. In the first 30-sec acquisition period, infants in the 250-msec group produced significantly more sucks than the control group, $t(34) = 2.27$, $p < .05$. No other significant group differences obtained in this period. Unlike this response pattern in the first 30-sec acquisition period, the 400-msec group sucked at a rate that increased much more rapidly across 10 min than did the 250-msec or the control group (see n. 2 above).

As figure 3 illustrates, stable differences between group means began to emerge only after the sixth minute of the acquisition period (cf. figs. 1–3). This difference between the pattern of results in experiments 1 and 2 versus 3 is probably due to the change in the paradigm, which could have increased the task difficulty. In experiment 3, the shift to a dishabituation vowel occurred in only half of the vowel pairs, rather than in all vowel pairs, within each 30-sec shift block of stimuli on an experimental tape. Consequently, infants in experiment 3 may have required a longer period of exposure in order to achieve a comparable level of discriminative performance. In a group (3) × period (20) ANOVA, the main effect for group was not significant, but there was a significant group × period interaction, $F(38,513) = 1.44$, $p < .05$. Trend analyses on two groups at a time revealed a significant group × period interaction in the analysis of the 400-msec versus control data, $F(19,342) = 2.04$, $p < .05$, as well as a significant group × linear trend across periods, $F(1,18) = 5.94$, $p < .05$. However, the 250-msec group did not differ in any of the analyses from the 400-msec group or the control group. Thus, the data suggest that infants reliably discriminated the stimulus change with an SOA of 40 msec but not with an SOA of 250 msec. Perhaps because of the relatively small sample size used in this experiment, there was no significant difference between the 250-msec and 400-msec groups despite a sizable difference in means (see fig. 3).

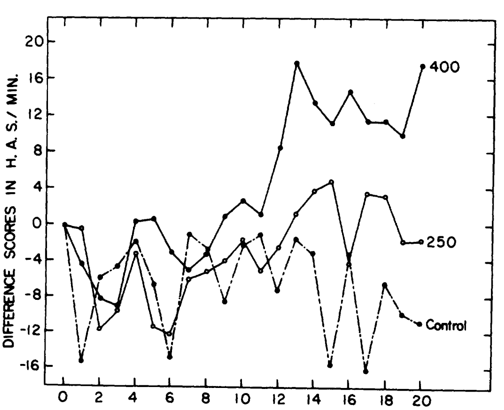

FIG. 3.—Experiment 3: difference scores across 10 min in HAS per minute for the 400-msec, 250-msec, and control groups.

A comparison across experiments 2 and 3 reveals an orderly pattern in infants' backward-masking results: with a 50-msec SOA, performance of the experimental group was most like its control; with a 250-msec SOA it was intermediate; and with a 400-msec SOA, performance of the experimental group was least like its control. The *forward*-masking condition of experiment 2, in which discrimination was robust, may alternatively be viewed as a 950-msec SOA backward-masking condition, because 950 msec was the time from the onset of the target vowel to the onset of the next stimulus. Thus, the pattern of means across experiments 2–3 suggests an increase in discriminative performance when the SOA was extended beyond 250 msec.

The present finding of infants' improved performance with SOAs longer than 250 msec differs importantly from the adult finding of a point of asymptote at about 250 msec (Massaro 1972, 1973). Although a comparison of adults and infants within the present study confirms this developmental difference, however, additional work is needed to ensure that task constraints do not differentially affect adult and infant subjects.

A more theoretical issue is the correct interpretation of adult backward-masking studies. The 250-msec estimate of echoic storage based upon masking studies is in conflict with longer estimates based upon other types of paradigm. One possible resolution of this discrepancy is that some paradigms may actually index a partially analyzed memory rather than the unanalyzed echoic trace (Klatzky 1980). However, there are two alternative accounts of asymptotic performance in backward masking. It could reflect a constraint of storage: the echoic trace may fade by 250 msec. Alternatively, the point of asymptote could reflect a constraint of process: the subject may not be able to extract useful information from an echoic trace after 250 msec, although the trace itself might still be intact at that time. Accordingly, it is not clear whether the present infant data suggest that the echoic trace may last longer in infancy than in adulthood, or that the extraction of information from the trace may last longer in infancy.

A more general conclusion from the present work is that the useful lifetime of an echoic trace (i.e., either the duration of the trace or its participation in the recognition process) may last longer in 8–9 week old infants than in adults. This feature of echoic storage would have the advantage for infants, in comparison with adults, that extra time would be available for the analysis of any one auditory stimulus (e.g., a syllable). However, because an infant's preperceptual storage mechanism would be freed for information intake less frequently, the added processing time might be achieved at the cost of a reduced ability of infants to analyze rapid sequences of stimuli that exceed a single echoic trace (e.g., a multisyllabic string). These properties of infant perception would help to explain the utility of slowed speech addressed to infants by caretakers and siblings (Snow & Ferguson 1977).

References

Eimas, P. D.; Siqueland, E. R.; Jusczyk, P.; & Vigorito, J. Speech perception in infants. *Science*, 1971, **171**, 303–306.

Klatt, D. H. Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, 1980, **67**, 971–995.

Klatzky, R. A. *Human memory* (2d ed.). San Francisco: W. H. Freeman, 1980.

Lasky, R. E., & Spiro, D. The processing of tachistoscopically presented visual stimuli by five-month-old infants. *Child Development*, 1980, **51**, 1292–1294.

Massaro, D. W. Preperceptual images, processing time, and perceptual units in auditory perception. *Psychological Review*, 1972, **79**, 124–145.

Massaro, D. W. A comparison of forward versus backward recognition masking. *Journal of Experimental Psychology*, 1973, **100**, 434–436.

Snow, C. E., & Ferguson, C. A. (Eds.), *Talking to children*. Cambridge: Cambridge University Press, 1977.

Swoboda, P. J.; Kass, J.; Morse, P. A.; & Leavitt, L. A. Memory factors in vowel discrimination of normal and at-risk infants. *Child Development*, 1978, **49**, 332–339.

Swoboda, P. J.; Morse, P. A.; & Leavitt, L. A. Continuous vowel discrimination in normal and at risk infants. *Child Development*, 1976, **47**, 459–465.

Trehub, S. E., & Chang, H. -W. Speech as reinforcing stimulation for infants. *Developmental Psychology*, 1977, **13**, 170–171.