Infant Speech Recognition in Multisyllabic Contexts

Jan V. Goodsitt, Philip A. Morse, and James N. Ver Hoeve
University of Wisconsin—Madison

Nelson Cowan
University of Massachusetts at Amherst

GOODSITT, JAN V.; MORSE, PHILIP A.; VER HOEVE, JAMES N.; and COWAN, NELSON. *Infant Speech Recognition in Multisyllabic Contexts.* CHILD DEVELOPMENT, 1984, **55**, 903–910. In 2 infant speech recognition experiments using trisyllabic sequences, the amount of redundancy within nontarget, context syllables was manipulated. Infants 6½ months old were trained to discriminate the syllables [ba] versus [du] in contexts that were either redundant (e.g., [ko ba ko] or [ti ba ti]) or mixed (e.g., [ko ba ti] or [ti ba ko]). A visually reinforced head-turning procedure was employed. In Experiment 1, context was manipulated between subjects, but in Experiment 2 each subject received all 4 contexts (2 redundant, 2 mixed). Infants consistently recognized the familiar target in all contexts, but did so more successfully in redundant than in mixed trisyllabic contexts. These results suggest that amount of speech redundancy may be an important factor in infants' perceptual capabilities.

Although most of the infant's early exposure to language occurs in the context of multisyllabic speech, studies of infant speech perception in the first year of life have concentrated primarily on isolated, single-syllable contrasts (for recent reviews, see Jusczyk, 1981; Morse & Cowan, 1982). However, as infants develop a receptive understanding of language, they must begin to apply these discriminative skills to multisyllabic speech. Infants must be able not only to discriminate and recognize familiar speech sound patterns but also to segment these patterns within the caretaker's speech stream.

The few infant studies that have explored the discrimination of speech contrasts in contexts other than isolated syllables have raised a number of interesting questions about the stimulus features and organization that facilitate perception of embedded syllables. Jusczyk and Thompson (1978) examined 2-month-old infants' ability to discriminate place of articulation ([ba] vs. [ga]) within bisyllabic strings. They observed that infants could discriminate this contrast in both the initial ([bada] vs. [gada]) and final ([daba] vs. [daga]) syllable positions. However, in a study of voicing discrimination, Trehub (1973) failed to find evidence that infants 4–17 weeks of age discriminate [ba] versus [pa] when presented in the trisyllabic sequence [ataba] versus [atapa] (although they could discriminate [aba] vs. [apa]). Taken together, these studies suggest that structural features (e.g., spectral differences in voicing vs. place contrasts) may affect in-context discrimination.

Thus, although it is likely that infants' discrimination of embedded syllables is modulated by the context syllables, the mechanisms by which context affects performance remain largely unexplored. The present study manipulates two basic features of context strings: the information content of two context syllables and the position of a target syllable within these context syllables. Previous work supports the expectation that redundancy (Fee & Ingram 1982; Ferguson, 1977; Koffka, 1935, pp. 483–486) and

This research was supported by NIMH grant MH35000 and NICHD grant HD08240 to Philip Morse. We are indebted to Kit Hoffmann and Debbie Brown for their help in data collection and analysis. We would also like to thank Cliff Gillman, Dave Wilson, and Gary Weismer for useful comments. Reprint requests should be addressed to Jan V. Goodsitt, Institute of Child Development, University of Minnesota, Minneapolis, Minnesota 55455, or Philip A. Morse, Director of Neuropsychology Service, New England Rehabilitation Hospital, Rehabilitation Way, Woburn, Massachusetts 01801.

[*Child Development*, 1984, **55**, 903–910. 0009-3920/84/5503-0021$01.00]

syllable position (Brown, 1973; Crowder, 1976; Kuczaj, 1979; Slobin, 1973) are important aspects of actual language learning environments.

In the present experiments, infants 6½ months of age were first trained, using a conditioned head-turning procedure, to discriminate a very salient speech contrast, [ba] versus [du]. In subsequent sessions, infants were tested for their recognition of this familiar contrast when it was embedded within two "context" syllables that were either redundant (i.e., identical: [ko-ko] or [ti-ti]) or mixed (i.e., different: [ko-ti] or [ti-ko]). We also examined whether the position of the target syllable within the trisyllabic string (i.e., initial, medial, or final position) would affect recognition.

It was expected that, although infants might be able to discriminate the target syllables reliably in either context, the performance level would be lower in a mixed context than in a redundant context. This expectation was based on the assumption that infants would encounter more information in a mixed context, making it relatively more difficult to recognize a syllable contrast in this condition. Experiment 1 manipulated redundancy in a between-subjects design; Experiment 2 extended these procedures to the more naturalistic and complex situation in which the redundant and mixed trial types of the first experiment were presented within subjects. It was also assumed that, if infants perceive each trisyllabic sequence as a separate unit, they might discriminate the target syllables more effectively when they occurred in the initial or final position rather than the medial position (Crowder, 1976; Cornell & Bergstrom, Note 1).

Experiment 1

Method

Subjects

Infants 6½ months of age, who were born in and near Madison, Wisconsin, were recruited from birth announcements in local newspapers. Of the 64 infants whose parents agreed to participate, 40 were excluded from the study for the following reasons: excessive fussiness/crying (three), equipment/experimenter error (six), failure to meet a predetermined training criterion during the first session in the laboratory (14), and difficulties scheduling the infant on subsequent testing sessions (17). The remaining 24 full-term healthy infants (14 female, 10 male) completed the study in a total of three or four sessions scheduled 5–10 days apart. These subjects ranged in age from 183 to 214 days (mean = 201.1) at their first session and from 203 to 232 days (mean = 217.2) at their final session. Twenty-two subjects required three sessions; the remaining two subjects completed the study in four sessions.

Stimuli

One exemplar of each of the four syllables, [ba], [du], [ko], and [ti], spoken by an adult female was selected as the speech tokens employed in both experiments. The stimuli were matched on the basis of duration (350 msec), peak intensity (54 dB-B), and fundamental frequency contour. Fundamental frequency was analyzed using the algorithm suggested by Henke (Cooper & Sorensen, 1981). The analysis indicated that the context stimuli ([ko] and [ti]) both had fundamental frequencies beginning at 225 hertz and declining to 175 hertz in a smooth, linear fashion. The target stimuli ([ba] and [du]) exhibited a similar linear decline from 200 to 150 hertz.

These stimuli were digitized, and audio tapes containing stimulus sequences were prepared using the VOCAL stimulus patterning program of the Waisman Center Computing Facility at the University of Wisconsin—Madison (Gillman, Wilson, Hirsch, & Morse, Note 2).

Apparatus

Infants' speech discrimination was tested using the apparatus described in Miller, Younger, and Morse (1982). Briefly, stimuli were delivered with a Sony TC-756 stereo tape deck, a Crown D-60 amplifier, and an AR 2AX speaker. Stimulus intensity was manipulated via a Hewlett-Packard 51-C attenuator. Subjects were tested in an Audio-Suttle AS-126 sound-attenuated chamber. Reinforcement was provided by illuminating an otherwise opaque smoked Plexiglas box below the speaker and activating a motorized animal figure inside the box. Trial duration and delivery of reinforcement were controlled by a custom-designed microprocessor.

Procedure

Using the conditioned head-turning paradigm of Eilers, Wilson, and Moore (1977), infants were tested in two phases: a training phase and a recognition phase. Training occurred within the first session, whereas recognition testing was conducted in the two or three subsequent sessions. During all sessions the infant was seated on the parent's lap. The continuously repeating background stimulus was presented at 54 dB (B) SPL (20 microbars as measured with a Bruel & Kjaer Type 2203 SPL meter). An assistant, seated across from the infant, maintained the infant's attention with an assortment of toys on a small table directly in front of the infant. Trials were initiated only when the infant was judged to be actively engaged in viewing the toys presented by the assistant. This judgment was made by the experimenter (outside of the test chamber), who monitored the infant over a closed circuit video system. Reinforcement was delivered contingent on the infant responding to a stimulus change with a head turn toward the loudspeaker. Throughout all sessions, both the assistant and the infant's parent listened to music over headphones that masked the types of stimuli presented to the infant.

Training.—The training stimulus tape contained repetitions of [ba] on one channel, separated by 1-sec intervals and synchronized with a sequence of [du] syllables on the second channel. The duration of a trial was timed by the microprocessor to contain five syllables and lasted approximately 5.75 sec. Half of the infants who successfully completed the study received [ba] as the repeating, background stimulus (T−) and [du] as the target stimulus (T+), whereas the remaining infants received [ba] as T+ and [du] as T−. During the first portion of the training session, each infant was shaped to turn his or her head on trials that contained a change from the background (T−) channel to the target (T+) channel. Initially, the T+ stimuli were presented at a level 12 dB greater than the T− stimuli. After two consecutive head turns to T+ at this and subsequent levels, the intensity difference between T+ and T− was systematically decreased in 4 dB steps. Failure to meet this criterion of two consecutive head turns at a given intensity level resulted in increasing the intensity differences in 4 dB steps up to 12 dB and/or the use of the reinforcer as an additional cue. When the infant successfully responded at +4 dB, the two tape channels were subsequently equated at 54 dB, and a second part of the training phase began, during which the infant was trained to a criterion level of performance.

In criterion testing, the experimenter initiated change (T+) and no-change control (T−) trials. The microprocessor arranged these trials in a random sequence with approximately half T+ and half T− trials. In the test chamber the assistant (who was continually monitoring the infant's head turning) depressed a small, concealed vote button whenever the infant produced a head turn toward the speaker. The experimenter, monitoring the infant over the video system, also depressed a vote button whenever a head turn occurred on a trial. Reinforcement for a head turn on T+ trials required the agreement of the assistant and experimenter. The criterion for successful completion of training consisted of nine out of 10 consecutively correct responses on both change and control trials. Since infants would require several subsequent recognition test sessions, only infants who met this criterion on their first visit to the laboratory were included in the study. Furthermore, based on pilot work, infants were not included if they failed to produce two consecutive head turns after 30 shaping trials at +12 dB, or if they displayed an obvious and persistent state change (e.g., crying).

Recognition testing.—The four recognition test sequences each contained the syllable [ba] on one channel and [du] opposite it on the second channel, in either the initial, medial, or final syllable position of a three-syllable string. Each trisyllabic string contained 50 msec of silence between syllables and a 2-sec interval separated consecutive strings. For one of the two recognition test sequences of the redundant condition, [ba] and [du] were embedded among two identical tokens of [ko], whereas in the other they were embedded among two tokens of [ti]. For example, the latter sequence would contain the trisyllables [titiba], [tibati], and [batiti] on one channel and [titidu], [tiduti], and [dutiti] on the opposite channel (see Table 1). In contrast, the two

TABLE 1

TRISYLLABIC SEQUENCES PRESENTED TO SUBJECTS

Redundant Conditions	
[*x* ko ko]	[*x* ti ti]
[ko *x* ko]	[ti *x* ti]
[ko ko *x*]	[ti ti *x*]
Mixed Conditions	
[*x* ko ti]	[*x* ti ko]
[ko *x* ti]	[ti *x* ko]
[ko ti *x*]	[ti ko *x*]

NOTE.—*x* = [ba] or [du], depending on condition.

recognition test sequences of the mixed condition contained two different context syllables in which [ba] and [du] were embedded. One sequence contained the order [ko] and then [ti], whereas the context of the other consisted of the order [ti] and then [ko]. Thus, as Table 1 illustrates, the latter sequence contained the following trisyllables on one channel: [tikoba], [tibako], and [batiko]. The recorded trisyllables in all sequences were randomized with three restrictions. First, the two channels of the tape were identical except for the [ba] versus [du] contrast. Second, three identical trisyllables formed a block within the sequence. This was done to provide the infant a sufficient trial window (9.6 sec) within which to execute a head turn. And third, the random sequence of trisyllabic blocks was limited to a maximum of three consecutive identical blocks. Examples of possible trial sequences in Experiment 1 are shown in the top half of Table 2.

The first recognition test session was scheduled approximately 1 week after training. Each recognition session began with a brief review of the familiar training contrast (two trials at each of the three shaping intensity levels). Immediately following the review, recognition testing commenced. Half of the subjects were tested for the recognition of their training contrast in a redundant context; the remaining infants were tested with their training contrast embedded in a mixed context. Furthermore, the infants in each context condition were evenly divided (six per group) between the two sequences for that condition. Change (T+) and control (T−) trials were initiated by the experimenter only at the boundary between blocks of trisyllabic strings. Recognition testing was continued using the randomized tape sequences until the infant had received a total of 10 change and 10 control trials for each of the three target syllable positions (i.e., a minimum of 60 trials). The average total number of recognition trials across sessions was 78.1.

Results and Discussion

Training

Between group differences during the training session were evaluated with ANOVAs containing factors for context structure (redundant vs. mixed) and target (T+) syllable ([ba] vs. [du]). No reliable between-subjects differences, $p > .2$, were observed in the analyses of the number of trials to completion of the shaping phase, the number of trials to criterion, or the percentage correct during the second part of the training phase.

Recognition Test Performance

Percentage correct head turning was tabulated for the total number of test (change and control) trials presented to each infant for each of the three target positions (shown in Table 3). Separate t tests on the percentage correct of the two context structure and two target syllable groups indicated that each group reliably, $p < .01$, exceeded chance (50%) levels of performance. These percentage correct scores were also subjected to a mixed ANOVA with context structure (redundant vs. mixed) and target syllable ([ba] vs. [du]) as between-subjects factors and position (initial, medial, or final)

TABLE 2

Examples of Possible Tape Sequences in Experiments 1 and 2

Experiment and Condition	Tape Segment $N-1$	Trial N	Tape Segment $N+1$	Eligible Subjects
Experiment 1:				
Redundant control	[titiba]	[tibati]	[batiti]	Redundant group ([titi])
Redundant experimental	[titiba]	[tiduti]	[batiti]	Redundant group ([titi])
Mixed control	[tikoba]	[tibako]	[batiko]	Mixed group ([tiko])
Mixed experimental	[tikoba]	[tiduko]	[batiko]	Mixed group ([tiko])
Experiment 2:				
Redundant control	[tikoba]	[kobako]	[kobati]	All subjects
Redundant experimental	[tikoba]	[koduko]	[kobati]	All subjects
Mixed control	[tikoba]	[kobati]	[kobati]	All subjects
Mixed experimental	[tikoba]	[koduti]	[kobati]	All subjects

Note.—Each trisyllabic sequence was presented three times in succession. This table illustrates only a few of several possible trial types.

as a within-subject factor. The two tape conditions within each level of context structure were combined in this analysis. The results of the analysis revealed significant effects for context structure, $F(1,20) = 11.73$, $p < .005$, and for target syllable, $F(1,20) = 10.21$, $p < .005$, with no other reliable effects. Infants in the redundant context correctly recognized the training contrast on 75.2% of the trials; the mean performance for subjects in the mixed condition was only 67.1%. The target syllable differences reflected the relative ease experienced by infants in recognizing [ba] as the T+ stimulus (75.2%) compared with [du] (67.2%).

In sum, the results of this experiment were consistent with the expectation that the infant's recognition of a familiar syllable would be facilitated by uniformity or reduced complexity of the background context in which it occurs. Performance in the redundant condition was superior to that observed in the mixed condition. A reliable difference in recognition testing was also observed for the target syllable [ba] versus [du]. This stimulus difference did not occur during the training session, nor did it interact with context structure during recognition testing.

Experiment 2

A second study was undertaken to replicate the redundant versus mixed trisyllabic context effect with both context types presented to each infant. This within-subject sequencing provided a potentially more demanding and more naturalistic situation for the infant, reproducing some of the complexity of speech in the infant's environment. In addition to presenting redundant versus mixed trial types, this design provided a "mixed" sequence of stimuli for each subject in that subjects heard the contrast embedded in not just a single context (i.e., any quadrant of Table 1), but in a mixture of all four of the contexts of Experiment 1 (i.e., all four quadrants of Table 1).

TABLE 3

Percentage Correct for Experiment 1

Context	Position: Initial	Medial	Final	Mean
Redundant	78.2	75.1	72.3	75.2
	(13.5)	(14.9)	(16.4)	(15.0)
Mixed	66.5	64.8	69.9	67.1
	(12.2)	(14.4)	(13.1)	(13.2)
Mean	72.4	70.0	71.1	
	(14.0)	(15.4)	(14.7)	

Note.—The figures in parentheses are standard deviations.

Method

Subjects

Thirty-seven infants were tested to obtain the final sample of 16 subjects (eight males, eight females). These infants ranged in age from 183 to 195 days (mean = 189.2 days) at their first session and from 196 to 216 days (mean = 204.2 days) on completion of the final test session. Of these 16 infants, 10 required three sessions, five needed four sessions, and one required a fifth session to receive the minimum number of trials to complete testing. The remaining 21 infants were not included in the results because of equipment problems or experimenter error (three), persistent crying (three), problems scheduling the multiple recognition test sessions (five), and failure to meet the predetermined training criterion in the first session (10).

Stimuli, Apparatus, and Procedure

The stimuli, apparatus, and procedures followed during the training session were identical to those employed in Experiment 1. Furthermore, the general features of the recognition test sessions were similar to those of the first experiment. As in Experiment 1, the target (T+) syllable ([ba] or [du]) was balanced across subjects and the target position (initial, medial, or final) varied within subjects. However, in this experiment context structure (redundant vs. mixed) also varied within subjects. Two recognition tapes with different randomizations were constructed such that they included all of the trial types depicted in Table 1. In addition, the randomization of blocks of three trisyllabic strings was restricted to provide more variation across trials. Whereas, in Experiment 1, subjects could hear as many as nine consecutive identical trisyllabic strings (three blocks), in this experiment the same context or the same target position was permitted to occur in a maximum of two consecutive blocks, and no instances of six consecutive trisyllables containing the same context and target position were allowed. Examples of possible trial sequences in Experiment 2 are shown in the bottom half of Table 2. In Experiment 1 each infant received a minimum of 60 trials (10 trials for each of six trial types) with the target contrast embedded in a particular context (e.g., [ti-ti]). In contrast, in Experiment 2 infants were presented with the target stimuli on change (T+) trials and control (T−) trials for all four of the contexts of Table 1. Infants received a minimum of 20 trials per context type, or a minimum of 80 trials in all. Within each set of 20 trials, the infant received an approximately equal number of trials per target position. (A minimum of 240 trials would have been required of each infant if the subject had received 20 trials per position for each of the four contexts.) Infants in Experiment 2 required an average of 105.4 trials to meet the 20 trial per context type criterion.

Results

Training

As in Experiment 1, analyses were carried out on the number of trials to completion of the shaping phase and on both the number of trials to criterion and percent correct during the second portion of the training session. The performance of the two target syllable groups ([ba] or [du]) revealed no evidence of any group differences during the training session, *t* tests, p's $> .2$.

Recognition Test Performance

As found in the first experiment, *t* tests indicated that the performance of both target groups significantly exceeded chance levels, p's $< .01$. Subjects' percentage correct scores for total trials were also subjected to an ANOVA with context type (redundant vs. mixed) and position of the target (initial, medial, or final) as within-subjects factors and target type ([ba] vs. [du]) as a between-subjects factor. As in Experiment 1, the data for the two redundant conditions were combined and those for the two mixed contexts were combined. The pattern of results, shown in Table 4, replicated those obtained in Experiment 1. Overall performance for target syllables presented in redundant contexts (71.1%) was superior to mixed contexts (66.2%), $F(1,14) = 7.38$, $p < .05$. As in Experiment 1, recognition performance when the target syllable was [ba] (74.3%) was superior to recognition of [du] (62.9%) as the target, $F(1,14) = 13.48$, $p < .05$. Finally, no reliable main effect of target position nor any significant interaction obtained. Thus, the pattern of results was essentially unchanged when redundant and mixed trial types were presented to each subject in Experiment 2, rather than when presented as a between-subjects variable as in Experiment 1.

TABLE 4

Percentage Correct for Experiment 2

	Position			
Context	Initial	Medial	Final	Mean
Redundant	68.1	72.3	72.8	71.1
	(26.8)	(11.9)	(15.1)	(19.0)
Mixed	65.9	66.6	66.0	66.2
	(15.8)	(14.4)	(17.0)	(15.6)
Mean	67.0	69.5	69.4	
	(21.9)	(13.4)	(16.3)	

Note.—Figures in parentheses are standard deviations.

General Discussion

This research demonstrates that 6½-month-old infants are able to discriminate consonant-vowel syllables in varying positions within trisyllabic sequences. However, the data also indicate that the level of discrimination performance is affected by characteristics of the context syllables. Specifically, in both experiments performance was reliably better when the two context syllables within a trisyllable were identical (redundant condition) rather than different (mixed condition).

Contrary to the hypothesis that performance would be weakest when the target syllable was in the medial position within a trisyllable, no main effect or interaction with position approached significance in either experiment. There are several possible explanations for this result. First, to provide the infant with a sufficient opportunity to respond, each trial consisted of three repetitions of a trisyllable. Infants could have perceived a trial as a single, nine-syllable series rather than three trisyllables, obscuring the relative distinctiveness of position within each trisyllabic string. Alternatively, position effects may have failed to obtain because the test permitted immediate recognition and responding and thus may not have stressed the short-term memory abilities of the infant.

In both experiments, infant's performance was better when the T+ signaling a head turn was [ba] as opposed to [du]. This result was unexpected, although discrimination of a speech sound pair in only one direction has been obtained in other infant speech perception studies (Miller, Morse, & Dorman, 1977; Swoboda, Kass, Morse, & Leavitt, 1978). One implication of this target effect is that it permits a comparison of the redundant and mixed conditions at two levels of target difficulty. Discrimination was found to be superior in the redundant condition as compared with the mixed condition across both the [ba] and the [du] targets, in the absence of a reliable context × target syllable interaction. Therefore, it is likely that the observed advantage for the redundant condition is not restricted to a single level of target difficulty but instead reflects a more general principle in infants' auditory/speech processing.

The most striking finding of this study was the relative advantage of redundancy or reduced information in the infant's processing of speech. Infants in both experiments evidenced better discrimination of the target when the background context syllables were redundant as opposed to mixed. Furthermore, when infants were presented with redundant and mixed trials randomized together in Experiment 2, the advantage for the redundant syllabic context observed in Experiment 1 was attenuated, although not eliminated (redundant condition in Experiment 1 = 75.2%, in Experiment 2 = 71.1%). However, in a comparison between experiments this difference was not significant. Performance levels in the mixed conditions were nearly identical in the two experiments (67.1% vs. 66.2%).

The main effects of context type and the greater difference between context types in Experiment 1 may result from requirements of the task. If infants monitor the trisyllabic strings for occurrences of the target, then this monitoring may be affected by the number of possible syllables that could occur. The greater the number of different context syllables, the more likely the infant's attention may be divided between the context and the target stimuli, thereby decreasing the likelihood of target recognition. In the redundant condition of Experiment 1, subjects were presented with the target in the context of only one other syllable (either [ko] or [ti]). In contrast, in the mixed condition of Experiment 1, infants heard the target syllable in the context of two other syllables. Furthermore, the within-subject design of Experiment 2 presented all infants with the task of monitoring for occurrences of the target in a sequence that contained one or two other syllables, depending on the trial type. Therefore, Experiment 1 provided some subjects with redundancy within and across trials, whereas Experiment 2 provided within-trial redundancy only.

In summary, this study has demonstrated that infants' recognition of a target speech sound is facilitated by the redundancy of the context in which it occurs. These findings are consistent with research in other aspects of early language development that describes redundancy in the caregiver's reduplicative phrasing of words to young children (e.g., "kiki" for kitty, "mama") and their routine use of a small number of familiar sentence frames (Ferguson, 1977).

Reference Notes

1. Cornell, E., & Bergstrom, L. *Serial position effects in infants' recognition memory.* Paper presented at the International Conference on Infant Studies, Austin, Texas, 1982.
2. Gillman, C., Wilson, D., Hirsch, J., & Morse, P. *VOCAL: A computer-based system for preparing auditory stimuli* (Research Status Report No. 1). University of Wisconsin—Madison, Infant Development Laboratory, 1975.

References

Brown, R. *A first language.* Cambridge, Mass.: Harvard University Press, 1973.

Cooper, W. E., & Sorenson, J. M. *Fundamental frequency in sentence production.* New York: Springer-Verlag, 1981.

Crowder, R. G. *Principles of learning and memory.* Hillsdale, N.J.: Erlbaum, 1976.

Eilers, R., Wilson, W., & Moore, J. Developmental changes in speech discrimination in three-, six-, and twelve-month-old infants. *Journal of Speech and Hearing Research,* 1977, **20,** 766–780.

Fee, J., & Ingram, D. Reduplication as a strategy of phonological development. *Journal of Child Language,* 1982, **9,** 41–54.

Ferguson, C. Baby talk as a simplified register. In C. Snow & C. Ferguson (Eds.), *Talking to children.* New York: Cambridge University Press, 1977.

Jusczyk, P. Infant speech perception: A critical appraisal. In P. Eimas & J. Miller (Eds.), *Perspectives on the study of speech.* Hillsdale, N.J.: Erlbaum, 1981.

Jusczyk, P., & Thompson, E. Perception of a phonetic contrast in multisyllablic utterances by 2-month-old infants. *Perception and Psychophysics,* 1978, **23,** 105–109.

Koffka, K. *Principles of Gestalt psychology.* New York: Harcourt, Brace, & World, 1935.

Kuczaj, S. A., II. Evidence for a language learning strategy: On the relative ease of acquisition of prefixes and suffixes. *Child Development,* 1979, **50,** 1–13.

Miller, C., Morse, P., & Dorman, M. Cardiac indices of infant speech perception: Orienting and burst discrimination. *Quarterly Journal of Experimental Psychology*, 1977, **29**, 533–545.

Miller, C., Younger, B., & Morse, P. The categorization of male and female voices in infancy. *Infant Behavior and Development*, 1982, **5**, 143–159.

Morse, P., & Cowan, N. Infant auditory and speech perception. In T. M. Field, A. Huston, H. C. Quay, L. Troll, & G. E. Finley (Eds.), *Review of human development*. New York: Wiley, 1982.

Slobin, D. Cognitive prerequisites for the development of grammar. In C. Ferguson & D. Slobin (Eds.), *Studies of child language development*. New York: Holt, Rinehart & Winston, 1973.

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

Trehub, S. *Auditory linguistic sensitivity in infants*. Unpublished doctoral dissertation, McGill University, 1973.