JOURNAL OF EXPERIMENTAL CHILD PSYCHOLOGY 69, 133–149 (1998)
ARTICLE NO. CH982438

Intersensory Redundancy Facilitates Learning of Arbitrary Relations between Vowel Sounds and Objects in Seven-Month-Old Infants

Lakshmi J. Gogate

University of Delaware

and

Lorraine E. Bahrick

Florida International University

This study investigated 7-month-old infants' ability to relate vowel sounds with objects when intersensory redundancy was present versus absent. Infants ($N = 48$) were habituated to two alternating video-films of vowel-object pairs in one of three conditions. In the moving-synchronous condition, where redundancy was present, the movement of one object was temporally coordinated with the spoken vowel /a/ and that of the other with /i/, simulating showing and naming the objects to the infant. In the still and in the moving-asynchronous conditions, where redundancy was absent, infants saw static objects, and objects moving out of synchrony with the vowel sounds, respectively. The results indicated that infants detected a mismatch in the vowel-object pairs in the moving-synchronous condition but not in the still or the moving-asynchronous condition. These findings demonstrate that temporal synchrony between vocalizations and the motions of an object facilitates learning of arbitrary speech-object relations, an important precursor to the development of lexical comprehension in infancy.

Adults refer to objects and actions by labeling them with conventionally used words. Words (concrete nouns and verbs) and their referents are arbitrarily

This study was conducted by the first author during a postdoctoral traineeship at Florida International University. Portions of these data were presented at the International Conference on Infant Studies, Providence, RI, April, 1996, and at the Biennial Meeting of the Society for Research in Child Development, Washington, DC, April, 1997. This research was supported by an NICHHD Grant RO1 HD25669 to the second author. Special thanks are extended to Ivonne Ruiz, Sangeeta Mallik, Natasha Del Pino, Jilayne Watson, and Christina Esquijarosa for their help with data collection, and to Arlene Walker-Andrews, Robert Lickliter, Roberta M. Golinkoff, and the anonymous reviewers for their helpful comments and suggestions on earlier versions.

Address correspondence and reprint requests to Lakshmi J. Gogate, Department of Educational Studies, University of Delaware, Newark, DE 19716. E-mail:gogatel@udel.edu.

0022-0965/98 $25.00

related; for instance, if convention permitted, a 'ball' could just as easily be called a *bat*. Lexical learning entails the detection of arbitrary but conventional relations between labels and specific objects and events in the environment. The ability to detect arbitrary label-object (event) relations is, therefore, an important precursor to lexical comprehension. How do infants detect these relations? What are the conditions that facilitate learning of the relations between a spoken word and its referent? The present study addressed these questions by presenting 7-month-olds with simple verbal labels (the vowels, /a/ and /i/) paired with objects under different conditions.

Word recognition precedes word comprehension and production in infants. Young infants recognize the words they hear, particularly the ones they hear frequently. Thus, by 4½ months, infants are familiar with their own name (Mandel, Jusczyk, & Pisoni, 1995), and by 7 months, following a brief exposure to words in continuous text they attend longer to the familiar words than to novel words (Jusczyk & Aslin, 1995). Similarly, word comprehension precedes word production (Benedict, 1979; Huttenlocher, 1974; Reznick, 1990). However, the developmental course leading to lexical comprehension has received little attention in past research (Bates, 1993). Studies have yet to show whether infants, prior to the first year, know that a word may be paired or matched with an object (referent) and generally refers to the object.

What are the requisite abilities that culminate in infants' matching of spoken words and referents? The cognitive view of lexical development proposes that prior to acquiring first words, infants build up representations for objects and events. They detect the meaning of a word by simply matching the words they hear from adults to their active internal representation of these objects or events (Ellis & Wells, 1980; Stemmer, 1989). Exactly how infants match words with objects, however, has not been investigated. We suggest, in accordance with several researchers (Harris, 1992; Sullivan & Horowitz, 1983; Zukow-Goldring, 1997), that the infant's perceptual development plays a central role in emerging lexical comprehension. Specifically, we hypothesized that redundant or amodal information, uniting the spoken labels with visually available objects, might facilitate the detection of arbitrary relations between them (c.f., Bahrick, 1994; Gogate, Bahrick & Watson, submitted; Sullivan & Horowitz, 1983). At first, the infant may benefit from temporal synchrony between the motion of an object and the presentation of a verbal label to link them together. Thus, the detection of intermodal information could assist lexical learning.

The area of intermodal perception offers a theoretical framework as well as the methods to study the precursors to lexical comprehension. Research in this area has revealed that very young infants are able to abstract amodal relations uniting audible and visible stimulation (Bahrick & Pickens, 1994; see Gibson, 1969 for a review). Amodal information is information that is completely redundant across two or more senses. Young infants detect redundant information across auditory and visual events involving speech (Kuhl & Meltzoff, 1988; Walton & Bower,

1993) and nonspeech (Bahrick, 1983; 1987; 1988; 1992; Lewkowicz, 1992). For instance, infants detect the correspondence between the shape of the mouth they see and the vowel sound they hear across several vowel pairs by 4-months (Kuhl & Meltzoff, 1988; 1984; 1982; Legerstee, 1990). Furthermore, infants look more at a gradually opening mouth than at a closing mouth when the amplitude during vocalization increases (MacKain, Studdert-Kennedy, Spieker, & Stern, 1983). These findings in the speech domain are consistent with the finding that infants detect redundant intensity shifts across bimodal nonspeech events. For example, they match approaching and receding objects with an increase and a decrease in sound amplitude, respectively (Pickens, 1994; Walker-Andrews, & Lennon, 1985). Thus, infants are accomplished perceivers of amodal information, or redundancy across the senses.

Lexical comprehension involves infants' intermodal learning of a different kind; that which is based on arbitrary relations between speech sounds and visual objects. Arbitrary intermodal relations are available when information across the senses is not redundant, and may not predictably occur together in nature (Bahrick, 1994; Walker-Andrews, Bahrick, Raglioni, & Diaz, 1991; Walker-Andrews, 1994). For example, the relationship between words and their referents is entirely arbitrary. So is the relation between the color of an object and its sound at impact, or the appearance of a container and the taste, temperature or smell of its contents. Research has shown that preverbal infants also detect arbitrary relations across different modalities. For example, they detected the relations between the color/shape of an object and the pitch of its impact sound (Bahrick, 1994), the odor of an object and its visual appearance (Fernandez & Bahrick, 1994), the color of a substance and its taste (Reardon & Bushnell, 1988) and the color/pattern of an object and its shape (Hernandez-Reif & Bahrick, submitted).

The development of infants' sensitivity to amodal relations offers an explanation for how young infants may integrate the auditory and visual information necessary for detecting arbitrary speech-object relations. Bahrick (1992, 1994, see also Bahrick & Pickens, 1994) proposed that redundancy across the senses captures infants' attention, and directs and constrains intermodal learning in the first months of life. Because of this, amodal relations are detected developmentally prior to arbitrary intermodal relations. For example, 3-month-old infants were able to detect the amodal synchrony and information specifying composition of objects hitting a surface (Bahrick, 1992). However, they were unable to detect the arbitrary relation between the pitch of the impact sounds and the color/shape of the objects until 7 months of age (Bahrick, 1994). Similarly, Hernandez-Reif and Bahrick (submitted) found that infants were able to detect amodal shape information across vision and touch by 4 months of age. However, they were unable to detect the arbitrary relation between shape and color/pattern until 6 months of age. Thus, there appears to be a developmental lag between the detection of amodal and arbitrary relations in a given domain. Moreover, research has also shown that amodal relations can serve as the basis for learning arbitrary relations. For example, Hernandez-Reif and Bahrick (submitted) found that the 6-month-olds were able to relate an arbitrary color/pattern with an object only in the presence of amodal information. That is, infants required intermodal redundancy in the sense of seeing and feeling an object of a common shape and size in order to pair the object's color/pattern with its tactually experienced shape. Infants were unable to pair the color/pattern displayed on a poster board with the object they were feeling. Consistent with these findings, we propose that detection of amodal relations such as temporal synchrony, intensity shifts, and common rhythm uniting spoken labels with visual objects provides an important basis for initially linking words and their referents. Moreover, infant-directed communication provides abundant exemplars of these amodal relations termed "multimodal motherese" (Gogate et al., submitted).

Other research is also consistent with the view that detection of amodal relations serves as a basis for learning arbitrary relations. For example, several studies have shown that infants detected the arbitrary relation between the visible object and the type of sound it produced when amodal information such as temporal synchrony, rhythm and tempo were present (e.g., Bahrick, 1994; Lawson, 1980; Lyons-Ruth, 1977; Spelke, 1979; 1981). However, in these studies the importance of amodal information for the detection of arbitrary intermodal relations was not fully assessed. Specifically, controlled experiments have not tested infants' detection of arbitrary intermodal relations in the presence and in the absence of amodal information such as temporal synchrony. Such experiments using speech stimuli with objects are particularly important in the study of lexical comprehension because prior research suggests that amodal information conveyed during concurrent naming and showing objects or actions might facilitate lexical comprehension in infants (Messer, 1978; Sullivan & Horowitz, 1983; Zukow, 1990; 1991; 1997). Thus, we expected that infants would detect arbitrary relations between speech stimuli and objects in the presence but not in the absence of temporal synchrony, if intermodal perception plays an important role in the development of lexical comprehension.

Descriptive studies of early language development have reported that infants relate words with objects as early as 8- to 9-months. For example, 9-month-old infants use idiosyncratic expressions that have a consistent meaning (Blake & Fink, 1987; Halliday, 1975b). Studies of 8- to 9-month-olds from English and Spanish environments have suggested that infants demonstrate comprehension of words by visually orienting to an object when it is labeled (Jackson-Maldanado, Thal, Markman, Bates et al., 1993). A few experimental studies provide indirect evidence that preverbal infants relate words with objects prior to the onset of word production. One showed that 10-month-old preverbal infants looked longer at objects when adults had previously labeled them than when they had merely pointed to them (Baldwin & Markman, 1989). Another study showed that labeling highlights category membership for 9-month-olds (Waxman & Balaban, 1996). Although we may infer from these studies that preverbal infants can pair a word with an object, further research is needed to determine when and under what conditions this ability develops.

Experimental research has focused mainly on older infants' capacity to attach meaning to labels for concrete objects around the time of first word onset. At 13-months, infants learned verbal labels for objects when the target object was labeled several times while the infant attended to it (Woodward, Markman, & Fitzsimmons, 1994). Two prior studies showed that infants from 12- to 17-months matched multisyllabic real and nonsense words with objects following training for the label-object relations (Molfese, Morse, & Peters, 1990; Oviatt, 1980). Furthermore, 16-month-old infants showed evidence of word recognition. They related words (nouns) and visual events by looking longer at one of two dynamic displays that matched with the agent- and object-nouns of a sentence they had heard (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987). Lloyd, Cohen, Werker, Foster et al. (1994) found, following habituation to two word-object pairings, that 14-month-olds could relate syllables to moving but not still objects. Although these studies have enhanced our knowledge about infants' word recognition and beginning lexical learning during the second year, studies of younger infants are required to show how infants learn to detect the arbitrary relations between speech and objects.

In the speech domain, to date, no experiments have shown that preverbal infants learn arbitrary relations between verbal labels and objects, even though they detect arbitrary auditory-visual relations in nonspeech events (Spelke, 1979; Bahrick, 1994). Further, although the descriptive data suggest emerging word-object matching by 8-months (Jackson-Maldanado et al., 1993), the experimental data do not show word-object matching until 13 months of age (Woodward et al., 1994). Might infants also detect the arbitrary relations between speech-object pairings at 7-months of age, similar to their detection of arbitrary relations in audio-visual nonspeech events? Further, might intermodal redundancy facilitate learning of these relations, similar to the facilitation found in the visual-tactile domain (Hernandez-Reif & Bahrick, submitted)?

In the present study, intermodal redundancy was made available in a moving-synchronous condition, when speech sounds were presented in synchrony with moving objects. Intermodal redundancy was eliminated in moving-asynchronous and still conditions when the speech-sounds were presented out of synchrony with moving objects or with still objects. Infants were expected to learn to relate speech-sounds and objects in the moving-synchronous condition because intermodal redundancy would make the speech sound-object relations salient, whereas learning would be attenuated in the other two conditions because intermodal redundancy was not available.

The specific goal of this study was to investigate under what conditions infants detect the arbitrary relations between vowel-sounds and objects, following habituation training to two vowel-object pairs. The vowels /a/ and /i/ were selected as object-labels because young infants may perceive and represent them better than more complex consonant-vowel syllables (Bertoncini, Bijeljac-Babic, Jusczyk, Kennedy, & Mehler, 1988; Kuhl, 1992). Vowels also have the unique ability to constitute a basic syllabic unit in isolation and can be uttered with intonation contours typical of monosyllabic (English) words. Because 7-month-olds learn arbitrary relations between visual and acoustic stimulation in nonspeech events, it was predicted that infants at this age would learn the arbitrary relations in the domain of speech as well. Based on previous findings suggesting that amodal relations or intermodal redundancy provides the basis for learning arbitrary relations (Bahrick, 1994; Hernandez-Reif & Bahrick, submitted), it was predicted that synchrony between the timing of a vowel sound and the motion of an object would provide a basis for learning the arbitrary vowel-object relations. Thus, infants should learn to match speech sounds and objects when they are presented in synchrony, but learning should be attenuated when they occur together but lack a common temporal structure. To test this hypothesis infants' learning of vowel-object pairings was evaluated under three conditions. In the moving-synchronous condition, the objects were moved in synchrony with the vowel sounds. In the still condition, still objects were presented with the vowel sounds. Furthermore, it was hypothesized that it was not the movement *per se,* but the temporal coordination between the verbal label and the object's motion that would facilitate learning the vowel-object relations. Thus, a moving-asynchronous condition was included where the objects were moved out of synchrony with the vowel sounds, providing movement but no synchrony. We predicted that if synchrony played a crucial role in early learning, then learning would be facilitated in the synchronous condition and attenuated in the still and asynchronous conditions.

The infant-controlled habituation procedure was used to assess detection of arbitrary speech-object relations in two events. During the habituation/training phase, infants were presented with two alternating video-films of an object paired with /a/ and another object paired with /i/. Following habituation, infants viewed two change test trials where the vowel-object pairings were switched. Visual recovery to the change in vowel-object pairings was taken as evidence that they detected the vowel-object relations. This "switch design" has been used previously to assess young infants' detection of arbitrary and amodal relations in nonspeech events (Bahrick, 1992; 1994). It has also been used by Cohen, Werker and their colleagues to test infants' learning of syllable-object relations at 14 months (Lloyd et al., 1994), and at 8 months (Werker & Stager, in press). In these studies, however, a fixed-length trial and a fixed number of trials were used during habituation. In the present study, the trials were infant-controlled as in Bahrick's previous studies (1992; 1994), where infants successfully detected arbitrary auditory-visual relations in nonspeech events at 7 months but not at 3 or 5 months.

METHOD

Subjects. Forty-eight 7-month-olds (M = 220 days, SD = 5.14), 16 each in the moving-synchronous, the still, and the moving-asynchronous conditions, participated. Full-term, healthy infants with no complications at birth were recruited through the local birth records. Parents of infants were middle class and had at least 12-years of education. Infants from diverse ethnic groups were included in this experiment. Sixteen additional infants did not complete the procedure due to experimenter error (N = 1), equipment failure (N = 1), excessive fussiness (N = 7), fatigue (N = 1), external interference (N = 2), failure to habituate to the stimuli after 20 trials (N = 3), and for health reasons (N = 1).

Stimulus materials. Video-displays were made of the vowel sounds /a/ and /i/ each with four different colorful objects; a wooden star and a plastic lambchop, a plastic crab and a porcupine. Each object was of two contrasting colors. In addition to these displays, a display of a green and white plastic turtle, whose front legs spun around producing a whirring noise, was used to test for fatigue during the habituation phase (Bahrick, 1992; 1994).

A female native speaker of Standard American English vocalized the vowels in an erratic temporal pattern at a rate of 26 vocalizations per minute. The tokens of each vowel displayed large variations in intonation contour, intensity and pitch. However, the speaker attempted to keep the overall intensity and pitch changes of the vowel sounds constant across the two vowel types. No significant difference was found (two-sample $t(54) = 1.63$, $p > .1$) between the mean fundamental frequencies (across the entire waveform) of 28 random samples of /a/ (M = 1124.5 Hz; SD = 66.3) and that of 28 random samples of /i/ (M = 1035 Hz; SD = 284). Similarly, no significant difference was found (two-sample $t(54) = .93$, $p > .1$) between the mean loudness (across the entire waveform) of 28 random samples of /a/ (M = 13.61 dB; SD = .87) and that of 28 random samples of /i/ (M = 14.11 dB; SD = 2.71).[1]

For the moving-synchronous condition, each object was depicted moving in synchrony with the vowels /a/ and /i/. The actor's hands were visible holding and moving the object, forward and then back or laterally, in synchrony with a vowel sound to resemble an act of showing and naming the object to the infant. The actions and synchronous vocalizations occurred in an unpredictable temporal pattern. A vowel sound was spoken during each object motion. The mean time lag between vocalization and object motion across 20 randomly selected occurrences was .45 s (SD = .19) for the onset and .04 s (SD = .11) for the offset. For the still condition, the actor's hand was not visible and the vowel sounds were spoken while each object was static. An audio segment of /a/ and another of /i/ from the moving-synchronous videos was dubbed onto each still object display.

[1] For the purpose of acoustic analyses the vowels were recorded onto a cassette tape recorder placed a foot away from and in front of the speaker. This recorded amplitude differed from that at which the infants had heard the vowel sounds, when seated on a table under which the speaker was centrally located.

Dubbing of vocalizations from the moving-synchronous to the still videos controlled for the frequency, rhythm, tempo, intensity shifts, and duration of vocalizations across conditions. For the moving-asynchronous condition, each object was depicted hand-held and moving, but out of synchrony with the vowels /a/ and /i/. The vowel sounds were recorded onto the videos of the asynchronous condition to occur during the pauses between object motions. Thus, the vocalizations were temporally asynchronous with the motions of the objects. The mean time lag between 20 randomly selected vocalizations and object motion was .82 s (SD = .68) for the onset and .90 s (SD = .86) for the offset. The frequency, intensity shifts, duration, and intonation of vocalizations were controlled across the conditions by using the moving-synchronous vocalizations as a model for the moving-asynchronous vocalizations. The vocalizations in all conditions occurred in an irregular temporal pattern.

Apparatus. Three video-decks (Panasonic, AG 7750) connected to a 4 × 4 input matrix switcher (Kramer, SV-6E) were used to present the displays on a 20″ color video-monitor (Sony, KV 20M10). The vocalizations were presented at 65-70 dB via a speaker (Trans Audio) situated under and in front of the monitor. Infants were seated in an infant seat facing the TV monitor approximately 55 cm away. Hidden from the infants, one or two observers recorded visual fixations through two peepholes located in a black screen above the television monitor. Each observer recorded visual fixations by pressing one of two buttons on a box connected to a computer (AT 386, IBM compatible) and a printer located in an adjacent room.

Procedure. Infants were habituated to two alternating vowel-object pairs in the moving synchronous, still, and moving-asynchronous conditions. They received /a/ paired with the porcupine and /i/ with the crab or the reverse vowel-object pairing, or /a/ with the star and /i/ with the lambchop, or the reverse pairing. The trial length was controlled by the infants. Each trial terminated when the infants looked away for 1.5 s or longer. The inter-trial interval was 6 s. The habituation criterion was defined as looking on two consecutive trials less than 50% of the mean of the first two trials. Following habituation, infants received two no change post-habituation trials to allow for spontaneous regression effects (Bertenthal, Haith & Campos, 1983). Following the post-habituation trials, infants received four test trials. These consisted of two no change (control) and two change (mismatch) trials in which the vowel-object pairings were switched. For example, if /a/ were vocalized with the star and /i/ with the lambchop, /a/ was now vocalized with the lambchop and /i/ with the star under the respective condition. Half the subjects in each condition received the control trials first and the remaining half received the mismatch trials first. Prior to the habituation sequence and after the test trials, subjects received a single trial of the moving turtle to test for fatigue. Infants whose visual fixation to the last turtle display was less than 20% of looking to the first turtle display were not included in the final sample. Trained observers recorded infants' visual fixations by initiating all trials

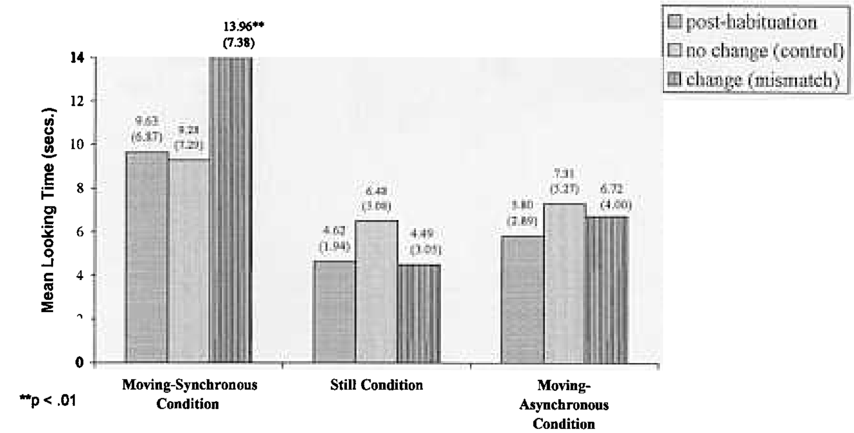

FIG. 1. 7-month-old infants' mean looking (and SD) to vowel-object pairs in the moving-synchronous, still, and moving-asynchronous conditions as a function of trial type.

when the infant fixated on the monitor and terminating the trials when the infant looked away for 1.5 s. Interobserver reliability, was calculated for 17 out of 48 subjects (35%). The number of seconds of looking recorded by the primary observer in comparison to that of a second observer was correlated across trials for each subject. The mean correlation across subjects was .985 (SD = .029).

RESULTS AND DISCUSSION

The results of this experiment are depicted in Fig. 1. In order to determine under what conditions infants learned the vowel-object relations, planned matched sample t-tests on infants' looking to the change trials versus the no change trials were conducted for each condition. Results were significant for infants in the moving-synchronous condition ($M = 4.68$, $t(15) = 3.23$, $p < .01$) but not for those in the still condition ($M = -1.99$, $t(15) = 1.73$, $p > .1$) or the moving-asynchronous condition ($M = -.59$, $t(15) = .47$, $p > .1$). Thus, consistent with our hypothesis, infants in the moving-synchronous condition showed evidence of learning the vowel-object relations but those in the still and moving-asynchronous conditions did not.

In order to compare across conditions, visual recovery to the mismatch trials was calculated by subtracting the mean looking on the two control (no change) trials from that of the two mismatch (change) trials in each condition. A mean recovery score was calculated by averaging the visual recovery of subjects for each condition (see Fig. 2). An analysis of variance on the visual recovery to the mismatch trials for the moving-synchronous, still, and moving-asynchronous conditions revealed a significant main effect of condition ($F(2, 45) = 7.42$, $p =$ .002). Tukey's post hoc tests indicated that visual recovery to the mismatch trials

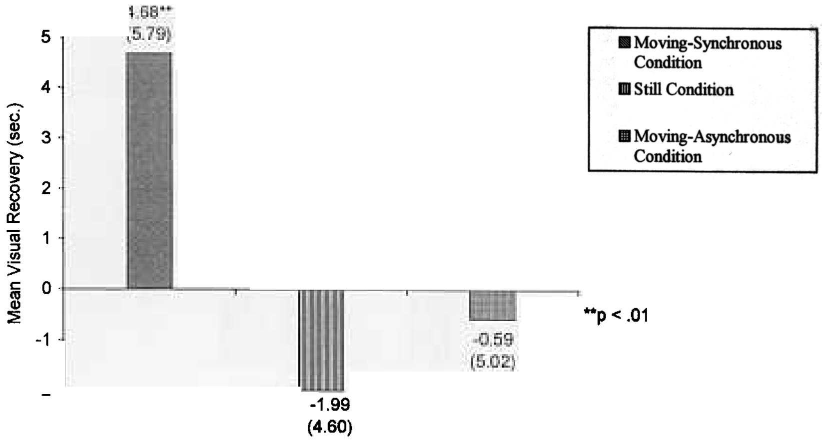

FIG. 2. Mean visual recovery to change (and SD) in vowel-object relations in the moving-synchronous, still, and moving-asynchronous conditions.

for the moving-synchronous condition differed significantly from that of the still and the moving-asynchronous conditions ($p < .05$).

An examination of individual subjects' visual recovery scores indicated that in the moving-synchronous condition, 11 out of the 16 infants showed a positive visual recovery (detected the change in the vowel-object relations). However, in the moving-asynchronous and still conditions only 4 and 3, out of the 16 infants in each condition, respectively, showed a positive visual recovery. A comparison of infants' success (positive versus negative visual recovery treated as a dichotomous variable) across conditions, indicated that infants in the moving-synchronous condition showed a positive visual recovery more often than those in the moving-asynchronous ($z = 2.76$, $p = .003$), and the still conditions ($z = 3.30$, $p = .0005$).

Secondary analyses assessed the effects of type of vowel-object pairing and order of presentation of no change (control) and change (mismatch) test trials on the visual recovery across conditions. A three-way analysis of variance of condition (3), type of vowel-object pairing (4), and test trial order (2) indicated a main effect of condition ($F(2, 24) = 7.48$, $p = .003$), but no effect of type of vowel-object pairing ($F(3, 24) = 1.83$, $p > .1$) or order of test trials ($F(1, 24) = .037$, $p > .1$). No interaction effects were found ($ps > .1$).

Furthermore, separate analyses assessed subjects' performance during the habituation phase to determine whether the following five measures, depicted in Table 1, differed as a function of condition. The measures assessed were (1) baseline, defined as the mean visual fixation on the first two habituation trials, (2) mean fixation on the last two habituation trials (seconds), (3) average fixation on the two post-habituation trials (seconds), (4) the mean total time taken to reach the habituation criterion (seconds), and (5) the mean number of trials required to

TABLE 1
Mean Looking (Seconds) and *SD* for the Baseline, the Last Two Habituation Trials, the Post-habituation Trials, the Total Time Taken to Habituation, and the Mean Number of Trials to Reach the Habituation Criterion for the Moving-Synchronous, Still, and Moving-Asynchronous Conditions

	Condition		
	Moving-Synchronous (N = 16)	Still (N = 16)	Moving-Asynchronous (N = 16)
Baseline	30.29 (18.47)[a]	12.44 (7.25)[a]	20.16 (14.19)
Last two habituation trials	6.92 (3.98)[b]	3.35 (1.76)[b]	5.47 (4.25)
Post-habituation trials	9.63 (6.87)[c]	4.62 (1.94)[c]	5.80 (2.89)
Total time to habituation	153.94 (72.66)[d]	84.95 (59.28)[d]	117.20 (57.86)
Number of trials to habituation	8.38 (2.58)	8.81 (3.19)	9.25 (3.53)

[a,b,c,d] Tukey's two-tailed post hoc, $p < .05$.

reach the habituation criterion. No main effect of condition was found for the mean number of trials to reach habituation criterion ($p > .1$). However, a significant main effect of condition was obtained for the baseline ($F(2, 45) = 6.47, p = .003$), the mean fixation on the last two habituation trials ($F(2, 45) = 4.19, p = .022$), the mean fixation on the post-habituation trials ($F(2,45) = 5.54$, $p = .007$), and the mean number of seconds to reach the habituation criterion ($F(2, 45) = 4.71, p = .014$). Post hoc *t* tests (Tukey) indicated that infants' visual fixation was significantly greater for the moving-synchronous condition than for the still condition for these four variables ($p < .05$, two-tailed). However, they did not differ across the moving-synchronous and moving-asynchronous conditions ($p > .1$; see Table 1). This suggests that the movement attracted infants' attention to the objects. Although one might argue that greater learning time could account for the superiority of matching under the moving-synchronous condition over the still condition, it should be noted that infants under both conditions controlled the length of their viewing times and reached the habituation criterion. Further, the lack of difference between looking times during habituation for the moving-synchronous and moving-asynchronous conditions demonstrates that differential learning opportunity cannot account for the superiority of matching in the moving-synchronous over the moving-asynchronous condition. Rather, the synchronous relation between the movement and sounds is responsible for the enhanced learning. This was further supported by the finding that mean looking time on the control (no change) test trials did not differ as a function of condition (F (2, 45) = .93, $p > .1$).

In summary, the 7-month-olds detected the arbitrary relations between simple speech sounds and objects when the visual and acoustic information was dynamic and temporally coordinated. However, infants failed to detect these relations when the objects remained still (even though they were concurrent with the speech sounds), or when the objects moved but were asynchronous with the vocalizations. These findings show that infants' detection of arbitrary relations between auditory and visual patterns of stimulation, such as a vowel sound and an object, is facilitated by movement and synchrony between the vocalization and the object.

GENERAL DISCUSSION

This study underscores the importance of intersensory redundancy for the detection of arbitrary vowel-object relations. The results showed that 7-month-old infants detected the arbitrary intermodal relations between temporally synchronous vocalizations and moving objects after a relatively short exposure (2–3 min). These arbitrary relations were detected when the object was shown moving in synchrony with the speech sounds but not when it was still or moving out of synchrony with the speech sounds. Thus, movement *per se* was not sufficient for learning. Rather, the synchrony between the moving object and the speech sound was necessary to specify the link between the vowels and the objects. These findings suggest that temporal synchrony served as the basis for learning the arbitrary relation between the speech sounds and objects. Infants showed learning in the presence of amodal information but not in its absence. When learning spoken label-referent relations, preverbal infants seem to benefit from temporal synchrony during repeated presentations in order to establish a link between labels and objects. Apparently, intersensory redundancy provides a basis for learning arbitrary spoken label-object relations, which, in turn, may be an important precursor to lexical comprehension.

Furthermore, these findings, from 7-month-olds, support the hypothesis that infants detect arbitrary relations between speech patterns and objects at approximately the same age as they detect arbitrary relations in nonspeech events (Bahrick, 1994; Hernandez-Reif & Bahrick, submitted; Reardon & Bushnell, 1988). Further, if infants perceived the vowel sounds as "speech," these parallel findings in the bimodal speech and nonspeech domains would suggest the use of similar perceptual-cognitive processes for the detection of arbitrary relations in bimodal events across the two domains. It is likely that infants perceived the vowel sounds (verbal labels) as speech sounds rather than as sounds differing along a single simple dimension such as overall pitch. Since the vowels were approximately equated for their overall pitch and intensity, infants could not have used these simple dimensions to discriminate between them. Further, the vowel sounds, although equated for overall intensity and pitch, showed a great deal of within type variability in terms of intonation contour (rising versus falling), intensity shifts, and duration, similar to natural speech. Infants were therefore required to generalize across these varied examples of each vowel. It may also be that infants, in the present study, discriminated the vowel sounds on the basis of their characteristic acoustic properties such as formant frequencies, sonority levels, and phonemic qualities. Detection of these properties by 7-month-olds would qualify as "special treatment of speech" in the presence of objects, in keeping with Baldwin's (1995) thesis. If so, infants would have perceived the vowels as different from nonspeech sounds such as the impact sounds of objects hitting a surface used in prior studies.[2] Of potential relevance, we noted that many infants in this study responded to the vocalizations with babble and imitation, especially of /a/, during the habituation trials. No such responses were reported in prior studies where infants were presented with objects arbitrarily paired with nonspeech sounds (Spelke, 1979, 1981), or objects making impact sounds that arbitrarily varied in their pitch (Bahrick, 1994). Whether or not infants treated the stimuli as speech sounds is, of course, an empirical question that deserves further investigation. Equally intriguing is the question as to whether and when infants treat speech sounds as names for objects rather than as mere auditory covariates.

The present study showed that infants detect the arbitrary relations between speech-sounds and objects. Our findings along with studies of word recognition show a readiness for word learning by the age of 7- to 8-months. Seven-month-old infants showed recognition of words in a passage or a list (Jusczyk & Aslin, 1995). Further, given a 2-min exposure to three-syllable nonsense words and surrounding syllables (e.g., bidakupado), 8-month-olds detected the boundaries of the nonsense words based solely on their probability of occurrence with the neighboring syllables during exposure (Saffran, Aslin, & Newport, 1996). Eight-month-olds also oriented longer to specific objects when mothers labeled them, indicating increased salience for labeled objects (Jackson-Maldanado et al., 1993).

Furthermore, the current findings in conjunction with those of Werker and her colleagues (Lloyd et al., 1994; Werker & Stager, in press) suggest that the importance of amodal information such as temporal synchrony for the detection of speech-object relations may decline with age. The results of the present study, demonstrated that learning did not occur in 7-month-olds on the basis of movement alone, but that synchrony was required for learning vowel-object relations. However, Lloyd et al. (1994) showed that learning did occur in 14-month-olds but not in 8-month-olds (see Werker & Stager, in press) when syllables were presented in conjunction with moving objects, despite a lack of synchrony. Further, object motion also seems to be important for learning label-object relations through the age of 14-months (Lloyd et al., 1994). This was demonstrated when infants failed to learn these relations when still objects were presented along with labels at 7-months, in the current study, and at 14-months (Lloyd et al., 1994). Taken together, these findings suggest that temporal synchrony between speech sounds and object motion is important for learning speech-object relations early in development (7- to 8-months).

[2] A comparison of initial interest levels (baseline) of 7-month-olds from Bahrick's (1994) study ($N = 16$), where synchrony between sound and object motion was provided, and that of infants ($N = 32$) tested under the moving-synchronous condition of this and author study revealed the following. A two-sample t test on baseline looking (seconds) was marginally significant ($t\ (1, 48) = 1.91, p = .06$), suggesting that infants at 7-months show greater attention to objects in the presence of speech ($M = 32.3$ s, $SD = 18.7$) compared to nonspeech sounds ($M = 22.84$, $SD = 8.85$).

Later (by 14-months), when infants have learned that speech sounds refer to objects, synchrony may become less important while movement is still necessary. Movement may serve to highlight which of the many simultaneously available visual events the infant might attend to in the presence of a sound, whereas synchrony conveys that a sound and an object belong together (Bahrick & Pickens, 1994). Further developmental research with objects and more word-like vocalizations that are synchronous, asynchronous or still are needed to examine how the requirements for word learning change with age.

The current findings, demonstrating infants' early ability to match dynamic visual events with simple spoken monosyllables, provide evidence for an ecological-perceptual basis for emerging lexical comprehension. The results show that, given synchronous speech patterns and moving objects, even 7-month-old-infants can perceive the arbitrary relations between the speech sounds they hear and the objects or events they see. We conclude that such perceptual learning likely lays the foundation for the development of lexical comprehension and production during infancy. This conclusion is consistent with that of prior descriptive studies, suggesting that lexical learning is facilitated when infants are guided by gestures from adults (such as showing objects) to attend to conventional relations between words and objects or events (Zukow, 1991, 1990, 1997; Dent, 1990). According to Zukow (1991, 1997), caregivers socially transmit knowledge of linguistic conventions by directing infants' attention to objects and events via speech and accompanying gestures. The results of the present study lend empirical support for Zukow's (1991, 1997) proposal that coordination of speech and gesture during infant-directed communication highlights speech-object-meaning relations for infants and facilitates early lexical learning. The detection of temporal synchrony between audible speech patterns and visible object motions may be an important basis for detecting word-object relations in the natural environment. If so, maternal communication to pre-lexical infants may contain multimodally coordinated and temporally synchronous auditory, visual, and even tactile components. This hypothesis was supported by a recent study of "multimodal motherese" (Gogate, Bahrick, & Watson, submitted). Mothers were asked to teach their infants novel names for objects and actions. Results indicated that mothers primarily used temporal synchrony between the words and object motions to teach word-referent relations. Further, they used temporal synchrony more often for infants of 5–8 months than for infants of 21–30 months. This is consistent with the notion that temporal synchrony is important for teaching spoken label-object relations early in development, but is less important later on. Other observational studies support this hypothesis. For example, Zukow (1991) reported several instances of what appear to be multimodally coordinated communication in her observations of maternal communication across cultures. Similarly, Messer (1978) reported that maternal communication to infants consists of temporally coordinated speech and object motion. Further systematic analyses of maternal interactions with young infants would reveal more about the nature and extent of redundancies between vocal- and gestural-components provided by mothers.

We suggest that infants' detection of intersensory redundancy uniting arbitrarily related speech patterns and events is an important means by which lexical comprehension develops. The ability to detect arbitrary speech-object relations across the auditory and visual modalities, is an important milestone for the development of lexical comprehension. The knowledge that a spoken word 'stands for' an object most likely begins with the perceptual ability to relate specific speech sounds to objects or events on the basis of redundant information such as temporal synchrony in the second half of the first year of life.

REFERENCES

Bahrick, L. E. (1983). Infants' perception of substance and temporal synchrony in multimodal events. *Infant Behavior and Development,* **6,** 429–451.

Bahrick, L. E. (1987). Infants' intermodal perception of two levels of temporal structure in natural events. *Infant Behavior and Development,* **10,** 387–416.

Bahrick, L. E. (1988). Intermodal learning in infancy: Learning on the basis of two kinds of invariant relations in audible and visible events. *Child Development,* **59,** 197–209.

Bahrick, L. E. (1992). Infants' perceptual differentiation of amodal and modality-specific audio-visual relations. *Journal of Experimental Child Psychology,* **53,** 180–199.

Bahrick, L. E. (1993). Intermodal learning and transfer in three-month-old infants. Presented at the Biennial Meeting of the Society for Research in Child Development, March. New Orleans.

Bahrick, L. E. (1994). The development of infants' sensitivity to arbitrary intermodal relations. *Ecological Psychology,* **6**(2), 111–123.

Bahrick, L. E., & Pickens, J. (1994). Amodal relations: The basis for intermodal perception and learning in infancy. In D. Lewkowicz & R. Lickliter (Eds.), *The development of intersensory perception: Comparative perspectives* (pp. 205–233). Hillsdale, NJ: Erlbaum.

Baldwin, D. A., & Markman, E. M. (1989). Establishing word-object relations: A first step. *Child Development,* **60,** 381–398.

Baldwin, D. A. (1995). Understanding the link between joint attention and language. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 131–158). Hillsdale, NJ.: Erlbaum.

Bates, E. (1993). Comprehension and production in early language development: A commentary. In S. Savage-Rumbaugh, J. Murphy, R. A. Sevcik, K. E. Brakke et al. (Eds.), *Monographs of the Society for Research in Child Development* (Vol. 3–4, pp. 222–242).

Benedict, H. (1979). Early lexical development: Comprehension and production, *Journal of Child Language,* **6,** 183–200.

Bertenthal, B. I., Haith, M. M., & Campos, J. J. (1983). The partial-lag design: A method for controlling spontaneous regression in the infant-conrolled habituation paradigm. *Infant Behavior and Development,* **6,** 331–338.

Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P. W., Kennedy, L. J., & Mehler, J. (1988). An investigation of young infants' perceptual representations of speech sounds. *Journal of Experimental Psychology: General,* **117**(1), 21–33.

Blake, J., & Fink, R. (1987). Sound-meaning correspondences in babbling. *Journal of Child Language,* **14,** 229–253.

Dent, C. H. (1990). An ecological approach to language development: An alternative functionalism. *Developmental Psychobiology,* **23**(7), 679–703.

Ellis, R. & Wells, C. G. (1980). Enabling factors in adult-child discourse. *First Language,* **1,** 46–62.

Fernandez, M. & Bahrick, L. E. (1994). Infants' sensitivity to object-odor pairings. *Infant Behavior and Development,* **17,** 463–466.

Gibson, E. J. (1969). *Principles of perceptual learning and development.* New York: Appleton-Century-Crofts.

Gibson, E. J. (1991). *An odyssey in learning and perception.* Cambridge, Mass: MIT Press.

Gogate, L. J., Bahrick, L. E. & Watson, J. D. *A cross-cultural study of maternal multimodal motherese: The use of temporal synchrony in maternal communication with infants.* Manuscript submitted for publication.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: lexical and syntactic comprehension in a new paradigm. *Journal of Child Language,* **14,** 23–45.

Halliday, M. A. K. (1975b). *Learning how to mean: Explorations in the development of language.* New York: Elsevier-North Holland.

Harris, M. (1992). *Langauge experience and early language development: From input to uptake.* Hillsdale, NJ: Erlbaum.

Hernandez-Reif, M. A., & Bahrick, L. E. *The development of visual-tactual perception of objects: Amodal relations provide the basis for learning arbitrary relations.* Manuscript submitted for publication.

Huttenlocher, J. (1974). The origins of language comprehension. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium.* Hillsdale, NJ: Erlbaum.

Jackson-Maldanado, Thal, D., Markman, V., Bates, E., Guttierrez-Clellen, V. (1993). Early lexical development in Spanish-speaking infants and toddlers. *Journal of Child Language,* **20,** 523–549.

Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology,* **29,** 1–23.

Kuhl, P. K. (1992). Psychoacoustics and speech perception: Internal standards, perceptual anchors, and prototypes. In L. A. Werner & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 293–332). Washington, DC: American Psychological Association.

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science,* **218,** 1138–1141.

Kuhl, P. K., & Meltzoff, A. N. (1984). The intermodal representation of speech in infants. *Infant Behavior and Development,* **7,** 361–381.

Kuhl, P. K., & Meltzoff, A. N. (1988). Speech as an intermodal object of perception. In A. Yonas (Ed.), *Perceptual development in infancy: The Minnesota Symposia on Child Psychology* (Vol. 20, pp. 235–266). Hillsdale, NJ: Erlbaum.

Lawson, K. R. (1980). Spatial and temporal contiguity and auditory-visual integration in infants. *Developmental Pyschology,* **16,** 185–192.

Legerstee, M. (1990). Infants' use of multimodal information to imitate speech sounds. *Infant Behavior and Development,* **13,** 343–354.

Lewkowicz, D. J. (1992). Infants' responsiveness to the auditory and visual attributes of a sounding/moving stimulus. *Perception and Psychophysics,* **52**(5), 519–528.

Lloyd, V. L., Cohen, L. B., Werker, J. F., Foster, R., & Swanson, C. S. (1994). *Gender and motion: Important factors in infants' ability to learn word-object associations.* Presented at the International Conference on Infant Studies, Paris, June.

Lyons-Ruth, K. (1977). Bimodal perception in infancy: Response to auditory-visual incongruity. *Child Development,* **48,** 820–827.

MacKain, K., Studdert-Kennedy, M., Spieker, S., & Stern, D. (1983). Infant intermodal speech perception is a left-hemisphere function. *Science,* **219,** 1347–1349.

Mandell, D.R, Jusczyk, P.W., & Pisoni, D.B. (1995). Infants' recognition of the sound patterns of their own names. *Psychological Science,* **6**(5), 314–317.

Messer, D. J. (1978). The integration of mothers' referential speech with joint play. *Child Development,* **49,** 781–787.

Molfese, D. L., Morse, P. A., & Peters, C. J. (1990). Auditory evoked responses to names for different objects: Cross-modal processing as a basis for infant language acquisition. *Developmental Psychology,* **26**(5), 780–795.

Oviatt, S. (1980). The emerging ability to comprehend language: An experimental approach. *Child Development,* **51,** 97–106.

Pickens, J. (1994). Perception of bimodal distance relations by 5-month-old infants. *Developmental Psychology,* **30**(4), 537–544.

Reardon, P. & Bushnell, E. W. (1988). Infants' sensitivity to arbitrary pairings of color and taste. *Infant Behavior and Development,* **11,** 245–250.

Reznick, J. S. (1990). Visual preference as a test of infant word comprehension. *Applied Psycholinguistics,* **11,** 145–166.

Saffran, J. R., Aslin R. N., & Newport, E. L (1996). Statistical learning by 8-month-old infants. *Science,* **274,** 1926–1928.

Spelke, E. S. (1979). Perceiving bimodally specified events in infancy. *Developmental Psychology,* **15,** 626–636.

Spelke, E. S. (1981). The infant's acquisition of knowledge of bimodally specified events. *Journal of Experimental Child Psychology,* **31,** 279–299.

Stemmer, R. (1989). Empiricist versus prototype theories of language acquisition. *Mind and Language,* **4,** 201–221.

Sullivan, J. & Horowitz, F. D. (1983). Infant intermodal perception and maternal multimodal stimulation: Implications for language development. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 2, pp. 183–239). Norwood, NJ: Ablex.

Walker-Andrews, A. S. (1994). Taxonomy for intermodal relations. In D. Lewkowicz & R. Lickliter (Eds.), *The development of intersensory perception: Comparative perspectives* (pp. 39–56). Hillsdale, NJ: Erlbaum.

Walker-Andrews, A. S., Bahrick, L. E., Raglioni, S. S., & Diaz, I. (1991). Infants' bimodal perception of gender, *Ecological Psychology,* **3**(2), 55–75.

Walker-Andrews, A. S., & Lennon, E. (1985). Auditory-visual perception of changing distance by human infants. *Child Development,* **22,** 373–377.

Walton, G. E., & Bower, T. G. R. (1993). Amodal representation of speech in infants. *Infant Behavior and Development,* **16,** 233–243.

Waxman, S.R., & Balaban, M. T. (1996). *Ursines and felines: Novel words support object categorization in 9-month-old infants.* Paper presented at the International Conference on Infant Studies, Providence, RI.

Werker, J. F., & Stager, C. L. Developmental changes in infant speech perception and early word learning: Is there a link? In J. Pierrehumbert & M. Broe (Eds.), *Papers in Laboratory Phonology V.* To appear.

Woodward, A., Markman, E., & Fitzsimmons, M., (1994). Rapid word learning in 13- to18-month-olds. *Developmental Psychology,* **30,** 553–566.

Zukow, P. G. (1990). Socio-perceptual bases for the emergence of language: An alternative to innatist approaches. *Developmental Psychobiology,* **23**(7), 705–726.

Zukow, P. G. (1991). A socio-perceptual/ecological approach to lexical development: Affordances of the communication context. *anales de psicologia,* **7**(2), 151–163.

Zukow-Goldring, P. (1997). A social ecological realist approach to the emergence of the lexicon: educating attention to amodal invariants in gesture and speech. In C. Dent-Read & Zukow-Goldring (Eds.), *Evolving explanations of development: Ecological approaches to organism-environment systems* (pp. 199–252). Washington, DC: American Psychological Association.

Received: September 12, 1997; revised: March 3, 1998