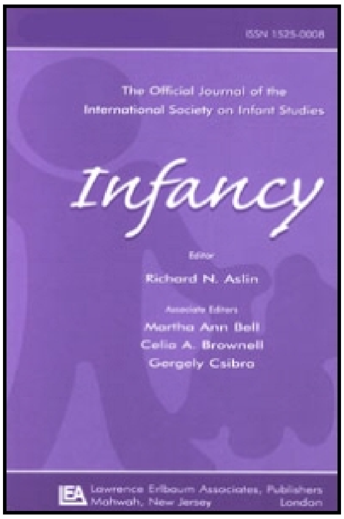

Infancy
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t775653654

Attention to Maternal Multimodal Naming by 6- to 8-Month-Old Infants and Learning of Word-Object Relations

Lakshmi J. Gogate [a]; Laura H. Bolzani [b]; Eugene A. Betancourt [b]
[a] Department of Psychiatry, State University of New York Health Science Center at Brooklyn. [b] Department of Psychology, Florida International University.

Online Publication Date: 01 January 2006

To cite this Article Gogate, Lakshmi J., Bolzani, Laura H. and Betancourt, Eugene A.(2006)'Attention to Maternal Multimodal Naming by 6- to 8-Month-Old Infants and Learning of Word-Object Relations',Infancy,9:3,259 — 288

To link to this Article: DOI: 10.1207/s15327078in0903_1

URL: http://dx.doi.org/10.1207/s15327078in0903_1

INFANCY, *9*(3), 259–288

RESEARCH ARTICLES

Attention to Maternal Multimodal Naming by 6- to 8-Month-Old Infants and Learning of Word–Object Relations

Lakshmi J. Gogate
Department of Psychiatry
State University of New York Health Science Center at Brooklyn

Laura H. Bolzani and Eugene A. Betancourt
Department of Psychology
Florida International University

We examined whether mothers' use of temporal synchrony between spoken words and moving objects, and infants' attention to object naming, predict infants' learning of word–object relations. Following 5 min of free play, 24 mothers taught their 6- to 8-month-olds the names of 2 toy objects, *Gow* and *Chi,* during a 3-min play episode. Infants were then tested for their word mapping. The videotaped episodes were coded for mothers' object naming and infants' attention to different naming types. Results indicated that mothers' use of temporal synchrony and infants' attention during play covaried with infants' word-mapping ability. Specifically, infants who switched eye gaze from mother to object most frequently during naming learned the word–object relations. The findings suggest that maternal naming and infants' word-mapping abilities are bidirectionally related. Variability in infants' attention to maternal multimodal naming explains the variability in early lexical-mapping development.

Correspondence should be addressed to Lakshmi J. Gogate, Psychiatry #1203, SUNY at Brooklyn, Brooklyn, NY 11203. E-mail: lgogate@downstate.edu

Several studies have underscored the important role that caregivers play in enabling infants' attention to spoken language (Cooper & Paccia-Cooper, 1980; Fernald, 1992; Garnica, 1977; Gogate, Bahrick, & Watson, 2000; Lewkowicz, 1996a; Sullivan & Horowitz, 1983; Zukow-Goldring, 1997). However, studies have yet to examine how mothers provide the conditions that preverbal infants require for learning word–object relations, and how preverbal infants learn novel word–object relations from their mothers during interactive play. Researchers have yet to assess the specific nature of preverbal infants' attention to maternal naming. For example, although some studies have reported that preverbal infants' gaze-following ability and word comprehension during the second year are related (e.g., Morales, Mundy, & Rojas, 1998), little is known about the extent to which preverbal infants engage in joint attention or switch eye gaze from their mother to an object during maternal naming. The research reported here assessed how mothers could teach their preverbal infants two novel word–object relations. In particular, we examined how mothers' predominant naming styles and infants' attention to maternal naming contribute to infants' learning of word–object relations.

Several factors, extrinsic or intrinsic to the infant, are potential contributors to word learning in early infancy. One major extrinsic factor that has been identified is that mothers initially scaffold their infants' attention to objects or actions and their names during social interactions (Zukow-Goldring, 1990, 2001; also see Nelson, 1988; Tomasello & Farrar, 1986; cf. Akhtar, Dunham, & Dunham, 1991; Rollins, 2003). Mothers begin to direct their infants' attention to objects by socializing object attention (i.e., they make attending to an object a social activity) during naming. They do this by "embedding the object within the interpersonal sphere" (Bakeman & Adamson, 1984, p. 1288), long before infants can attend simultaneously to the mother and an object by themselves. Prior to 6 months of age, mothers engage their infants in passive joint visual attention where the mother displays an object of interest to her infant, enabling the infant to attend to and explore it. The infant, however, may not necessarily glance at the mother while exploring the object. With development, there is a gradual shift in the infant's participation in the mother–infant–object triad. At about 6 months of age, infants begin to switch their attention between the caregiver and the object (Newson & Newson, 1975; also see Butterworth & Jarrett, 1991; D'Entremont, Hains, & Muir, 1997; Scaife & Bruner, 1975; for a review, see Butterworth, 1998; cf. Carpenter, Nagell, & Tomasello, 1998). When the mother displays an object, the infant switches eye gaze from the mother to the object as long as the object is within the infant's visual field. This qualitative developmental shift in infants' joint visual attention may contribute significantly to the onset of lexical development (Nelson, 1979). The ability to switch eye gaze and attention from mother to object should enable infants to attend to word–object relations during *en face* interactions. Looking at (and presumably listening to) the mother first could focus the infant's attention on the utterance, while shifting gaze from the mother to the object might assist in linking the heard word and the seen object. Eye gaze switching from mother to object might be a precursor to infants' later developing ability to follow maternal gaze to an object, and language comprehension in the second year (see Morales et al., 1998; Silven, 2001).

Another major extrinsic factor in mother–infant naming complements the developmental shift in joint attention and could potentially impact infants' early lexical development. Some studies have shown that when mothers name novel objects or actions for their preverbal infants, they use intersensory redundancy, in the form of temporal synchrony, or simultaneous spoken words and visual object motions (Gogate et al., 2000; Messer, 1978; Zukow-Goldring, 1997). During temporally synchronous naming, they speak a word while holding and moving an object rather than moving it out of phase with the spoken word. Mothers of preverbal infants (5–8 months) used temporally synchronous naming more often in a play episode than the mothers of toddlers (21–30 months; Gogate et al., 2000). They used it to teach novel words more often to young infants who were beginning to detect word–referent relations and likely needed maternal assistance to perceive the object and its name as a unified event. They used temporal synchrony less often for toddlers who were likely adept at detecting word–referent relations. Mothers named static objects more often for these toddlers. These findings suggested that mothers' use of temporal synchrony in multimodal naming is well adapted to infants' developing lexical mapping abilities. Further, because mothers spontaneously used temporal synchrony to teach the names of target word labels for objects and actions more often compared to when simply mentioning nontarget object labels in passing, the authors hypothesized that mothers perceptually highlight novel word–referent relations for young infants during multimodal naming. This hypothesis is implicit in the view that learning of word–object relations is "accomplished in the context of the infant's multimodal interactions with its mother" (Sullivan & Horowitz, 1983, p. 210). Temporal synchrony differs from temporal contiguity, which also plays a role in word learning. Temporal contiguity merely requires that a static object be present when a word is spoken (see Slater, Brown, & Badenoch, 1997). In contrast, temporal synchrony requires object motion and a temporal alignment between an utterance and the object's motion.

Complementing the findings on maternal use of intersensory redundancy during multimodal naming, controlled experiments have shown that young infants rely heavily on perceptual cues such as intersensory redundancy for word mapping. Thus, temporal synchrony between spoken syllables and objects' motions facilitates 7- to 8-month-olds' learning and memory for two syllable–object pairings on video displays (Gogate & Bahrick, 1998, 2001; Gogate, Bewley, Wan, Kremen, & Chen, 2003). Following habituation to two synchronous syllable–object pairs, infants looked longer to a switch in the pairings relative to control test trials.

When object motion alone was provided during habituation (Werker, Cohen, Lloyd, Casasola, & Stager, 1998), in the absence of temporal synchrony, infants did not learn the two syllable–object pairings until 14 months. One possible reason for the discrepancy between Werker's and Gogate's findings is that, in Werker's experiments, the lack of synchrony between the spoken syllables and object motions perceptually highlights the separation between syllables and objects for young infants. In contrast, in Gogate's experiments, the temporal synchrony perceptually unifies the syllables and objects and foregrounds the relation for young infants. A comparison of these results is difficult because the syllable contrasts and stimulus displays differ (the objects moved without an agent [Werker et al., 1998] or were held by a disembodied hand [Gogate & Bahrick, 1998]).

In general, the maternal and infancy research suggests that maternal naming and infant lexical-mapping abilities are bidirectionally related (Gogate, Walker-Andrews, & Bahrick, 2001). For example, mothers provide temporal synchrony during naming for their preverbal infants, and temporal synchrony facilitates young infants' learning of word–object relations. If this is the case, mothers should provide perceptual conditions during naming that contribute to young infants' lexical mapping. It is necessary, therefore, to examine maternal naming and infant perceptual-lexical and attentional systems in word learning contexts. They can be examined by observing mothers' spontaneous naming when they teach their infants the names of objects during play, and studying how infants attend to this naming and learn the word–object relations on a post-test. If mothers' spontaneous temporally synchronous object naming covaries with infants' learning of word–referent relations, it would support the view that maternal naming and infants' lexical-mapping abilities are bidirectionally related. Furthermore, anecdotal reports of word mapping around 8 to 9 months (Halliday, 1975; Jackson-Maldonado, Thal, Marchman, Bates, & Gutierrez-Clellan, 1993) and experimental findings from 6-month-olds who relate familiar referents (e.g., their mother and father) with the corresponding words (Tincoff & Jusczyk, 1999), warrant examining the developmental transition from prelinguistic beginnings to early word-mapping capabilities between 6 and 8 months. Systematic variability observed in maternal naming, infant attention, and learning of word–object relations during this transition might elucidate the developmental process of lexical mapping (see Smith & Thelen, 2003).

In this research, first, we hypothesized the following with respect to the play episode (see Table 1): (a) If temporal synchrony between spoken words and objects' motions highlights word–object relations for preverbal infants, mothers should use it more often than other naming styles (e.g., naming an object not simultaneous with object movement, or naming a static object) to highlight novel relations for 6- to 8-month-olds. (b) Next, we hypothesized that mothers' temporally synchronous object naming would predict infants' attention to object naming more than other naming types. A parallel hypothesis of this study was that (c) infants

TABLE 1
The Hypotheses, the Variables, and the Analyses of This Study

Hypotheses	*Variable 1 (Dependent)*	*Variable 2 (Independent)*	*Variable 3 (Covariate)*	*Variable 4 (Covariate)*	*Variable 5 (Covariate)*	*Type of Analyses*
1. Mothers use temporal synchrony more often than other naming types to highlight novel word–object relations for preverbal infants	Proportion of total target words (PTTW)	Maternal bimodal naming types (synchronous – *s*; asynchronous – *a*; static – *st*; infant holds object, *iho*)				Repeated measures analysis of variance (general linear models procedure)
2. Mothers' temporally synchronous object naming predicts in part infant attention to objects	Proportions of infant looks to objects and gaze switch from mother to object combined (Table 2) during *s, a, st*, and *iho*	Proportions of maternal bimodal naming types (*s, a, st*, and *iho*)				Linear regression analyses
3. More number of infants gaze-switch during mothers' temporally synchronous naming than other types	Infant attention success during maternal naming (e.g., gaze-switch from mother to object)	Maternal bimodal naming type (*s, a, st*, or *iho*)				Binomial tests comparing dichotomous variables
4a. Mothers who use temporal synchrony more often have infants who learn word–object relations better	Infants' proportion of first looks (PFL) to word-matched object display for Blocks 1 and 2 on test	Proportions of maternal temporally synchronous and nonsynchronous naming (columns 1 and 5, Table 2)				Pearson correlations

(continued)

TABLE 1 *(Continued)*

Hypotheses	*Variable 1 (Dependent)*	*Variable 2 (Independent)*	*Variable 3 (Covariate)*	*Variable 4 (Covariate)*	*Variable 5 (Covariate)*	*Type of Analyses*
4b. Infants who attend more often to mothers' temporally synchronous object naming learn word–object relations better	Infants' PFLs to the word-matched object display for Blocks 1 and 2 on test	Proportions of infant attention by maternal naming (e.g., gaze-switch from mother to object when mothers use temporal synchrony, Table 2)				Pearson correlations
5. Infants' word mapping is a function of multiple factors	Infants' PFLs to the word-matched object display on test for Blocks 1 and 2, and in each block	Infants' ability to gaze-switch from mother to object (two groups—high and low)	Maternal bimodal naming types (PTTW—*s, a, st*, or *iho*)	Infant attention to objects by maternal naming type	Infants' overall attention to test displays	Analyses of covariance (general linear models procedure)

would engage in joint attention behaviors such as switching eye gaze from mother to object during maternal naming of objects. We hypothesized that if perception of intersensory redundancy and gaze-switching behaviors are closely linked (Stein, Jiang, & Stanford, 2004), more infants would engage in gaze-switching behavior when mothers used temporal synchrony than during other naming types. To test these hypotheses, we asked mothers to teach their infants the names for two distinct objects during a semi-structured play episode. The episodes were videotaped and coded for maternal multimodal naming such as temporal synchrony during target word naming, similar to Gogate et al. (2000). Infants' attention to maternal naming during play was coded for infants' looks to the target objects, their mother, elsewhere, or eye gaze switch from mother to object or object to mother, when the mothers named the target objects.

Further, we hypothesized the following with respect to the test phase and its relation to the play episode: (d) If young infants' word learning is a function of their attention to maternal naming and mothers' extent of temporally synchronous naming, then variability in infants' learning of word–object relations during the play episode should be accounted for by variability in these factors. Thus, mothers who use temporal synchrony more often should have infants who learn word–object relations better (e). And infants who attend more often to their mothers' temporally synchronous naming of objects should also learn word–object relations better (f). To test these hypotheses, following the play episode, infants were given a two-choice intermodal word-mapping test (Spelke, 1979). Several versions of this method have been used to study word mapping (Gogate & Bahrick, 2001; Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Schafer & Plunkett, 1998; Tincoff & Jusczyk, 1999; also Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998, measured latency during word recognition). In this study, it was used for the first time to examine infants' learning of word–object relations from maternal naming during play. (g) Finally, we hypothesized that infants' word mapping is a function of multiple factors, such as maternal use of temporal synchrony, infants' attention to objects during temporally synchronous naming, infants' ability to switch eye gaze from mother to object, and infant attention to the test displays. Specifically, we predicted that infants' gaze-switching from mother to object would influence their word learning.

METHOD

Participants

Twenty-four healthy, full-term 6- to 8-month-old infants ($M = 218$ days, $SD = 22.5$, range = 183–249 days [6.1–8.3 months]), 10 boys and 14 girls, and their mothers, participated. The infants had no known history of birth defects, weighed at least

5.5 pounds at birth, and their Apgar scores at birth and 5 min later were at least 9 and 9. Seven additional dyads were excluded from the final sample due to the mother's continued naming of an object (rather than switching to the second one) after the time limit (1.5 min; *n* = 1), infants' excessive fussiness during the test and failure to complete at least 10 out of 12 trials (*n* = 1), experimenter error during test (*n* = 2), equipment failure during video recording of play episodes (*n* = 2), or infants' visual impairment (*n* = 1). The mothers and infants were recruited from birth records. They resided in the southern suburbs of the Miami metropolitan area and were from middle-class socioeconomic backgrounds. The mothers had at least 12 years of education (high school graduate). Of the 24 mothers, 6 reported that their infant heard English at home, 6 reported that their infant heard Spanish, and the remaining 12 mothers reported that their infant heard Spanish and English. Mothers and infants from mono- and bilingual backgrounds were recruited. Although some studies have found differences in the way bilingual infants listen to speech early on (Bosch & Sebastian-Galles, 2003), the methods used herein have not revealed word-mapping differences in infants (Gogate & Bahrick, 1998, 2001), or in maternal naming styles as a result of either language background (Gogate et al., 2000). Of the 24 dyads one was African American, 13 were Hispanic American, and 10 were Caucasian.

Play Episode

Procedure. The mothers and infants were seen during a half-hour visit to the laboratory. First, one of the experimenters introduced the two nouns and their referents to the mother. They were *Chi* and *Gow,* the names for one of two brightly colored hand puppets, a Martian or a raccoon (see Figure 1). Proper nouns were chosen because, typically, they are learned earlier than other nouns (4.5-month-olds are familiar with their own names; Mandell, Jusczyk, & Pisoni, 1995). The experimenter stood behind the infant, who was seated in a stroller, and held and named each object for the mother, saying, "This is called Gow, and this is called Chi." Perhaps infants heard the words, but they did not see the objects until the mothers presented them during the play episode. The experimenters held the objects static during naming to preclude maternal mimicking of temporally synchronous naming during the play episodes. Mothers were asked to teach the names for the two objects using any means or language that they normally used to communicate with their infant.

Next, the infants were seated in an infant seat placed on a large table (5' × 5' × 2.5'), and the mothers sat cross-legged on the table facing their infant and approximately 15 cm away (the distance between the front edge of the infant seat and a mother's legs in sitting position) during the play episode. A camera (Sony DCR-VX1000) was positioned in front of the table so that the mother, the infant,

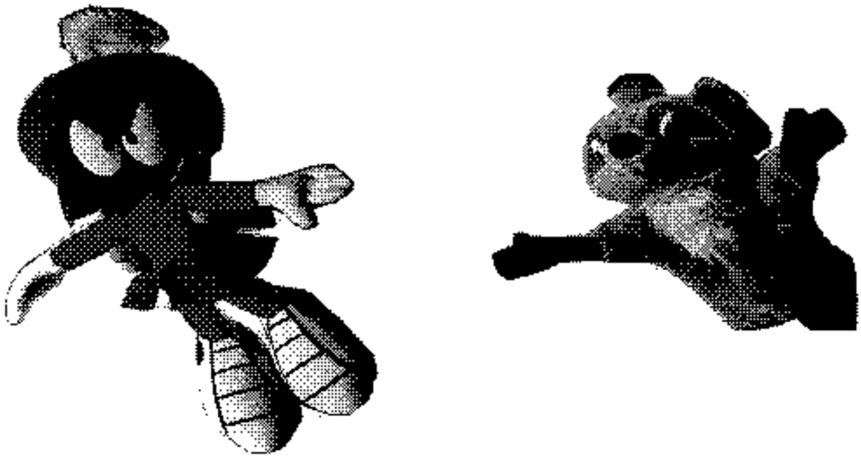

FIGURE 1 Stimuli: The toy objects mothers used to teach the word–object relations: the Martian and the raccoon named Chi or Gow.

and the camera formed three points of a right-angled triangle, with the camera at the 90° base. Positioned to the infant's right and to the mother's left, the camera focused on the mother and her infant during play. A mirror (36" wide × 60" high) was placed to the far end of the table against a wall and facing the camera, and to the infant's left and the mother's right. The mirror captured most mother–infant interactions that were hidden from direct view of the camera by either participant. For the first 5 min, mothers engaged their infant in free play with a set of toys. These toys remained on the table during the session. For the next 3 min, mothers taught the two target word–object relations one at a time to their infant during semi-structured play. An experimenter entered the room with the first target object at the start of the 3-min play period, and the second object 1.5 min later, via a door located behind the infant, and placed each object on the table within the mother's reach. This served as a cue for the mothers to place all other objects (including the first object when the second one was introduced) on the table (out of infants' immediate focus), and name the target objects for their infants. All infants received both words, with order of teaching counterbalanced across dyads. Twelve infants received Chi first, and 12 infants received Gow first. Further, the word–object relations were counterbalanced so that 12 infants received Chi with the raccoon and Gow with the Martian, and 12 others received Chi with the Martian and Gow with the raccoon. During the episodes, 10 mothers spoke English, 8 spoke Spanish, and 6 spoke English and Spanish to their infants.

Coding and Scoring of Play Episodes

Maternal naming of target objects. The videotaped 3-min semi-structured play episodes were coded to examine the multimodal (auditory and visual) properties of target words and gestures when mothers taught the object labels to their infants. A trained observer identified and coded the target word tokens of each dyad into four types of maternal bimodal naming using her best judgment: naming of an object synchronous with object movement (synchronous; *s*), naming asynchronous with object movement (asynchronous; *a*), naming an object without motion (static; *st*), and naming an object when the infant held and manipulated it (infant holds object, *iho*). A second trained observer identified and independently coded a portion (25%) of these dyads into the same four types. Some target word tokens (97 of 1,339, or 8.5%) were not coded either because mothers named the objects out of the camera's view, or named them while causing no deliberate action such as when wearing or adjusting the puppet on their hand, or due to observers' difficulty in categorizing the occurrences.

A bimodal naming occurrence was coded as synchronous if the mother uttered the target word while moving the object with a small discrepancy (< 150 msec, as verified later by a descriptive analysis; see later) between word onset or offset and object motion. A bimodal occurrence was coded as asynchronous if the word onset immediately preceded or followed the onset of an object's motion resulting in a greater discrepancy (> 150 msec, again, as verified by the descriptive analysis) between target word onset or offset and onset or offset of object motion. In the case of asynchronous occurrences, the relative closeness in time between auditory and visual elements clearly distinguished them from cases where mothers named a static object. For instance, if a mother named a static object and moved it soon after, the occurrence was coded as asynchronous if the onset of movement occurred immediately after word onset (within 500–700 msec). If the onset of movement occurred relatively later than word onset, it was coded as static naming (*st,* typically after 700 msec).

The temporal properties of a proportion (14%) of the synchronous (*s*) and asynchronous (*a*) target word tokens were further analyzed. In this descriptive analysis, the temporal discrepancy between auditory and visual segments was measured for five synchronous and five asynchronous occurrences, randomly selected from each of 14 episodes that had at least five asynchronous occurrences of naming. The videotapes were played at a slow speed on a professional video player/recorder and a monitor. The utterances were played on the normal audio track while using the timing device in seconds and frames to measure the onset and offset of object motion relative to word onset and offset. The analysis yielded a significant difference between synchronous and asynchronous naming categories. In synchronous occurrences, the words occurred during object motions with only a small discrepancy between the onset of words and object motions

(*M* = 60 msec, *SD* = 20, range = 0–90 msec) and the offset of words and object motions (*M* = 80 msec, *SD* = 70, range = 0–150 msec). In asynchronous occurrences, the auditory and visual elements overlapped only slightly or not at all, resulting in a much greater discrepancy between words and object motions (*M* = 520 msec, *SD* = 100, range = 430–650 msec for the onset; and *M* = 560 msec, *SD* = 100, range = 450–670 msec for the offset). These categories fit within the temporal thresholds for infants' perception of synchrony (within 350 msec; see Lewkowicz, 1996b) versus asynchrony in auditory–visual events.

For each dyad, the proportions of total target word (PTTW) tokens of each type (*s, a, st, iho,* or uncoded occurrences) were derived by dividing the number of target word tokens of each bimodal naming type by the total number of tokens summed across all types (*s, a, st, iho,* and uncoded occurrences). Interobserver reliability was obtained by comparing the proportions of the first observer for each of the bimodal naming types (*s, a, st, iho,* and uncoded) to those from the second observer for these types. The mean Pearson correlation coefficient (*r*) between the two observers' proportions for the bimodal naming types across 6 dyads (25%) was .97 (*SD* = .01). Further, the mean agreement between two observers' coding of maternal naming into five naming types across 6 dyads was .87 (*SD* = .07).

Infants' attention to maternal naming. To examine how infants attended to maternal naming, the dyads were independently coded by one or two observers. Each instance of maternal target word naming was coded for infants' attention under one of five types: looks to the object, looks to the mother, gaze-switch from object to mother, gaze-switch from mother to object, or nonlooks. The temporal window within which a coder differentiated a look to the object versus a look from object to mother was the duration of the word. These categories were mutually exclusive. For example, if an infant switched gaze from the mother to the target object, that occurrence was included only under the gaze-switch from mother to object category, and not in the continuous look to the object category, resulting in independent measures of their frequency. Further, this dichotomous measure was most appropriate because infants never looked more than once at the object during a target word utterance.

We used the number or frequency, and not the length, of looks to preclude mothers' prosodic idiosyncrasies during target word naming such as vowel elongation from confounding our measure of infant attention to naming. The number of looks during maternal naming was a more accurate measure, because length of infants' look or duration (the more typical measure) in our case was confounded by the length of maternal utterance. Cross-modal research suggests that when speech accompanies a moving target object infants attend to it more than in silence (e.g., Johnson, 2005). We obtained the proportion of attention types by dividing the sum total of each type by the total looks and nonlooks across all types. Interobserver reliability was calculated between the two observers' proportions of infant attention. The mean Pearson correlation averaged across 8 dyads (33%) was .96 ($SD = .04$). The mean agreement between two observers' coding of the five attention types across the 8 dyads was .84 ($SD = .04$).

To examine which different maternal naming types correlated with infants' attention to object naming, individual proportions of infants' attention during maternal synchronous, asynchronous, static, naming, and naming during infants' manipulation of objects were calculated. The sum of each infant's looks to the target objects, or mother or elsewhere, or gaze switches from mother to object or object to mother, when the mother named the objects using each naming style (*s, a, st,* or *iho*) was divided by the sum of maternal target words for each naming style to obtain these individual proportions. The proportion of looks to target objects and gaze switches were collapsed to obtain an overall measure of infants' attention to objects during maternal naming.

Test for Learning of Word–Object Relations

Stimuli. Videotapes were made of the objects (Figure 1), the Martian and the raccoon, and the words, *Gow* or *Chi,* spoken in a female voice in infant-directed speech. An actor held and moved the objects simultaneously while uttering each word in synchrony with the motions, simulating an act of showing and naming the objects to the infants. Only a portion of the actor's forearm and the objects were visible on film. The objects were randomly moved back and forth or laterally, similar to Gogate and Bahrick's (1998) stimuli. The movements occurred in an irregular temporal pattern, rhythm, and tempo. The words occurred at the rate of 26 tokens per minute and in temporal synchrony with the objects' motions. These videos were edited to obtain four 6-min displays of the Martian and the raccoon, each paired with the words *Gow* and *Chi*.

Procedure. Immediately following the play episode, infants received the intermodal word-mapping test to see if they had learned the word–object relations from their mothers. Infants were seated in an infant seat, in an adjacent dimly lit room, in front of and equidistant from two 20-in. television monitors (Sony, KV20M10) placed 25 in. apart. The visual displays of the Martian and the raccoon were played synchronously from two videotape recorders (Panasonic, AG 7750), using an editing controller (Panasonic, AGA 750). The words were played at a sound pressure level of 65 dB to 70 dB via a central speaker (Trans Audio) located under a large table on which the infant seat was placed. Infants received two blocks of six 10-sec trials (Blocks 1 and 2), with no interval between them. They were required to complete at least 10 out of the 12 trials. On each trial, they viewed the two objects moving identically (in tandem) and side-by-side on the monitors, in temporal synchrony with each other. The motions of both objects were synchronized with utterances of Chi or Gow, on each trial. The lateral positions of the two object displays remained the same for the first block of six trials and were switched for the second block of trials. The lateral display position was counterbalanced across participants so that half the infants (n = 12) received the Martian on the right monitor and the raccoon on the left on the first block of trials, and the other half received the raccoon on the right monitor and the Martian on the left on the first block of trials. Infants received the two words in a random order that was repeated across blocks. In each block of trials, infants received the two words an equal number of times. Infants' attention was centered before each trial using a pair of manually operated bells located centrally between the two monitors. To elicit anticipatory or first looks to the word-matched objects on each trial, two occurrences of the word (*Gow* or *Chi*) appeared approximately 4 sec before the object displays and then continued to occur during the displays.

When the objects appeared, infants were expected to look first at the object display that matched the word, *Gow* or *Chi,* if they had learned the word–object relation from their mother during play. In two prior studies that used an intermodal preference procedure (Gogate & Bahrick, 2001; Hirsh-Pasek, & Golinkoff, 1996), the first look was a more sensitive measure of word mapping than the length of look to the word-matched object display.

One or two trained observers, hidden from the infants by a vertical screen that surrounded the monitors on the sides, blind to the lateral positions of the object displays, recorded infants' visual fixations to the right and left video displays from one of two peepholes located above the monitors. To ensure observers' blindness to the objects' lateral positions, a third experimenter set up the videotapes prior to test and played the videos to the infants only after the observers positioned themselves behind the monitors. When the objects appeared, and the infants visually fixated on them, the observers pressed one of two "silent" buttons on a button box to record infants' looks to the right or left monitors. Observers could not see each other's button presses in the dark. The button boxes were connected to a computer and a printer located in an adjacent room that printed the data online.

Measure of infants' learning of word–object relations. The proportion of infants' first looks (PFLs) to the word-matched display was calculated by dividing the number of trials on which infants looked first to the word-matched object by the total number of (10–12) trials (Blocks 1 and 2), and for each block of six trials, where the lateral positions of the objects remained the same. Interobserver reliability was calculated between two observers' PFLs per trial for 11 out of 24 infants. The mean Pearson correlation between observers' first looks was .95 (SD = .05).

RESULTS

Maternal Multimodal (Auditory–Visual) Naming

If temporal synchrony between spoken words and objects' motions[1] highlights novel word–object relations for preverbal infants, then mothers should use it more often than other naming types to teach their infants novel words. To test this hypothesis, the proportion of total target words for the four maternal naming types, *s* (naming synchronous with object motion), *a* (naming asynchronous with object motion), *st* (naming of a static object), and *iho* (naming when the infant holds the object) were subjected to a one-way repeated measures analysis of variance (ANOVA; general linear models procedures weighted the proportions according to the number of target words contributing because some mothers uttered the target words more often than others, range = 15–136). The analysis revealed a highly significant main effect of naming type, $F(3, 69) = 87.91$, $p < .0001$, estimated effect size = .79. Post-hoc analyses (Scheffe's multiple comparison, $p < .05$) revealed that the mothers named the novel objects using a greater proportion of temporal synchrony (*s*) relative to the other coded bimodal naming types (*a, st,* and *iho*; see last two rows of Table 2 for mean proportions and frequencies). A further repeated measures analysis of maternal naming types with temporally synchronous naming versus all other naming types collapsed into a single nonsynchronous category also showed a significant effect of naming types, $F(1, 23) = 24.90$, $p < .0001$, estimated effect size = .52. Analyses of each mother's naming showed that 23 of the 24 mothers used temporal synchrony[2] more often than the other types during target object naming. Only one mother named static objects more often than using other bimodal naming types. Thus, mothers taught words to their 6- to 8-month-old infants primarily using temporally synchronous verbal labels and gestures or "multimodal motherese" (also see Gogate et al., 2000). A secondary analysis of the proportions by naming type and ethnicity (Caucasian American or Hispanic American; one African American mother was excluded) revealed no main effect of ethnicity, $F(1, 21) = .05$, $p = .82$, or an interaction between factors, $F(3, 63) = .035$, $p = .99$. Because Hispanic

[1]Of the target words, 61% were uttered in multiword sentences (similar to Aslin, Woodward, LaMendola, & Bever, 1996; 39% in isolation). Further, 76% of the target word occurrences in multiword sentences were in sentence-final position. The findings lend cross-cultural support for Fernald's (1992) view that, in infant-directed speech, focused words occur regularly in sentence- or phrase-final position where the pitch peaks are also most often exaggerated. However, we found no correlation between these variables and the PFLs (*p*s > .1), suggesting no relation between word-mapping ability and maternal use of target words in multiword utterances or in word-final position. Neither was frequency of maternal naming positively correlated with infants' word-mapping ability (*p*s > .1).

[2]Mothers sometimes used auditory–visual–tactile synchrony ($M = .27$, $SD = .19$). These instances were collapsed with (bimodal) auditory–visual synchrony.

TABLE 2
The Number (Mean Frequency) and Mean Proportions (*SD*) of Target Words Mothers Used to Name Novel Objects as a Function of Bimodal Naming Styles and Infants' Attention

	Maternal Bimodal Naming Types				
Infants' Attention	*Moving-Synchronous Naming (s)*	*Moving-Asynchronous Naming (a)*	*Naming of a Still Object (st)*	*Naming When the Infant Held an Object (iho)*	*Nonsynchronous Object Naming (a, st, iho)*
Looks to mother					
Raw total (*M* frequency)	54 (2.25)	8 (.33)	9 (.38)	12 (.50)	29 (1.21)
M proportions (*SD*)	.04 (.06)	.10 (.18)	.07 (.12)	.20 (.29)	.08 (.12)
Looks to object					
Raw total (*M* frequency)	570 (23.8)	45 (1.9)	93 (3.9)	47 (2.0)	185 (7.71)
M proportions (*SD*)	.59 (.21)	.60 (.39)	.48 (.27)	.49 (.35)	.56 (.25)
Gaze-switch from mother to object					
Raw total (*M* frequency)	64 (2.7)	7 (0.3)	19 (0.8)	2 (.08)	28 (1.17)
M proportions (*SD*)	.09 (.08)	.10 (.23)	.12 (.16)	.10 (.04)	.08 (.08)
Gaze-switch from object to mother					
Raw total (*M* frequency)	50 (2.1)	3 (.13)	5 (.21)	0 (0.0)	8 (.33)
M proportions (*SD*)	.05 (.06)	.03 (.06)	.05 (.09)	.00 (.00)	.02 (.04)
Looks elsewhere					
Raw total (*M* frequency)	166 (1.96)	19 (.79)	47 (1.96)	11 (.46)	77 (3.21)
M proportions (*SD*)	.23 (.21)	.17 (.22)	.28 (.23)	.21 (.19)	.22 (.18)
Grand total[a] (*M* frequency)	904 (37.67)	82 (3.42)	173 (7.21)	72 (3.0)	327 (14.18)
M proportions (*SD*)	0.63 (0.19)[b]	0.07 (0.07)[c]	0.14 (0.11)[c]	0.07 (0.10)[c]	.28 (.16)[c]

[a]Out of 1,339 target word tokens, 97 (.085) were not codable for maternal naming, and 11 additional occurrences (.01) were not codable for infant attention.
[b,c]Scheffe's multiple comparison, two-tailed $p < .05$.

TABLE 3
Regression Coefficients (R^2) of Infants' Attention to Objects on Maternal Bimodal Naming Types (Asynchronous-*a*, Synchronous-*s*, Static-*st* Naming, and Naming When Infant Held and Explored the Object -*iho*)

	a	*s*	*st*	*iho*
1. Attention during asynchronous naming (*a*)	.01	—	—	—
2. Attention during synchronous naming (*s*)	—	.29*	—	—
3. Attention during naming of static objects (*st*)	—	—	.005	—
4. Attention when infants held object—during maternal naming (*iho*)	—	—	—	.11

*$p < .01$.

American and Caucasian American mothers used similar naming styles for their infants, subsequent analyses were collapsed across ethnicity.

Infants' Attention to Maternal Multimodal Naming

The next analyses focused on infants' attention to maternal naming of the target words. A distributional analysis showed that infants attended to either the mother or the object 76% of the time, and looked elsewhere (23%) during object naming. When they attended to mother or object, infants looked to the object more frequently (59%) than they switched their gaze from mother to object (8%), or from object to mother (5%) during maternal naming. They also looked at their mother (4%) without switching direction of gaze. A small portion of naming occurrences (1%) was not coded because one mother obscured her infant's face from the camera's view during a portion of the play episode. Consistent with the results of prior studies (Morales et al., 1998; Silven, 2001), the preverbal infants showed an ability to switch eye gaze from mother to object when mothers named novel objects within the infants' visual field.

To test the hypothesis that mothers' temporally synchronous object naming predicts infants' attention to object naming, the proportions of infants' looks to target objects during naming and gaze switches combined (from Table 2) during synchronous (*s*), asynchronous (*a*), static (*st*), and when infants held an object (*iho*) were regressed on the corresponding proportions of maternal naming types. Infants' gaze switches from object to mother were included in this object attention analysis simply because the infant looked at the object during word utterance. A linear regression of infants' attention during synchronous naming on maternal use of temporal synchrony alone was statistically significant (Table 3), $F(1, 23) = 8.8$, $p = .007$. Similar regression analyses of object attention during other naming types yielded no significant effects (ps > .1). These results suggest that maternal

TABLE 4
Number of Infants ($N = 24$) Evidencing Each Attention Type During Maternal Temporally Synchronous Naming (s) Versus Other Naming Types (Asynchronous-*a*, Static-*st*, or Naming When the Infant Held an Object-*iho*)

Attention Type	*s*	*a*	*st*	*iho*
Gaze switch Mother to object	17	7 $z = 3.175$**	13 $z = 1.21$	10 $z = 2.13$*
Object to mother	11	4 $z = 2.296$*	6 $z = 1.54$	8 $z = .89$
Looks to object	23	21 $z = 1.056$	22 $z = .598$	22 $z = .598$

*$p < .05$. **$p < .01$, z test of infants' attention success during maternal temporally synchronous versus other naming types.

temporal synchrony during naming is a significant predictor of preverbal infants' attention to objects.

To test the hypothesis that more infants switch eye gaze from mother to object during mothers' temporally synchronous naming, we compared infants' attention success (an attention type or its absence treated as a dichotomous variable) across maternal naming types. A greater number of infants (17 of 24) switched gaze from mother to object during temporally synchronous maternal naming than during other naming types (but number of infants who gaze-switched during synchronous vs. static object naming did not differ statistically; see Table 4).

Infants' Learning of Word–Object Relations

Maternal temporal synchrony and infants' word mapping. If young infants' word mapping is a function of mothers' temporally synchronous naming, then mothers who use temporal synchrony more often should have infants who learn word–object relations better. To test this hypothesis, first, we examined the correlation between the proportion of maternal temporal synchrony and infants' PFL to the word-matched object display on test for Blocks 1 and 2 combined. A Pearson correlation between these factors was significant for the PFLs of Blocks 1 and 2, $r(23) = .42$, $p = .038$. In contrast, the correlation between the proportion of maternal nonsynchronous naming and the PFLs (Table 2, column 5) yielded no significant results (ps > .1). Thus, mothers who used temporal synchrony more often had infants who showed better learning of word–object relations.

Infants' attention to temporally synchronous naming and word mapping. If word mapping is a function of infants' ability to attend to mothers' temporally

TABLE 5
Pearson Correlations (*r*) Between Infants' Proportion of First Looks to the Word-Matched Objects (PFLs) on the Word Mapping Test and the Proportion of Looks to Target Objects During Various Types of Maternal Object Naming in the Play Episode

	PFL (Blocks 1 & 2)
Attention to maternal synchronous naming	
1. Looks to object	–.070
2. Looks from mother to object	.467*
3. Looks from object to mother	.093
Attention to maternal asynchronous naming	
1. Looks to object	–.248
2. Looks from mother to object	.147
3. Looks from object to mother	.216
Attention to maternal static naming	
1. Looks to object	–.124
2. Looks from mother to object	.294
3. Looks from object to mother	.343
Attention to maternal naming when infant holds an object	
1. Looks to object	–.182
2. Looks from mother to object	–.025
3. Looks from object to mother	–.070
Total looks during naming	.24

*$p < .05$; two-tailed.

synchronous naming, then infants who attend more often to their mothers' temporally synchronous naming should learn word–object relations better. To test this hypothesis, Pearson correlations were computed between the different proportions of infants' attention types when mothers used each naming type (from Table 2) and infants' PFL on Blocks 1 and 2 taken together. The results indicated that all types of attention to synchronous naming were not associated with better word mapping (Table 5). No positive correlation was observed between infants' total proportion of looks to target objects during maternal naming regardless of naming type (total looks during naming) and infants' PFLs (bottom row Table 5). The PFL was positively associated only with infants' gaze-switching from mother to object when she used temporal synchrony during the play episode (Table 5). In addition, no positive correlation was found between infants' age (in days) and their ability to gaze-switch from mother to object during her synchronous naming, Pearson $r(23) = .16$, $p = .45$.

To examine whether infants' gaze-switching from mother to object during play influenced their learning of word–object relations (i.e., to be able to draw a causal rather than a correlational conclusion from the results), the 24 infants were assigned to one of two randomized[3] groups ($n = 12$) based on their gaze-switching ability. Infants were assigned to a high gaze-switching group (range = .07–.24, $M = .13$, $SD = .06$) if their proportion of gaze-switching from mother to object was greater than the median (.06) or to a low gaze-switching group (range = .0–.06, $M = .025$, $SD = .025$) if their proportion of gaze-switching from mother to object was less than or equal to the median. To compare word-mapping performance of the two groups of infants, and to analyze whether age was a confounding factor in infants' word-mapping ability, the PFLs for Blocks 1 and 2 and for each block were subjected to separate one-way ANOVAs as a function of group using infants' age as a covariate. The analyses of the proportions for Block 1 alone and Blocks 1 and 2 taken together revealed a significant main effect of group (Figure 2), $F(1, 21) = 4.9$, $p = .038$ (estimated effect size = .19); $F(1, 21) = 6.21$, $p = .02$ (estimated effect size = .21); but no effect of age, $F(1, 21) = .49$, $p = .49$; $F(1, 21) = 1.87$, $p = .19$, respectively. An analysis of the first looks for Block 2 revealed no group or age effects ($p > .1$).[4] In summary, success in word mapping was influenced to a great degree by infants' emerging ability to switch eye gaze from mother to object during maternal naming. However, the variance in word mapping could not be explained in terms of age differences between 6 and 8 months.

Multiple factors in word mapping. Next, we tested the hypothesis that infants' word mapping is a function of multiple factors such as maternal use of temporal synchrony, infants' attention to objects during temporally synchronous naming, infants' ability to switch gaze from mother to object, and ability to attend to the test displays.[5] Infants' PFLs to the word-matched object were submitted to analyses of covariance (ANCOVAs),[6] using as covariates the proportion of

---

[3]The term *randomized group* or *block* is used here to mean that in repeated samples the median is not fixed but is based on the proportion of attention observed in each sample (Edwards, 1985).

[4]*T* test versus chance (50%), performed using the PFLs for Blocks 1 and 2, for the high and low gaze-switching groups, yielded above-chance performance of the high gaze-switching group ($M = .59$, $SD = .11$, $t = 2.80$, $p = .02$). For this group, 10 of 12 infants' PFLs were above chance, 1 performed at chance, and 1 below chance. Only 4 of the 12 infants of the low gaze-switching group performed above chance. Further, an ANOVA of the PFLs by ethnicity (Caucasian, Hispanic) and gaze-switching ability for 23 dyads (1 dyad was African American) yielded no significant effect of ethnicity, $F(1, 19) = .28$, $p = .64$, or interactions, $F(1, 19) = 2.08$, $p = .17$.

[5]We tested whether infants' success in word mapping was in part task specific or based on interest in participating, indexed by the infants' visual attention to test displays (total looking /120 sec; range = .28–1.0, $M = .83$, $SD = .17$). Infants' PFLs for Blocks 1 and 2, $r(23) = .48$, $p = .018$, and for Block 1, Pearson $r(23) = .483$, $p = .017$, were positively correlated with infants' attention to test displays. Thus, infants' word-mapping success was in part task specific or based on infants' interest in participating or both. We included this factor in the subsequent multiple factors analyses.

[6]The ANCOVA served two purposes. Inclusion of the randomized block design allowed assessment of a causal effect of gaze-switching on word mapping. Further, similar to a regression analysis, "comparisons among treatments (covariates) in ANCOVA were tantamount to comparisons among intercepts of regression equations in which the dependent variable is regressed on the covariate" (Pedhazur, 1982, p. 494).

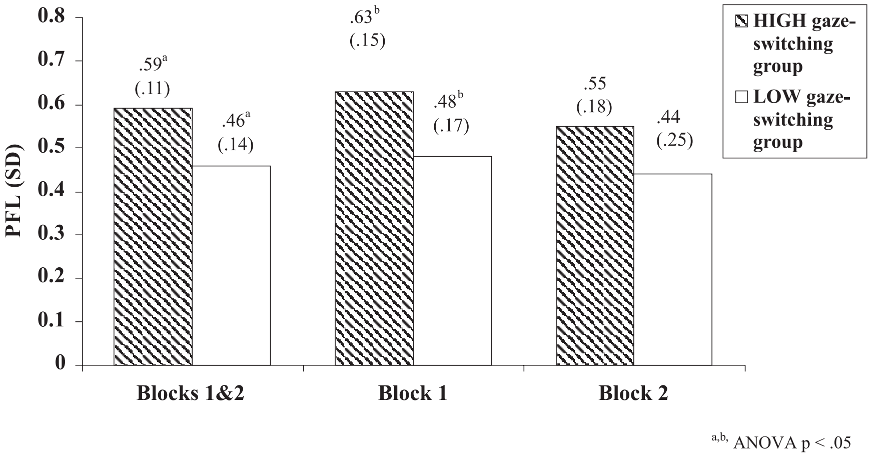

FIGURE 2 The proportions of first looks (PFLs and standard deviations) to the word-matched object displays on the word-mapping test as a function of infants' gaze-switching ability.

temporally synchronous maternal bimodal naming during the play episodes (PTTW-*s*), the proportion of infants' looks to objects when mothers used temporal synchrony (from Table 2), and the proportion of infants' total looking to the test displays. Infants' ability to switch eye gaze from mother to object during naming was the independent variable with two groups (high and low). The analysis for Blocks 1 and 2 taken together pointed to a role for maternal temporal synchrony, $F(1, 19) = 4.0$, $p = .06$ (estimated effect size = .20), and the proportion of infants' total looking to the test displays, $F(1, 19) = 6.37$, $p = .02$ (estimated effect size = .23), but no effect of infants' looks to the target objects when mothers used temporal synchrony, $F(1, 19) = .90$, $p > .1$. In addition, factoring out the two covariates, maternal temporally synchronous naming, and infants' total looking to the test displays, resulted in a more significant group effect on the PFLs as a function of infants' ability to switch gaze from mother to object during naming, $F(1, 19) = 6.22$, $p = .022$ (estimated effect size = .24). A similar ANCOVA for Block 1 revealed significant effects of the covariates, maternal use of temporal synchrony, $F(1, 19) = 16.7$, $p = .001$ (estimated effect size = .43), and the proportion of infants' total looking to test displays, $F(1, 19) = 10.83$, $p = .004$ (estimated effect size = .34), but not infants' looks to target objects when mothers used temporal synchrony during the play episodes, $F(1, 19) = 1.09$, $p > .1$. Factoring out the covariates, maternal temporally synchronous naming, and the proportion of infants' looking to test displays, resulted in a more significant between-group difference in the PFLs for Block 1 as a function of infants' ability to switch eye gaze from mother to object during her naming, $F(1, 19) = 8.22$, $p = .01$. An analysis for Block 2 revealed no significant effects of the covariates or a main effect of gaze-switching (all *p*s > .25), likely owing to infants being bored. (Further, the proportion of infants' gaze-switching was not significantly correlated with any of the three covariates.) These results support our hypothesis that multiple factors predict word mapping in infancy. These factors include maternal use of temporal synchrony during object naming and infants' attention to the test displays, but not infants' looks to objects alone during mothers' temporally synchronous object naming. Rather, infants' ability to switch eye gaze from mother to object, when mothers named the objects, significantly influenced their learning of the word–object relations, indexed by infants' ability to look first at the word-matched object (Figure 2).

Separate ANCOVAs of the PFLs using as covariates the proportions of other maternal bimodal naming types (*st, a,* and *iho*; see Table 2), the proportion of infant looks to objects when mothers used these naming types during play (Table 2), and the proportion of infants' total attention to test displays, and using gaze-switching from mother to object as the independent variable, yielded no significant effects (all *p*s > .1). The null findings, taken together with the significant effects of temporal synchrony, suggest that no measured maternal bimodal naming type except temporal synchrony was a predictor of infants' word-mapping ability.[7]

DISCUSSION

Between 6 and 8 months, during the transition from non-word-mapping to early word-mapping abilities, some infants learned two word–object relations from their mother during a play episode (cf., Oviatt, 1980; Thomas, Shucard, Ramsay, & Shucard, 1981). This learning was demonstrated in infants' anticipatory or first looks more often to the word-matched objects on an intermodal word-mapping test. Learning of word–object relations was significantly influenced by the ability in some infants to switch eye gaze from mother to object during target object naming. Infants' word-mapping ability was also predicted by the extent of mothers' temporally synchronous naming. These results support the general hypothesis that maternal multimodal naming and infant perceptual-lexical systems are bidirectionally

[7]When infants completed the word-mapping test, mothers were asked to complete the MacArthur Communicative Development Inventory (Infant version; Fenson et al., 1994). A raw score of each infant's receptive vocabulary for toy object names was obtained from 53 possible words ($M = 2.73$, $SD = 3.83$, range = 0–12). Infants' attention to the target objects during play was positively correlated with mothers' reports of infants' knowledge of toy names ($r = .47$, $p = .028$), but not all words. Infants, who attend to objects more often, inside or outside the laboratory, are likely better at learning names for objects.

related. Mothers used temporal synchrony between spoken words and the motions of objects in abundance during naming, perhaps as a by-product of their attempts to gain infants' attention. Reciprocally, infants benefited from their mothers' use of temporal synchrony. More infants switched their gaze from mother to the object during temporally synchronous naming than during other maternal naming types. This manner of gaze-switching, in turn, facilitated the learning of word–object relations (see results of earlier ANCOVA). Thus, the developmental process of word mapping in infancy can be attributed to the ongoing reciprocal or bidirectional interplay between multiple factors that are either extrinsic (e.g., the extent of temporal synchrony in maternal naming) or intrinsic to the organism (e.g., infants' ability to switch eye gaze from the mother to an object being named in temporal synchrony with object motion; also see Kelso, 1997; Lickliter, in press; Thelen & Smith, 1994).

These multiple factors contributed in meaningful ways to the developmental process. Mothers scaffolded their young infants' attention to novel word–object relations by providing intersensory redundancy in the form of temporal synchrony between spoken words and the motions of objects (i.e., multimodal motherese) during play. This manner of scaffolding predicted infant attention to word–object relations during play (Table 3). A greater number of infants switched their eye gaze in the direction of the named object during temporally synchronous maternal naming than during other naming types (Table 4). Thus, maternal temporally synchronous naming is associated with the very type of attention (gaze-switching from mother to object) that infants might need to begin mapping words onto objects. These findings speak to the importance of individual variability in explaining the development process of early word mapping and show that variability is not just noise (Bloom, 1998; Smith & Thelen, 2003). To the contrary, variability in mother–infant interaction accounts for variability in word mapping by preverbal infants. Systemic variations in multiple factors across multiple systems account for developmental change (Kelso, 1997; Lickliter, in press; Smith & Thelen, 2003; Thelen, 1988; Thelen & Smith, 1994), from prelinguistic to beginning word-mapping capabilities (Gogate et al., 2001). They also support, in part, the emergentist coalition view of word comprehension, which predicts that early word learning is facilitated by multiple cues that are in alignment with one another (Hollich, Hirsh-Pasek, & Golinkoff, 2000).

Furthermore, contrary to findings from older infants that follow-in labeling or child-centered utterances (caregiver contingent communications that follow the child's visual line of regard) support word learning (Akhtar et al., 1991; Olsen-Fulero, 1982; Tamis-LeMonda & Bornstein, 1989), the findings reported here suggest that directive utterances play an important role in word mapping (see also Collis, 1977; Harris, Jones, & Grant, 1983). It is possible that each utterance type is useful at different times during development. Early on, between 6 and 8 months, when preverbal infants begin to put together words and referents, mothers direct their infants' attention to the relations using a great deal of temporal synchrony during object naming (Gogate et al., 2000; Messer, 1978; Zukow-Goldring, 1997). As is clear from Table 5, the equivalent of follow-in labeling, mothers' naming when infants explored an object or *iho,* had no impact on word mapping for 6- to 8-month-olds. When infants become more competent at learning word–referent relations with experience, around 9 months and later, mothers scaffold vocabulary development using child-centered utterances (Akhtar et al., 1991; Rollins, 2003). We take this as evidence that during the course of lexical development, mothers' communication directs and is directed by infants' behavior, once again illustrating reciprocal or bidirectional adaptations in mother–infant communication (Studdert-Kennedy, 1991).

Although the learning of word–object relations at 6 to 8 months was evident only under optimal conditions, these results show that at least some preverbal infants have a more sophisticated understanding of their proximal naming environment than was previously thought. For these infants, learning of word–object relations took place in the context of a brief play episode with the mother (1.5 min of exposure to each word–referent relation). Then, in an entirely different context, on the intermodal word-mapping test, the infants demonstrated learning by looking first at the word-matched object display (3-month-olds learn to relate two objects and their natural sounds also during habituation and generalize this learning to objects of a different color and shape on an intermodal preference test; Bahrick, 2002).

How might infants have learned the word–object relations at such an early age in this study? First, infants' familiarity with their own mother's voice and naming style likely promoted early word mapping. Prior studies that have demonstrated word mapping, but not until the end of the first year or early during the second year (Hollich et al., 2000; Oviatt, 1980; Schafer & Plunkett, 1998; Thomas et al., 1981; Werker et al., 1998; Woodward, Markman, & Fitzsimmons, 1994; also see Stager & Werker, 1997; cf. Balaban & Waxman, 1997; Tincoff & Jusczyk, 1999), have not exploited infants' familiarity with maternal naming. Second, "multimodal motherese" was highly salient and relevant to preverbal infants (see Bloom, 1998; also see Gogate & Bahrick, 1998, for infants' longer looking during habituation to synchronous relative to static syllable–object displays). Infants are quite adept at detecting intersensory redundancy (amodal properties) when the same information is presented simultaneously to more than one sensory modality (see reviews by Bahrick & Pickens, 1994; Gogate et al., 2001; Walker-Andrews, 1994; also Bahrick & Lickliter, 2000; Hollich, Newman, & Jusczyk, 2005; Prince & Hollich, 2005). When unique information is presented along with redundant information simultaneously to two modalities, the redundant temporal properties highlight the arbitrary relations between otherwise unrelated patterns of stimulation for infants (Gogate & Bahrick, 1998). In this study, when mothers provided temporal synchrony between otherwise unrelated spoken words and moving objects, the redundant information likely captured infants' attention and highlighted the arbitrary relation. Third, it appears that infants who were able to attend to multiple cues simultaneously—namely the mother, her naming of an object, and the object—indexed by their ability to switch gaze from mother to object, succeeded in pairing the word with the object. As discussed earlier, some studies have shown a positive correlation between gaze-following (which entails gaze-switching from the mother to an object) early in the first year and standardized measures of word comprehension by 12 months or later (Morales et al., 1998; Silven, 2001). The present study links infants' gaze-switching and word-mapping abilities during the same developmental period. This finding is ecologically relevant. As infants learn to sit on their own, they might increasingly encounter naming of objects during *en face* interactions with adults.

This is the first evidence to suggest that infants' perception of intersensory redundancy and their ability to gaze-switch are linked in the service of word learning. To reiterate, more infants switched gaze from mother to object during temporally synchronous naming than during other naming types. This type of gaze-switching alone facilitated word learning (gaze-switching from object to mother did not result in learning). How might these abilities be linked? Recent cross-species studies of cats, rhesus monkeys, and rodents suggest a common neural basis in the superior colliculus for gaze-switching and perception of temporally and spatially aligned auditory–visual stimuli (see review by Stein et al., 2004). Neuronal evidence from these mammalian species suggests that multisensory neurons, predominant in the superior colliculus (and some cortical regions), show response enhancement to simultaneous auditory–visual stimulation compared to visual or auditory stimulation alone (see Stein et al., 2004. for a review). This midbrain region is also a major control center for gaze-shifting behaviors. Owing to this shared neuronal space, on hearing a rustling in the bushes, an animal can shift eye gaze in the direction of the rustling and visually locate its prey during hunting. We speculate that human infants might be exploiting these sophisticated abilities for learning word–object relations. When a mother names and moves an object in temporal synchrony, the infant, on hearing the utterance, shifts gaze in the direction of the moving object. In so doing, the infant visually locates the object and links word with referent. Further examination of the possible link between these abilities in word-mapping contexts might better elucidate the developmental process of word mapping.

The findings support, in part, the thesis that during the development of lexical comprehension, infants first learn that a specific word and an object go together. Only later do infants learn that a word stands for an object, an index of referential ability (Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979; Golinkoff, Mervis, & Hirsh-Pasek, 1994). Thus, knowing that specific words go with specific objects (word mapping) is an important precursor to knowing that the words stand for or are symbols for the objects. Perception of the "goes with" relation appears to be fragile and constrained at first. In our view, during the developmental course of word mapping, a gradual shift occurs in infants' preference from perceptual to referential cues (also see Hollich et al., 2000). Early on, prior to detecting word–object relations on their own, preverbal infants rely heavily on perceptual cues, such as intersensory redundancy between words and dynamic objects (Gogate & Bahrick, 1998, 2001; Gogate et al., 2003). Later, as infants become adept at detecting word–object relations on their own, they rely less on intersensory redundancy but still rely on object motion (Werker et al., 1998). Their perception of word–referent relations becomes more stable and flexible as development progresses. During the second year, infants learn word–object relations under less scaffolded conditions. Thus, 12-month-old infants link a familiar label with a category of perceptually similar three-dimensional static objects (Waxman & Markow, 1995). They use adults' pointing or eye gaze to objects or their depictions (e.g., pictures in books) to learn word–referent relations. Relatively freed from mnemonic constraints, these infants have begun to use words to stand for objects. Compared to their younger counterparts, 24-month-olds rely more on referential cues, such as a speaker's intended direction of gaze, to learn word–referent relations (Baldwin et al., 1996; Moore, Angelopoulos, & Bennett, 1999; see also Hollich et al., 2000). Intrinsic development of the organism coacts with environmental experience, permitting learning that is less constrained by systemic and contextual factors, eventually giving way to referential capability during the second year.

In conclusion, these findings and those of others support a domain-general intersensory perceptual basis for early lexical learning. In this study, some preverbal infants learned arbitrary word–object relations (in the language domain). Others have shown that in the first year, infants detect arbitrary visual-acoustic relations in nonlanguage stimuli, such as the color of an object and the pitch of its sound on impact against a surface (Bahrick, 1994), diagonal red or green lines, and a male or female voice (Slater et al., 1997), and arbitrary visual-tactual relations such as the color pattern and the shape of an object (Hernandez-Reif & Bahrick, 2001). At this early age, intersensory perceptual sensitivities are not likely to be specialized solely for learning word–object relations. Rather, domain-general intersensory perceptual abilities assist in early lexical development. Infants continue to exploit domain-general abilities during the second year. They treat nonwords (buzzes, squeaks, or gestures) as labels for objects (Namy & Waxman, 1998; Woodward & Hoyne, 1999). These findings have led some researchers to argue that infants are not truly referential at this point in development; they do not differentiate between words and nonwords in the same manner as older children or adults. They link words to objects in an "associative" manner rather than in a referential manner (Golinkoff et al., 1994; Woodward et al., 1994). Still others hold the view that infants "associate" words and objects only when they map them in the absence of intersensory redundancy, similar to adults (Werker et al., 1998).

It is our view that early domain-general intersensory capabilities and referential abilities fall on a developmental continuum and that intersensory capabilities facilitate infants' mapping of words onto objects and reference. Early intersensory abilities therefore deserve further examination because they may well be a stepping stone to referential capabilities in the infant (also see Bates et al., 1979). Longitudinal examinations of the domain-general sensitivities of infants can yield important clues about how development proceeds gradually from domain-general perceptual abilities, which encompass early lexical mapping, to more language-specific capabilities such as reference. Longitudinal studies of mother–infant interaction could also elucidate the dynamic bidirectional processes of maternal and infant communication systems during ongoing lexical development.

ACKNOWLEDGMENTS

This project was funded by grants to Lakshmi J. Gogate from the Florida International University Foundation, the Thrasher Research Fund (Award No. 02819–1), and a Dean's Research Initiative Award from The College of Medicine, SUNY at Brooklyn. Special thanks are extended to Jilayne Watson for coding maternal naming types, Inna Kremen for coding infants' attention to naming, to the mothers and infants who participated, and to Lorraine Bahrick for her support. Christopher Prince, Arlene Walker-Andrews, Kathy Hirsh-Pasek, George Hollich, Samantha Berkule, and three anonymous reviewers provided valuable comments on an earlier version. We are grateful to the late Esther Thelen for her invaluable discussion on how variability in the data can explain development. A portion of these results was presented at the International Conference for Infant Studies, Chicago, May 2004.

REFERENCES

Akhtar, N., Dunham, F., & Dunham, P. (1991). Directive actions and early vocabulary development: The role of joint attentional focus. *Journal of Child Language, 18,* 41–49.

Aslin, R. N., Woodward, J., LaMendola, N., & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. In J. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to language in early acquisition* (pp. 117–134). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

Bahrick, L. E. (2002). Generalization of learning in three-and-a half-month-old infants on the basis of amodal relations. *Child Development, 73,* 667–681.

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology, 36,* 190–201.

Bahrick, L. E., & Pickens, J. N. (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: Lawrence Erlbaum Associates, Inc.

Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother–infant and peer–infant interaction. *Child Development, 55,* 1278–1289.

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology, 64,* 3–26.

Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R., Irwin, J. M., Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child Development, 67,* 3135–3153.

Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols: Cognition and communication in infancy.* New York: Academic.

Bloom, L. (1998). Language acquisition in its developmental context. In W. Damon, D. Kuhn, & R.S. Siegler (Eds.), *Handbook of child psychology* (pp. 309–370). New York: Wiley.

Bosch, L., & Sebastian-Galles, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language & Speech, 46,* 217–243.

Butterworth, G. E. (1998). The origins of joint visual attention in infancy: A commentary. *Monographs of the Society for Research in Child Development, 63*(4), 144–166.

Butterworth, G. E., & Jarrett, N. L. M. (1991). What minds have in common is space: Spatial mechanisms for perspective taking in infancy. *British Journal of Developmental Psychology, 9,* 55–72.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development,* 63(4).

Collis, M. (1977). Visual co-ordination and maternal speech. In H. R. Schafer (Ed.), *Studies in mother–infant interaction* (pp. 355–375). London: Academic.

Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech.* Cambridge, MA: Harvard University Press.

D'Entremont, B., Hains, S. M. J., & Muir, D. W. (1997). A demonstration of gaze following in 3- to 6-month-olds. *Infant Behavior and Development, 20,* 569–572.

Edwards, A. L. (1985). The analysis of variance for a randomized group design. In A. L. Edwards (Ed.), *Experimental design in psychological research* (pp. 88–108). New York: Harper & Row.

Fenson, L., Dale, P., Reznick, S., Bates, E., Thal, D., & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5).

Fernald, A. (1992). Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In J. H. Barkow, L. Cosmides, & M. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 391–428). Oxford, UK: Oxford University Press.

Fernald, A., Pinto, J., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science, 9*(3), 72–75.

Garnica, O. (1977). Some prosodic and paralinguistic features of speech to young children. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 63–88). New York: Cambridge University Press.

Gogate, L. J., & Bahrick, L. E. (1998). Intersensory redundancy facilitates learning of arbitrary relations between vowel-sounds and objects in 7-month-olds. *Journal of Experimental Child Psychology, 69,* 133–149.

Gogate, L. J., & Bahrick, L. E. (2001). Intersensory redundancy and seven-month-old infants' memory for syllable-object relations. *Infancy, 2,* 219–231.

Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development, 71,* 876–892.

Gogate, L. J., Bewley, S., Wan, P., Kremen, I., & Chen, L.-F. (2003, June). *Preverbal infants' learning of arbitrary relations between complex syllables and objects.* Poster presented at the Biennial Meeting of the Society for Research in Child Development, Tampa, FL.

Gogate, L. J., Walker-Andrews, A. S., & Bahrick, L. E. (2001). The intersensory origins of word comprehension: An ecological-dynamic systems view. *Developmental Science, 4,* 1–18.

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.

Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language, 21,* 125–155.

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

Harris, M., Jones, D., & Grant, J. (1983). The non-verbal context of mothers' speech to infants. *First Language, 4,* 261–268.

Hernandez-Reif, M., & Bahrick, L. E. (2001). The development of visual-tactual perception of objects: Amodal relations provide the basis for learning arbitrary relations. *Infancy, 2,* 51–72.

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). The intermodal preferential looking paradigm reveals emerging language comprehension. In D. McDaniel, C. McKee, & H. Cairns (Eds.), *Methods for assessing children's syntax* (pp. 105–124). Cambridge, MA: MIT Press.

Hollich, G. J., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development, 65*(3, Serial No. 262).

Hollich, G. J., Newman, R. S., & Jusczyk, P. W. (2005). Infants' use visual information to segment speech in noise. *Child Development, 76,* 598–613.

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

Johnson, S. (2005, April). Intersensory redundancy facilitates abstract rule learning. In S. Johnson (Chair), *Multiple perspectives on multimodal contributions to cognitive development.* Symposium presented to the Society for Research in Child Development, Atlanta, GA.

Kelso, S. (1997). *Dynamic patterns: The self-organization of brain and behavior.* Cambridge, MA: MIT Press.

Lewkowicz, D. J. (1996a). Infants' response to the audible and visible properties of the human face: 1. Role of lexical-syntactic content, temporal synchrony, gender, and manner of speech. *Developmental Psychology, 32,* 347–366.

Lewkowicz, D. J. (1996b). Perception of auditory-visual temporal synchrony in human infants. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1094–1106.

Lickliter, R. (in press). Dynamical systems and psychological science. In B. D. Midgley & E. K. Morris (Eds.), *Modern perspectives on J. R. Kantor and interbehaviorism.* New Haven, CT: Greenwood.

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

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

Moore, C., Angelopoulos, M., & Bennett, P. (1999). Word learning in the context of referential and salience cues. *Developmental Psychology, 35,* 60–68.

Morales, M., Mundy, P., & Rojas, J. (1998). Following the direction of gaze and language development in 6-month-olds. *Infant Behavior and Development, 21,* 373–377.

Namy, L., & Waxman, S. R. (1998). Words and gestures: Infants' interpretation of different forms of symbolic reference. *Child Development, 69*, 295–308.

Nelson, K. (1979). The role of language in infant development. In M. Bornstein & W. Kessen (Eds.), *Psychological development from infancy: Image to intention* (pp. 307–334). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Nelson, K. (1988). Constraints on word learning? *Cognitive Development, 3,* 221–246.

Newson, J., & Newson, E. (1975). Intersubjectivity and the transmission of culture. *Bulletin of the British Psychological Society, 28,* 437–445.

Olsen-Fulero, L. (1982). Style and stability in mothers' conversational behavior: A study of individual differences. *Journal of Child Language, 9,* 543–564.

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

Pedhazur, E. J. (1982). *Multiple regression in behavioral research* (2nd ed.). New York: CBS College Publishing.

Prince, C. G., & Hollich, G. J. (2005). Synching infants with models: A perceptual-level model of infant synchrony detection. *Journal of Cognitive Systems Research, 6,* 205–228.

Rollins, P. R. (2003). Caregivers' contingent comments to 9-month-old infants: Relationships to later language. *Applied Psycholinguistics, 24,* 221–234.

Scaife, M., & Bruner, J. S. (1975). The capacity for joint attention in the infant. *Nature, 253,* 265–266.

Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development, 69,* 309–320.

Silven, M. (2001). Attention in very young infants predicts learning of first words. *Infant Behavior and Development, 24,* 229–237.

Slater, A., Brown, E., & Badenoch, M. (1997). Intermodal perception at birth: Newborn infants' memory for arbitrary auditory-visual pairings. *Early Development and Parenting, 6,* 99–104.

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Science, 7,* 243–348.

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

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388,* 381–382.

Stein, B. E., Jiang, W., & Stanford, T. (2004). Multisensory integration in single neurons of the midbrain. In G. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of intersensory perception* (pp. 243–264). Cambridge, MA: MIT Press.

Studdert-Kennedy, M. (1991). Language development from an evolutionary perspective. In N. Krasnegor, D. M. Rumbaugh, R. Scheifelbush, & M. Studdert-Kennedy (Eds.), *The biological and behavioral determinants of language development* (pp. 5–28). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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.

Tamis-LeMonda, C. S., & Bornstein, M. H. (1989). Habituation and maternal encouragement of attention in infancy as predictors of toddler language, play, and representational competence. *Child Development, 60,* 738–751.

Thelen, E. (1988). On the nature of developing motor systems and the transition from prenatal to postnatal life. In P. Smotherman & S. R. Robinson (Eds.), *Behavior of the fetus* (pp. 207–224). Caldwell, NJ: Telford.

Thelen, E., & Smith, L. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press.

Thomas, D., Shucard, D., Ramsay, D., & Shucard, J. (1981). Semantic comprehension in infancy: A signal detection analysis. *Child Development, 52,* 798–803.

Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science, 10,* 172–175.

Tomasello, M., & Farrar, J. (1986). Joint attention and early language. *Child Development, 57,* 1454–1463.

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

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology, 29*(3), 257–302.

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology, 34,* 1289–1309.

Woodward, A., & Hoyne, K. L. (1999). Infants' learning about words and sounds in relation to objects. *Child Development, 70,* 65–77.

Woodward, A., Markman, E. M., & Fitzsimmons, C. M. (1994). Rapid word learning in 13- to 18-month-olds. *Developmental Psychology, 30,* 553–566.

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

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 & P. Zukow-Goldring (Eds.), *Evolving explanations of development: Ecological approaches to organism–environment systems* (pp. 199–252). Washington, DC: APA.

Zukow-Goldring, P. (2001). Perceiving referring actions: Latino and Euro-American infants and caregivers comprehending speech. In K. L. Nelson, A. Aksu-Koc, & C. Johnson (Eds.), *Children's language* (Vol. 11, pp. 139–165). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.