Developmental Psychology
1990, Vol. 26, No. 5, 780-795

Copyright 1990 by the American Psychological Association, Inc.
0012-1649/90/$00.75

Auditory Evoked Responses to Names for Different Objects: Cross-Modal Processing as a Basis for Infant Language Acquisition

Dennis L. Molfese
Southern Illinois University at Carbondale

Philip A. Morse
Boston University and
New England Rehabilitation Hospital

Cindy J. Peters
Southern Illinois University at Carbondale

Evoked potential techniques were used to study the acquisition of names for different objects in 14-month-old infants. Auditory evoked responses (AERs) were recorded from each infant by scalp electrodes positioned over frontal, temporal, and parietal regions of each hemisphere before and after several days of training in which nonsense consonant–vowel–consonant–vowel bisyllables were used by parents to consistently name novel objects. Analyses of the AERs collected during the posttraining session indicated that two portions of the brain response discriminated a "match" versus a "mismatch" occurring between the auditorily presented names and the objects held by the infants. Specifically, an early occurring portion of the AER recorded from bilaterally placed frontal electrodes and a late occurring response detected at only the left hemisphere electrode sites discriminated between situations when a match occurred between the object and its name versus those in which there was a mismatch between the name and object. No such differences were found in the pretraining AER data. Results are viewed as a preliminary step in the neuropsychological study of word concept development using electrophysiological measures.

The interest of developmental psychologists in the early stages of language perception has increased markedly over the last two decades. For the most part, much of this interest has focused on the infant's perception of speech sounds prior to the development of comprehension for single words or the use of one- and two-word utterances (see Eimas, Miller, & Jusczyk, 1987; Kuhl, 1985; and Morse, 1974, 1979; for reviews of this literature). Other investigations have begun to probe the nature of the older infant's early word meanings (Bloom, Lahey, Hood, Lifter, & Fiess, 1980; Clark, 1983; Retherford, Schwartz, & Chapman, 1981; Snyder, Bates, & Bretherton, 1981). However, only recently have the very beginning stages of the infant's ability to perceive and remember the names for objects and events received direct study (Bates, Bretherton, Snyder, Shore, & Volterra, 1980; Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Kamhi, 1986; Miller & Chapman, 1981). Moreover, virtually nothing is known about the role that the brain plays in the early acquisition of such word meanings. It is these last two points that are of concern in this article.

Support for this work was provided by the National Science Foundation (BNS 8004429 and BNS 8210846) and the Office of Research Development and Administration (2-10947), Southern Illinois University at Carbondale.

We wish to thank Leslie MacGregor for her assistance in testing the infants. The contribution of Teong Hian Thew in assisting the development of the EPACS© brain response analysis programs for the Macintosh microcomputer system is also gratefully acknowledged.

Correspondence concerning this article should be addressed to Dennis L. Molfese, Department of Psychology, Southern Illinois University, Carbondale, Illinois 62901.

Infant Speech Perception and Cross-Modal Matching

Studies of infant speech perception have identified a number of speech perception abilities present during the first 7 months of life (for reviews of this literature, see Eimas et al., 1987; Kuhl, 1985; and Morse & Cowan, 1982). Although these studies of infant speech perception document an impressive array of speech perceptual abilities in early infancy, most of this work has focused on the perception of single syllable contrasts. Monosyllabic perception abilities, although necessary, are hardly sufficient for the development of an understanding of the words spoken by the infant's caregivers. The infant must also be able to (a) discriminate and recognize patterns of speech sounds embedded in and consisting of multisyllabic strings, (b) retain them in memory, and (c) match these auditory patterns with the objects that they identify. This latter ability involves not only general representation/symbolic skills (Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979; Ingram, 1978; Ramsey & Campos, 1978) but also cross-modal matching of associations (Luria, 1973). Unfortunately, studies of the infant's abilities for complex auditory and speech perception/memory and cross-modal matching are considerably fewer than those of the infant's basic monosyllabic skills.

The few studies of the infant's recognition of speech sounds in multisyllabic contexts have indicated that at least by 6 months of age infants possess some abilities in this area (Jusczyk & Thompson, 1978; Trehub, 1973; Goodsitt, Morse, Ver Hoeve, & Cowan, 1984). Although infants in these studies showed discrimination of syllabic contrasts embedded in multisyllabic contexts, the structural redundancy of the multisyllabic context appeared to be an important factor in the infant's recognition of a syllabic target in a complex string (Goodsitt et al.,

1984). Little is known, however, about the infant's auditory discrimination and memory over extended intervals for different sets of multisyllabic combinations as occur in the names for objects.

Studies of cross-modal processing in infants have generally not addressed the following two basic aspects of cross-modal association critical for word recognition and comprehension in human language: (a) the association of auditory with visual modality information, and (b) the arbitrary nature of this association (e.g., in the language that the infant and the adult learn, the visual object "pen" could just as easily be associated with "Kugelschreiber" (German), "plume" (French), "pero" (Serbo-Croatian), or even the auditory event "suitcase."

Many of the studies of infant cross-modal processing have focused on the matching of visual and tactile information (e.g., Meltzoff & Bornton, 1979; Wagner & Sakovits, 1986). Those studies that have addressed visual–auditory associations have generally studied what Wagner and his colleagues (Wagner & Sakovits, 1986; Wagner, Winner, Cicchetti, & Gardner, 1981) refer to as *amodal* or *metaphorical* associations. For example, Spelke (1976) observed that infants can match auditory and visual events on the basis of rhythmical patterns, and Lewkowicz and Turkewitz (1980) reported that neonates are able to match loudness to brightness. Wagner et al. (1981) further demonstrated with a preferential looking procedure that infants averaging 11 months of age match (a) a broken line to a pulsing tone and a continuous line to a continuous tone, (b) a jagged circle to a pulsing tone and a smooth circle to a continuous tone, and (c) an upward arrow to an ascending tone and a downward arrow to a descending tone. Although these studies of cross-modal matching indicate that prelinguistic infants can relate a number of amodal features of auditory and visual information, these are not the arbitrary relationships between auditory and visual information of which words are made. Knowledge of these amodal types of cross-modal relationships may be available at birth (e.g., Lewkowicz & Turkewitz, 1980; Meltzoff & Bornton, 1979). In other cases, differing amounts of familiarity or experience may be necessary for the infant to demonstrate knowledge of these relationships in a particular behavioral paradigm (Wagner & Sakovits, 1986) or to acquire this amodal knowledge. In contrast, knowledge of the arbitrary auditory–visual relationships underlying the comprehension of words is only acquired through experience.

Training and Cross-Modal Matching

The role of repeated cross-modal matching experiences in the infant's acquisition of auditory names for objects is apparent from observing parents' naming activities with their young children. Parents generally name objects to which their infants are attending by a variety of repetitive and stylized verbal and manual gestures (Benedict, 1975; Murphy, 1978; Stevenson, Leavitt, Roach, & Chapman, 1986). However, with the exception of Oviatt's (1980) study, no experimental work has been conducted to investigate the training effects of this method of teaching infants the names of objects. Experience consisted of a 3-min exploration of a novel object (rabbit), followed by the experimenter and mother naming the object approximately 24 times. After a 3-min distractor period with other toys, the infant was tested immediately and again 15 min later for comprehension of the target name (e.g., "Where's the rabbit?") versus nonsense names (e.g., "Where's the kawlow?") versus known names (e.g., "Where's the book?"). Oviatt noted that although 9- to 11-month-olds showed little receptive learning under these conditions, half of the 12- to 14-month-olds and 80% of the 15- to 17-month-olds reliably recognized the object after both the short (3 min) and long (15 min) distraction periods. Oviatt's interesting study suggests a number of important questions about the nature of training and experience effects on the early comprehension of names. Our study was designed to address these issues from methodological as well as neuropsychological perspectives.

The following methodological issues were considered. First, whereas in Oviatt's study very short-term effects of training were demonstrated, in our study we examined whether this type of cross-modal training had any long-term effects when carried out over a longer training period. Second, only one training object and name were used in Oviatt's study. In our study, two different objects and nonsense names were used with each infant. We counterbalanced the object–name relationships across infants to control specific auditory and visual confounding both within and across infants. Third, Oviatt presented each infant with only a single choice object during the testing period, which could have biased the infant to choose the only object available in the test situation. In our study, infants were presented with two objects and their names in a match/mismatch situation. Fourth, the visual stimulus used in the Oviatt study was interesting and dynamic (a rabbit). Although this helped to maximize the infants' attention during the study, similar comprehension and learning results may not obtain for less interesting objects in the child's experience. In our study, we used two wooden objects that were probably about as uninteresting as Oviatt's stimulus was interesting to infants in this age range.

Auditory Evoked Responses

In addition to extending our knowledge of cross-modal matching beyond Oviatt's pioneering work, we sought to investigate aspects of the neuropsychology associated with the infant's early word comprehension and learning. The procedures used in this study involved the recording of auditory evoked responses (AER) from scalp electrodes placed over areas of the left and right hemispheres. The AER is a synchronized portion of the ongoing EEG pattern that is detectable at the scalp and that occurs immediately in response to some sound (Callaway, Tueting, & Koslow, 1978; Rockstroh, Elbert, Birbaumer, & Lutzenberger, 1982). The AER is believed to reflect changes in brain activity over time as reflected by changes in the amplitude or height of the waveform at different points in its time course. Because of its time-locked relation to the evoking stimulus, the AER has been demonstrated to reflect both general and specific aspects of the evoking stimulus and the infant's perceptions and decisions regarding it (Molfese, 1983; Molfese & Betz, 1988; Molfese & Molfese, 1979a, 1979b, 1980, 1985; Nelson & Salapatek, 1986; Ruchkin, Sutton, Munson, Silver, & Macar, 1981). It is this time-locking feature that enables researchers to

identify portions of the brain's electrical response that occur while the infant's attention is focused on some discrete event.

The evoked potential technique would appear to be particularly well suited for the neuropsychological study of the infant's acquisition of word comprehension, given its previous successes in investigations of infant speech perception (Molfese, 1972; Molfese & Molfese, 1979a, 1980, 1985; see Molfese & Betz, 1988, for a review of this literature) and of semantic comprehension in older children and adults (Brown, Marsh, & Smith, 1979; Chapman, McCrary, Bragdon, & Chapman, 1979; Molfese, 1979; Molfese, Morris, & Romski, 1990). Indeed, two recent articles indicated that such procedures may be productive in attempts to study the developmental neuropsychology of early word comprehension (Molfese, 1989; Molfese, 1990). In one study designed to determine whether young infants could discriminate known from unknown words, Molfese (1989) recorded auditory evoked responses from frontal, temporal, and parietal scalp locations over the left and right hemispheres of 10 infants, 14 months in age, who listened to a series of words, half of which were determined to be known to the infants (based on behavioral testing and parental report) and half of which were believed not to be known to the infant. Analyses of the AER data isolated three regions of the evoked potential waveform that discriminated known from unknown words in this population. Initially, AER activity across both hemispheres (with the exception of the right parietal region) between 30 and 220 ms following stimulus onset discriminated between known and unknown words. This effect could be seen as a positive peak for the known words and a negative peak in this same region for the unknown words. This activity was followed shortly by a large positive to negative change in amplitude between 270 and 380 ms across all electrode sites for both the left and right hemispheres that was larger for the known than for the unknown words. Finally, a late negative peak between 380 and 500 ms that was detected only by electrodes placed over the left and right parietal regions was larger for the known than for the unknown words. When only familiar versus novel nonsense sounds were presented, however, no such effects were found. In a subsequent study with 16-month-old infants (Molfese, 1990), similar differences were found in a different group of infants in response to known and unknown words. Thus, there are some indications from studies with young infants that support the notion that AERs can be used to successfully discriminate between words that infants do and do not understand.

Studies of receptive language abilities in brain-damaged adult populations provide some suggestions about the possible localization of the infant's comprehension and learning of semantic knowledge. In general, patients with left hemisphere compromise have been found to be particularly impaired in language comprehension tasks (cf. Riedel, 1981, for a review of the relevant literature). On several measures of lexical and complex language comprehension, patients with more posterior damage (Wernicke's aphasics) tend to have more difficulty than Broca's aphasics (anterior damage). However, studies of phonemic perception (e.g., Baker, Blumstein, & Goodglass, 1981) and semantic processing of single words (Gainotti, Caltagirone, & Ibba, 1976; Pizzamiglio & Appicciafuoco, 1971) generally indicate impaired performance independent of left hemisphere locus. Riedel (1981), in her review of the auditory comprehension literature, even suggested that several sources of neuropsychological data support the inference of some degree of bilateral representation of semantic knowledge.

On the basis of these auditory comprehension findings in adult brain-injured patients, one would predict in this study that the most likely outcome would be for the matching of auditory/visual information to be most evident when recorded from sites over the left hemisphere. Pronounced anterior/posterior differences would not be expected, on the basis of the adult aphasia literature. Furthermore, there is some suggestion that bilateral effects might also be observed. Although these predictions are based on adult findings, studies of speech perception using AERs have demonstrated that the infant's brain responds to speech contrasts in a manner similar to the adult's brain (Molfese & Betz, 1988; Molfese & Molfese, 1979a, 1980, 1985). In addition, evidence of a bilateral effect is supported by the few studies of the infant's processing of semantic information that are currently available (Molfese, 1989, 1990).

Purpose of This Study

In sum, our study had several purposes. First, the study's counterbalanced design permitted the direct investigation of the infant's matching versus mismatching abilities for auditory–visual information. Second, the present study served to extend the findings of Oviatt's work by assessing longer term training effects with more than one training item and with less captivating stimuli. Third, it sought to determine whether electrophysiological procedures involving auditory evoked responses could be used to identify the emergence of general associations between auditory and visual stimuli. However, because the trained associations between the specific object names were counterbalanced across the different objects across children, no direct assessment of specific and particular associations could be assessed at the group level. Instead, this study attempted to identify a more general level of processing that would reflect whether any general associations or effects might emerge between auditory and visual stimuli that had been paired together versus those that had not. Fourth, given the ability of the AER procedures to allow spatial analyses of processes, this study attempted to identify the general brain regions involved in the acquisition of this association, as well as the manner and the order in which these regions responded to the stimuli. This point was addressed by placing electrodes over regions of both hemispheres. By comparing AERs collected from over the left hemisphere versus those collected from over the right hemisphere, conclusions could be reached concerning the differential roles that the hemispheres might play in the early acquisition of meaning. Moreover, by comparing AER activity recorded from various regions within a hemisphere, a determination could be made regarding the contribution of frontal, temporal, and parietal regions to this process. Fifth, because the AER represents a change in electrical activity over time, it was believed that this signal could be used to identify the order in which events occurred in the brain. Changes in different portions of the AER that occurred at different times following the onset of stimulation would indicate which events were processed earlier or later.

The Oviatt findings concerning age differences in processing

fit well with those reported by Thomas, Campos, Shucard, Ramsay, and Shucard (1981) in which 13-month-olds were reliably found to visually direct their eye fixations to objects in response to maternal cuing, whereas 11-month-olds could not. Given that both training and responding occurred markedly better in older infants (beyond 12 months of age), a decision was made in the present study to use a somewhat older population than those used by either Oviatt or Thomas et al. in order to enhance the likelihood that the infants would be capable of identifying relationships between words and objects.

Method

Subjects

By using county birth records, letters describing the research project were sent to the parents of prospective subjects, and a follow-up telephone conversation confirmed their willingness to participate in the study. In addition, some subjects' parents heard about the project and volunteered to participate.

In all, 25 infants were scheduled for testing over a 4-month period. However, the data from 11 infants were not included in these analyses because of a variety of problems. Of these 11 infants, 3 failed to return for the posttraining test session because of illness, 5 children would not allow electrodes to be applied to their heads during the initial test session, 1 infant failed to complete all of the final test session, the electrode for one site became detached repeatedly during testing for another child, which resulted in a decision to drop that child's data from the pool, and 1 child's data was lost because of experimenter error.

Fourteen infants (7 girls and 7 boys) did successfully participate in all phases of this experiment. The mean age was 14.72 months (SD = .63). Although infants were not screened on the basis of parental hand preferences, handedness questionnaires administered to all parents (Edinburgh Inventory; Oldfield, 1971) indicated that both parents of each infant were strongly right-handed (group mean laterality quotient = .67, SD = .19, range = +0.42 to +1.00, where +1.00 indicates that all tasks are performed with the right hand and −1.00 indicates that all tasks are performed with the left hand). Given the general belief of a link between handedness, genetic factors, and left hemisphere function for language processes (Bryden, 1982; Harris, 1988), the strong right-hand preferences of both parents would suggest that this group of infants were homogeneous in that respect.

Stimuli

Two consonant–vowel–consonant–vowel (CVCV) bisyllabic auditory stimuli were used in this study, "gibu" and "bidu." The CVCV stimuli were produced by an adult male speaker with a general American accent using a flat intonation. Recorded tokens of these stimuli were digitized using an APPLE IIe microcomputer and Mountain Hardware analogue-to-digital converter and edited to a duration of 438 ms with peak stimulus intensities equated. The CVCV stimulus tokens were then recorded on audiotape by using a 5-KHz low-pass filter. The stimulus tape contained a total of 160 presentations of the two CVCV bisyllables, in blocked sequences with each block consisting of five repetitions of the same CVCV. In this manner, 32 blocks of the two stimuli were arranged in a random order so that 80 repetitions of each CVCV bisyllable were recorded on the tape. The interstimulus interval within a block of 5 presentations varied randomly in duration and ranged from 2.0 to 3.0 s. The interblock interval varied randomly between 4.5 and 6.0 s. Both of these steps were taken to reduce the likelihood of expectation, habituation, and baseline shift effects on the AERs. The same tape was used in both electrophysiological phases of this study.

The two stimulus objects used in the present study consisted of (a) a squared dowel rod that measured 4 in. in length, with ⅛-in. sides, and (b) a 1-in. diameter cylinder that measured 2 in. in length.

Procedure

Pretraining electrophysiological test. Parents brought their infant to the laboratory for a pretest session 1 day prior to the beginning of the behavioral training session. Scalp recording electrodes were placed over the left and right sides of the head at the following locations: T3 and T4 of the Ten–Twenty System (Jasper, 1958), midway between the external meatus of the left ear and Fz (FL), midway between the right external meatus and Fz (FR), midway between the left external meatus and Pz (PL), and over the right side of the head midway between the right ear's external meatus and Pz. Thus, these electrode placements were over the left frontal, temporal, and parietal areas of the brain and the corresponding areas of the right hemisphere.[1] Additional electrodes were placed on the forehead supraorbitally and canthal to the right eye. The electrical activity recorded from these scalp electrode positions was referred to linked ears (A1, A2). Electrode impedances were under 5 kOhms and did not vary more than 1 kOhm between electrode sites as indicated by measurements before and after testing. The gain settings of the modified Tektronix differential amplifiers were placed at 80,000, and filters were flat between 0.1 Hz and 35 Hz.

The infant was seated in the parent's lap throughout the test session. The nonsense CVCV bisyllable auditory stimuli, "gibu" and "bidu," were presented in 5-trial blocks of the same stimulus (for a total of 160 stimulus presentations) through a speaker positioned over the midline and 1 meter above the infant's head. A line suspended from the midpoint of the speaker provided researchers with a reference point so that the infant could be maintained under the speaker's midline.[2] Stimulus loudness levels measured at the infant's ear were set at 75 dB SPL.

---

[1] The frontal and parietal electrode sites used in this study were chosen instead of the more standard 10–20 sites of Jasper (1958) for a number of reasons. First, prior testing had demonstrated to us that the more extended time required for all of the additional scalp measurements necessary for correct electrode placement using the 10–20 system were not tolerated well by infants in this age range. The present electrode arrangement met more effectively our need to place electrodes over the left and right frontal and parietal regions in as short a time period as possible. Second, the use of the 10–20 adult placement system described by the Committee of Clinical Examination in Electroencephalography (Jasper, 1958) did not seem warranted with this infant population because the relationship between electrode placements and the underlying topological structures of the brain does not correspond to that found with infants (Blume, Buza, & Okazaki, 1974).

[2] The overhead speaker arrangement was used instead of front- or side-positioned speakers to maintain a relatively constant distance from each ear to the speaker. Given that systematic differences in such distances could produce hemisphere effects that are simply an artifact of speaker–interaural distances, this concern is not a trivial one. The overhead speaker permitted head movements to the left and right that would not change the distance from either ear to the speaker, unlike speakers placed in a horizontal plane to the infant's ears. Although earphones would have been preferable for stimulus presentation, we found that the infants would not tolerate the earphones on their heads. Furthermore, the earphone cups would have masked the left and right temporal electrode sites, making placement of electrodes in these positions impossible. In addition, prior testing with adults and children indicated that the electrical signal for the acoustic stimuli generated through the earphones interfered with and masked the AER signal.

Stimulus presentation occurred when the infant was in a reasonably quiet awake state, based on continuous monitoring of the infant's ongoing EEG and EMG, as well as on behavioral observation. During periods of motor activity, stimulus presentation was suspended and later resumed when the infant's motor activity declined. Individual auditory evoked responses were recorded on-line to each auditory stimulus by using a Metaresearch Benchtop 2048 unit to digitize the electrophysiological signals for input to a Macintosh Plus microcomputer.

Stimulus training. The training of the nonsense words began on the day following the electrophysiological pretest session. Training occurred over a consecutive 5-day period and ended the day before the electrophysiological posttest portion of this study. As part of this training, the parents were given two objects varying in shape (short, round cylinder; long, square rectangle) and color (yellow, blue). Each object was assigned a name, either "gibu" or "bidu." The objects and their associated names were randomized across subjects with the exception that half of the infants were trained to associate "gibu" with one object and "bidu" with another, whereas the other half of the subjects were trained to associate the two CVCV words to the opposite objects. Parents were instructed to find two times during the day when their infants would be in a good mood and willing to play with the "gibu" and the "bidu" objects. During the training procedure, parents were asked to encourage their infant to manipulate the objects, in 10-min sessions for each object, two times a day (for a total time of 20 min per object per day) for 5 days. The parents were also asked to use the name of the objects in different sentence combinations such as, "Look at the gibu," or "Where is the bidu?" Parents were given a written protocol to follow in training their infant, which included instructions about counterbalancing the objects during presentations. To verify that training had occurred and that it followed the instructions to the parents, the play sessions were recorded on cassette by the parents. Parents were also given (a) copies of the Edinburgh Handedness Inventory (one per parent), (b) a handedness questionnaire for their infant, and (c) a questionnaire concerning their child's development and daily activities to be returned on the day that the electrophysiological testing occurred.

Posttraining behavioral test. This testing occurred on the day following the final day of behavioral training. The parent who played with the child most in the training sessions was asked to rate the child's understanding of the terms "gibu" and "bidu." As part of the evaluation, a rating scale designed to assess the parents' judgment as to whether the infant knew the meaning of a term and the parents' own confidence in their judgment of the infant's word knowledge was used. To accomplish this, the questions were administered in two parts. First, the parent was asked to indicate whether or not the infant knew the name of the object in question. The parent answered "yes" or "no" to this question. The parents were then instructed to rate their own confidence in that judgment by using a 5-point scale with 1 as *completely not confident* and 5 as *very confident.* Thus, if a parent thought that the infant did not understand the word but the parent was not very sure of their own judgment, they would answer "no" to the first question and would give a rating of 1. If, on the other hand, the parent believed that the infant understood the term but again was not confident of his or her answer, he or she would respond with a "yes" to the first question and would give a rating of 1. If the parent was confident, however, that the infant understood the term, an answer of "yes" would be given along with a rating of 5. The parents' responses for each term were then rescaled along a 10-point continuum with the *unknown* category making up the first 5 points of the scale and the *known* category the next 5. The confidence rating was then used to arrange the parents' decisions within each group. Consequently, after conversion, the parents' ratings were transformed to range from a *confidently unknown* rating of 1 to an *unknown but completely not confident* rating of 5 to a *known but completely not confident* rating of 6 to a *confidently known* rating of 10. Using this rating system, all parents rated both of the bisyllable terms as *known* by the infant. The mean confidence rating for "bidu" was 8.71 ($SD = .88$); for "gibu," the average rating was 8.79 ($SD = .94$), indicating that the parents as a group were confident that the infants understood the terms. There was no difference in parental ratings between the two CVCV items, $t(26) = .32$.

Posttraining electrophysiological test. This testing was conducted following the conclusion of the posttraining behavioral test. The electrophysiological techniques used were identical to those used during the pretest phase of this study, but this test session differed in two important respects from the procedure used in the pretest electrophysiology session.

First, the infant was trained on a simple task immediately prior to the electrophysiological testing. In this task, the mother was given a plastic transparent bottle with a wide, open mouth to hold next to the infant. The infant was then handed a small disk. One of the experimenters then placed an identical disk in the bottle. This action was repeated until the infant placed the disk in the bottle. At this point the infant was handed another disk. If the infant placed this second disk in the bottle, training was finished, the disks were put aside, and testing commenced. If the infant failed to imitate the placement of the disk into the bottle, the experimenter again demonstrated the placement and watched the infant's response. All infants learned this simple task within 3 to 4 trials. This task was used to keep the infant's attention focused on the objects throughout the test session. Within a few minutes following training, the auditory stimuli, which consisted of the nonsense CVCV bisyllables presented during the pretest and training periods, were presented through a speaker positioned over the midline of the infant's head at the same loudness level used in the pretest electrophysiological session.

Second, the auditory stimuli, "gibu" and "bidu," were presented in the five-trial blocks of the same stimulus while the infant held either the corresponding object or the object that matched the other auditory stimulus. On half of the trials, there was a match between the name the infant heard and the object that it was handed. On the other half of the trials, there was not a match. Immediately before the onset of each stimulus block, the infant was handed one of the training objects so that the infant would have the opportunity both to examine and to manipulate the object. The infant would usually hold the object and then place it in the bottle that was held by the parent. Prior to the presentation of each name within a block of five trails, identical objects were given to the infant in sequence to help maintain the infant's attention. At the end of a block of trials, any objects the infant continued to hold were taken from the infant and placed in the bottle. The presentation of a "match" or a "no match" object during specific trials was randomized across blocks and across subjects. Because two different CVCV bisyllables were used in the training session, each CVCV was presented during electrophysiological testing and occurred an equal number of times as a match or as a mismatch to the object. In this manner, each CVCV bisyllable occurred 80 times, with 40 times as a match to the object and 40 times as a mismatch. In all, there were 160 total stimulus presentations. This step was taken to eliminate the possibility that subsequent results were due to acoustic or perceptual differences between the stimuli across the entire group of infants.

Analyses. For both the pre- and posttraining electrophysiological data sets, individual auditory evoked responses were digitized at 10-ms intervals for a 600-ms period following stimulus onset. This interval and sampling rate was selected on the basis of pilot work that indicated that most of the synchronized activity of the AER elicited from the infants in a match/mismatch paradigm had concluded at the end of the 600-ms poststimulus onset period. Digitized values were stored on-

Although shielded earphones might have corrected this last problem, the additional weight and the lack of calibration adjustments precluded use with the infants.

line during the data recording session by a Macintosh Plus microcomputer using the EPACS© (1988; Evoked Potential Analysis and Collection System) software package. Subsequent analyses using the EPACS© software package were conducted off-line after the testing session had been completed. Artifact rejection was carried out on the AER data for each electrode to eliminate from further analyses the AERs contaminated by motor movements. If an artifact (operationally defined as a shift in the voltage level in excess of 40 μV) occurred on any one electrode channel during the 600-ms poststimulus period, all of the AERs collected across all of the electrode sites for that trial were discarded from subsequent analyses. This procedure, which was based on the peak-to-peak amplitudes of single trial responses, resulted in an average rejection rate of 11.03% of the trials (SD = 7.0) for the pretest AER data and 16.74% of the trials (SD = 6.99) for the posttraining test. Rejection rates were comparable across infants for the different stimulus conditions. Following artifact rejection, the single trial data were then averaged separately for the pre- and posttraining AER test data for each of the six electrode sites and each of the two stimulus conditions. For each data set for each infant, 12 averages were obtained. Each average was based on 80 samples combining responses to "bidu" and "gibu" for the match condition and 80 samples across "bidu" and "gibu" for the mismatch condition. In this manner, 168 averaged AERs were obtained for the 14 infants for the pretraining data and another set of 168 averaged AERs were obtained for the posttraining AER data.

In the case of the posttraining data set, these included averages for the matched and mismatched trials for each of the six electrode sites. The pretraining AER data were sorted in exactly the same manner although obviously in this case, because no training had occurred, there were no actual match or mismatch trials. This was done to provide a control comparison for the posttraining AER data set. Although differences in AERs between the match and mismatch trials were anticipated for the posttraining data set, no such differences were expected for the pretraining AER data.

Each data set was next submitted to a two-step analysis procedure that first involved the use of a principal components analysis (PCA) and then an analysis of variance (ANOVA). This analysis sequence followed the procedures outlined and used successfully in previous studies. Although there are a variety of different analysis procedures that could be used to analyze the AER data, we decided to use a multivariate approach that has produced consistent results in programmatic research across a number of laboratories (Brown et al., 1979; Chapman et al., 1979; Donchin, Tueting, Ritter, Kutas, & Heffley, 1975; Gelfer, 1987; Molfese, 1978a, 1978b; Molfese & Molfese, 1979a, 1979b, 1980, 1985; Ruchkin et al., 1981; Segalowitz & Cohen, 1989). For example, Molfese, in a series of articles investigating speech perception cues such as voice onset time and place of articulation, noted consistent systematic effects across studies for each cue (Molfese, 1978a, 1978b, 1980, 1984; Molfese & Schmidt, 1983). Moreover, these effects have also been independently replicated by using comparable analysis procedures (Gelfer, 1987; Segalowitz & Cohen, 1989). The rationale for the use of this procedure is that it has proven successful both in identifying regions of the AER where most of the variability occurred across AERs and subjects and subsequently in determining if the variability characterized by the different factors was due to systematic changes in the independent variables under investigation. The PCA procedure behaves somewhat similarly to a factor analysis with the exception that it constructs the factors on the basis of variances instead of correlations (Rockstroh et al., 1982, p. 63). The PCA procedure itself is blind to individual experimental conditions and generates the same solution regardless of the order in which the AERs are entered. Once the PCA identified where within the AERs most of the variability occurred, the ANOVA was used to identify the cause of this variability. The ANOVA accomplished this task by determining whether the variability reflected in the factor scores assigned for each factor to each averaged AER differed as a function of changes in the independent variables. This procedure directly addresses the question of whether the AER waveshapes in the region characterized by the most variability for any one factor changed systematically in response to the match versus the mismatch word–object conditions recorded from the different electrode sites over each hemisphere.

Results

Twelve averaged AERs (two conditions, two hemispheres, and three electrodes per hemisphere) were recorded from each of the 14-month-old infants in this study during the electrophysiological pretraining test session, and 12 averaged AERs were recorded from each infant during the posttraining test session. For both data sets, each averaged AER was based on 80 stimulus repetitions (40 from "bidu" and 40 from "gibu" during the match condition and 40 from each in the mismatch condition) and consisted of 60 data points sampled at 10-ms intervals that began with stimulus onset and ended 600 ms later. In the following section, data will be presented first for the pretest session and then for the posttest session. To decrease the likelihood of a Type 1 error, only effects beyond the .01 level are reported to address the problem of experiment familywise error (Keppel, 1982, p. 145). In cases in which post hoc analyses were conducted, the conservative Scheffé Critical F test procedure with the same p level was used (Scheffé, 1959).

Pretraining AER Data Analysis

First, a group-averaged AER was constructed on the basis of all of the averaged waveforms from all of the infants. This grand average or centroid of the 168 averaged AER waveforms obtained from all of the 14 infants during the pretraining AER test session was characterized by a small initial positive wave that reached a peak at 120 ms (P120) following stimulus onset. This was followed by a large negative deflection that peaked at 220 ms (N240) and a large positive peak at 320 ms (P360). This second positive peak was followed in turn by a second large negative deflection that peaked at 490 ms (N490). Finally, a small positive peak occurred at 560 ms (P560) following stimulus onset. The group-averaged AERs for the different electrode sites and "stimulus conditions" are presented in Figure 1. These group-averaged AERs were recorded from the frontal (F), temporal (T), and parietal (P) regions of both the left and right hemispheres in response to presentations of the CVCV auditory tokens later used for training. Although no training for matching the CVCV to a specific object occurred during this pretraining session, the AERs presented in Figure 1 are arbitrarily grouped under MATCH and MISMATCH labels by using the same grouping procedure later used for constructing the averages for the data of the posttraining session. This was done to facilitate comparisons with the posttraining data presented in Figure 4.

The averaged AERs from the pretraining session formed the input matrix for the PCA using the BMDP4M program from the BMDP87 package (Dixon, 1987). This program first transformed the data into a covariance matrix and then applied the PCA to this matrix. Eight factors accounting for 54.86% of the total variance were selected for further analyses based on the Cattell Scree Test (Cattell, 1966). These factors were then rotated by

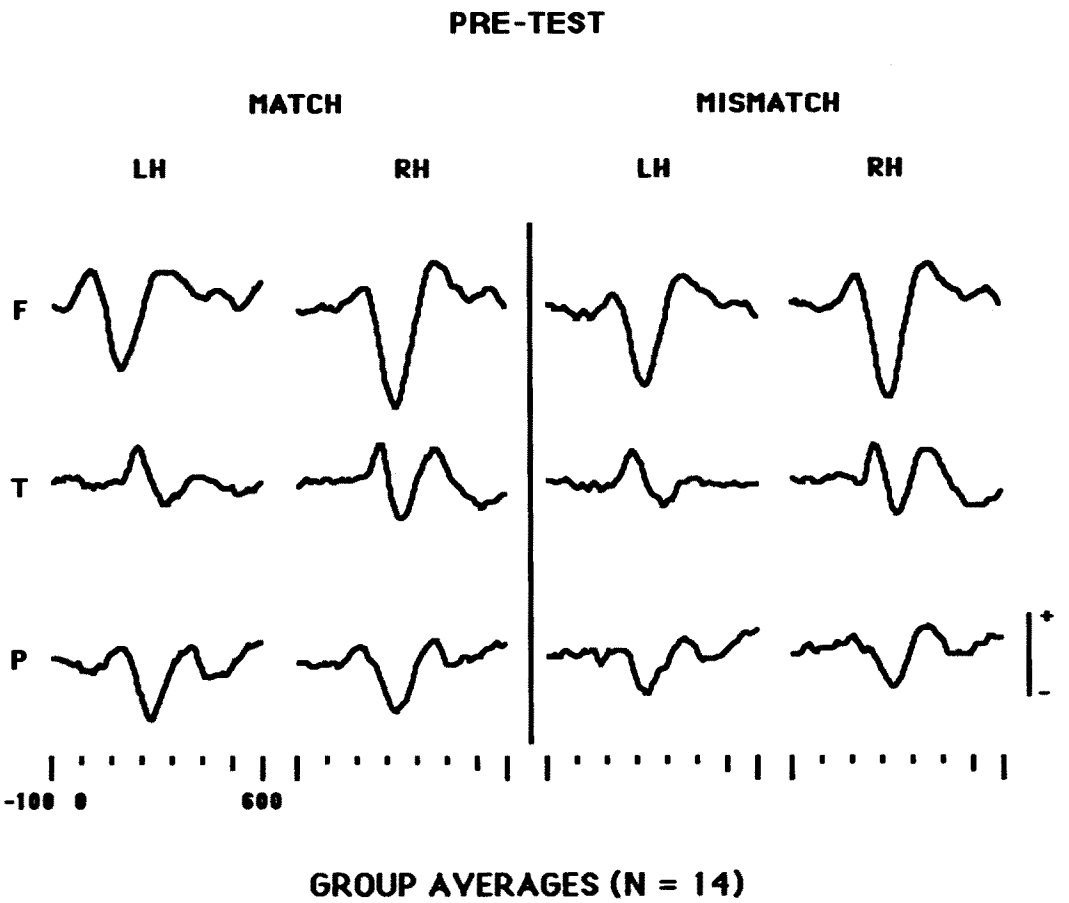

Figure 1. The group-averaged auditory evoked response (AER) waveforms from the 14 infants that were collected during the pretraining AER session. (The AERs were recorded from the frontal [F], temporal [T], and parietal [P] regions of both the left and right hemispheres in response to presentations of the consonant–vowel–consonant–vowel (CVCV) auditory tokens later used for training. Although no training for matching CVCV to a specific object occurred during this session, the AERs are grouped here under *match* and *mismatch* using the same sorting assignment as was used with the posttraining data to facilitate comparisons with the posttraining data presented in Figure 4. A 100-ms prestimulus period is presented as well as the 600-ms poststimulus onset period. Stimulus onset began at 0 ms. Positivity is up. The calibration marker is 10 μV LH = left hemisphere; RH = right hemisphere.)

using the normalized varimax criterion (Kaiser, 1958), which preserved the orthogonality among the factors while improving their distinctiveness. This analysis generated factor scores or weights for each of the 168 averaged AERs for each of the eight rotated factors. The variance isolated by the PCA was characterized by the eight factors (factor loadings). These factor scores, which reflected the amount of variability for that factor in an individual AER, constituted the dependent variables in the subsequent ANOVA described later. The peak for each factor and the area immediately surrounding it in time indicated that this region of the brainwave changed in amplitude or slope across some proportion of the AERs in our data. A minimum factor-loading cutoff of $\pm.35$ was used to estimate the region of variability for each factor. Thus, for example, if factor loading exceeded +.35 for a particular factor, it was estimated that the major region of variability occurred at this point. The centroid or grand averaged AER and the eight factors derived in this analysis are presented in Figure 2.

Next, a series of eight independent ANOVAs using the BMDP8V statistical package (Dixon, 1987) was conducted. The separate ANOVAs conducted for each factor were appropriate because the factors derived by the PCA were found to be orthogonal to each other as shown by the multivariate ANOVA (MANOVA). The univariate ANOVA design was $2 \times 7 \times 2 \times 3 \times 2$ (Sex Differences $\times$ Subjects $\times$ Word Match $\times$ Electrode Sites Within Hemispheres $\times$ Hemispheres). Because no interactions of sex with any of the independent variables were noted, the data were collapsed and the analyses redone. The final ANOVAs used in this study were based on the design of $14 \times 2 \times 3 \times 2$ (Subjects $\times$ Word Match $\times$ Electrode Sites Within Hemispheres $\times$ Hemispheres). These ANOVAs were conducted to determine if any of the regions of the AERs identified by the factors varied systematically as a function of the specific levels of the independent variables in this study.

One region of the AERs was found to discriminate between electrode sites across hemispheres. A main effect for electrodes

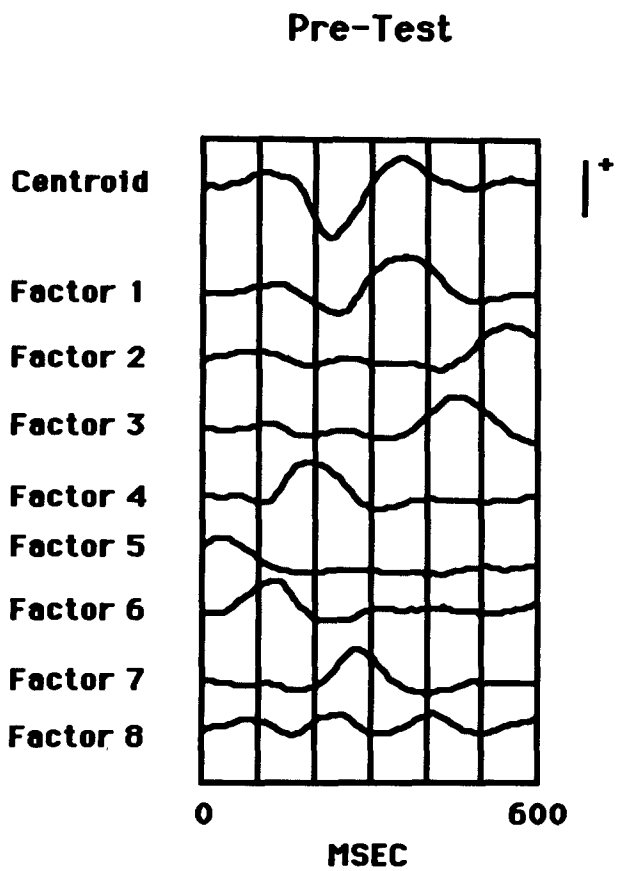

Figure 2. The centroid and the eight factors identified by the principal components analysis for the pretraining auditory evoked potential (AER) data set. (For the centroid, the calibration marker is 10 μV with positivity up. AER duration is 600 ms. The percentage of total variance accounted for by each factor is displayed to the right of that factor.)

was found for Factor 6, $F(2, 26) = 23.55$, $p < .0001$, that characterized variability in the AER waveforms between 160 and 260 ms. This effect resulted from differences between the temporal and frontal sites, $F(1, 26) = 46.39$, $p < .00001$, and between the temporal and parietal sites, $F(1, 26) = 17.3$, $p < .0005$. No match effects were found, even using the less conservative and more conventional significance level of .05.

Posttraining AER Data Analysis

The same analysis procedures used for the pretraining data were also applied to the posttraining data, with one exception: An additional MANOVA procedure was used to further evaluate match effects found with the posttraining AER data set.

First, a group-averaged AER was constructed on the basis of all of the averaged waveforms from all of the infants. This grand average or centroid of the 168 averaged AER waveforms obtained from all of the 14 infants during the AER test session following the 5 days of behavioral training was characterized by a small initial positive wave that reached a peak at 110 ms (P110) following stimulus onset. This was followed by a large negative deflection that peaked at 220 ms (N220) and by a large positive peak at 320 ms (P320). This second positive peak was followed in turn by a second large negative deflection that peaked at 440 ms (N440). Finally, a small positive peak occurred at 560 ms (P560) following stimulus onset. These group averages are presented in Figure 3.

Several analyses were conducted on this averaged data set. These AERs formed the input matrix for the PCA using the BMDP4M program from the BMDP87 package (Dixon, 1987). This program first transformed the data into a covariance matrix and then applied the PCA to this matrix. Eight factors accounting for 55.78% of the total variance were selected for further analyses based on the Cattell Scree Test (Cattell, 1966). The eight factors, the percentages of variance accounted for by each factor, and the centroid AER for the entire posttraining AER data set are presented in Figure 4.

Two factors found in subsequent analyses reflected changes in the AERs that were directly related to correct matches between the object and its name. Factor 4, accounting for 6.76% of the total variance, reflected waveform variability beginning about 30 ms after stimulus onset, peaking at 60 ms, and ending approximately 120 ms following stimulus onset. Given its temporal overlap with the first positive peak of the averaged AERs, it appears that this factor reflected changes (variability) in the waveform leading up to the peak of the P110 component. A second factor, Factor 6 (6.28% of the total variance), reflected waveform variability beginning 520 ms after stimulus onset that peaked at 580 ms and then diminished by 600 ms. This factor, therefore, reflected changes in the slope and amplitude of the P560 component. Four other regions of the AER waveform, as indicated later, were found to vary as a function of electrode sites. Factor 1, which accounted for 9.96% of the total variance, reflected changes in the AER waveform between 430 and 550 ms. Factor 2 (8.37% of the total variance) characterized the variability in the AER between 360 and 460 ms. Given their latencies, it appears that Factors 1 and 2 reflect variations in the waveforms in the region of the N440 component. Factor 3 (7.48% of the total variance) represented changes in the waveform that occurred between 260 and 350 ms following stimulus onset. This factor reflected changes in the region of the P320 component. Finally, Factor 7 (5.36 of the total variance) reflected changes in the AER (surrounding the N220 AER peak) that occurred between 150 and 230 ms.

MANOVA

The PCA, as noted previously, generated a set of factor scores for each averaged AER for each factor. Consequently, 168 factor scores were generated for Factor 1, 168 factor scores were generated for Factor 2, 168 for Factor 3, and so on. These factor scores, which reflected the amount of variability for that factor in an individual AER, constituted the dependent variables in two different analyses. First, a MANOVA was conducted by using the SPSSX procedure to assess the interrelationship between experimental factors. The univariate ANOVA design was $2 \times 7 \times 2 \times 3 \times 2 \times 8$ (Sex Differences $\times$ Subjects $\times$ Word Match $\times$ Electrode Sites Within Hemispheres $\times$ Hemispheres $\times$ Factors).

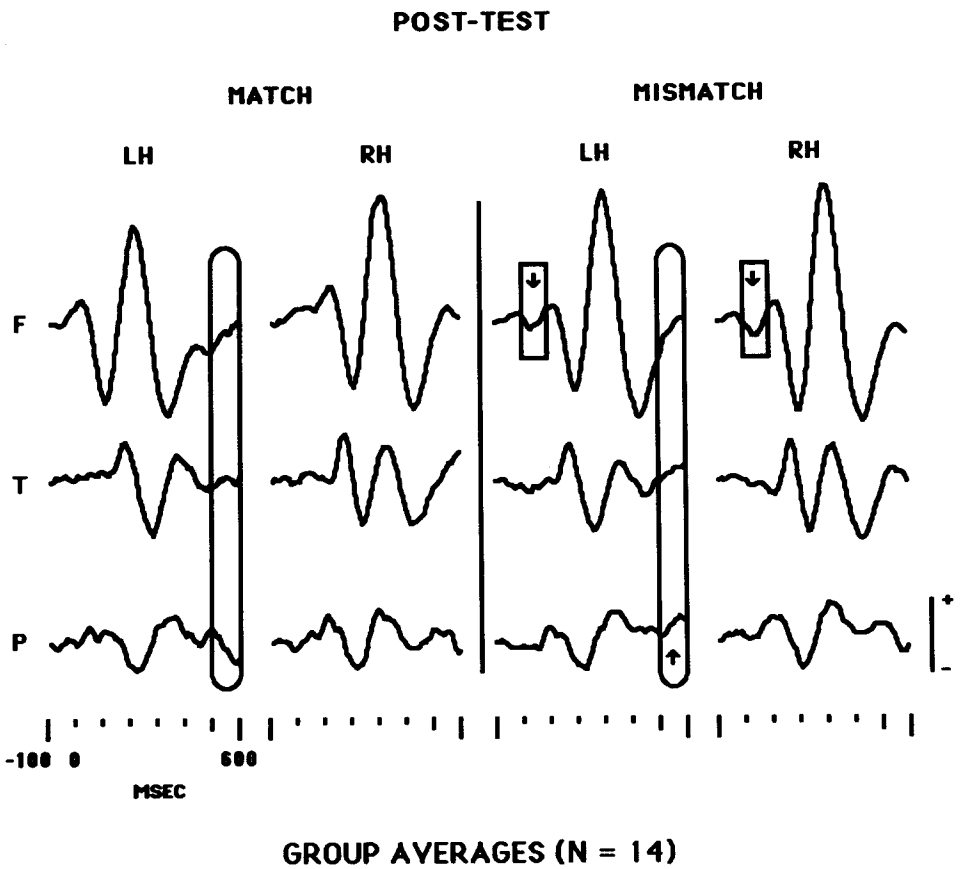

Figure 3. The group-averaged auditory evoked potential (AER) waveforms from the 14 infants that were collected during the posttraining AER session. (The AERs were recorded from the frontal [F], temporal [T], and parietal [P] regions of both the left and right hemispheres in response to presentations of the CVCV auditory tokens following 5 days of training. A 100-ms prestimulus period is presented as well as the 600-ms poststimulus onset period. Stimulus onset began at 0 ms. Positivity is up. The calibration marker is 10 μV. LH = left hemisphere; RH = right hemisphere.)

Because no interactions of sex with any of the independent variables were noted, the data were collapsed and the analyses redone.[3]

Univariate ANOVAs

Next, a series of eight independent univariate ANOVAs using the BMDP8V statistical package (Dixon, 1987) were conducted, one for each factor. These separate ANOVAs were appropriate because the factors derived by the PCA were found to be orthogonal to each other, as shown by the MANOVA. The univariate ANOVA design was 2 × 7 × 2 × 3 × 2 (Sex Differences × Subjects × Word Match × Electrode Sites Within Hemispheres × Hemispheres). Because no interactions of sex with any of the independent variables were noted, the data were collapsed and the analyses redone. The final ANOVAs used in this study were based on a 14 × 2 × 3 × 2 design (Subjects × Word Match × Electrode Sites Within Hemispheres × Hemispheres). These ANOVAs were conducted to determine if any of the regions of the AERs identified by the factors varied systematically as a function of the specific levels of the independent variables in this study. For the purposes of this study, the match effects are reported in the order in which they occurred in the AER waveform. Next, effects are reported for the electrode effects that did not interact with the word match effects.

Match effects. An interaction for Match × Electrode Site was noted for Factor 4, $F(2, 26) = 6.56$, $p < .0049$. A graph of the means for this interaction is presented in Figure 5. The factor scores or weights that served as the dependent variables in the analyses are the metric depicted along the ordinate of the graph. Post hoc Scheffé procedures indicated that only the portion of the AERs between 30 ms and 120 ms following stimulus onset (P110) that were recorded from over frontal electrode sites discriminated between the match and mismatch conditions, $F(1, 26) = 39.98$, $p < .001$. As can be seen in the graph, the largest difference between the match and mismatch conditions occurred for the frontal electrode sites.

[3] The specific concern focused on the match-related effects. However, no evidence of correlated factors were found. For example, in the Match × Electrode interaction, the Mauchly Sphericity Test generated a ω of .66107, $\chi^2(2) = 4.97$, $p < .083$. The Greenhouse-Geisser epsilon was .74687. Thus, the homogeneity assumption was upheld and subsequent univariate analyses of variance were considered appropriate.

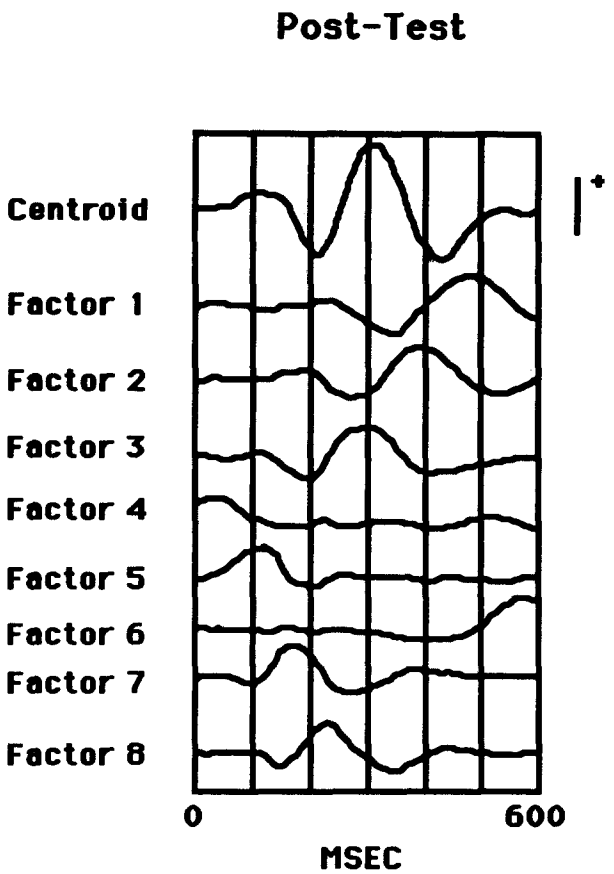

Figure 4. The centroid and the eight factors identified by the principal components analysis for the posttraining auditory evoked potential (AER) data set. (For the centroid, the calibration marker is 10 μV with positivity up. AER duration is 600 ms. The percentage of total variance accounted for by each factor is displayed to the right of that factor.)

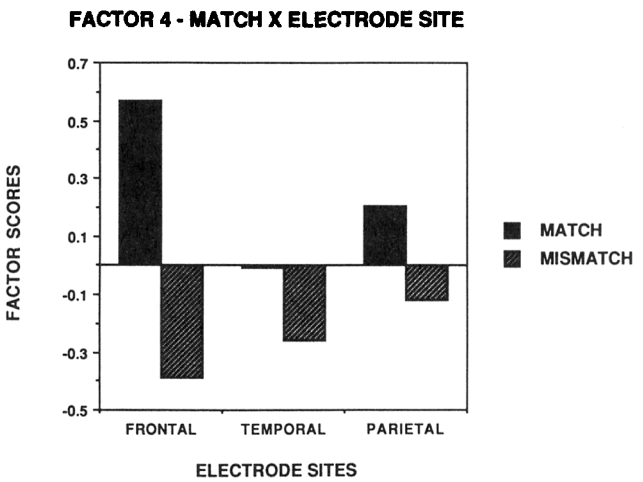

Figure 5. The graphic representation of the Match × Electrode interaction for Factor 4, depicting the differences in the frontal, temporal, and parietal electrode responses elicited during the match and mismatch conditions. (The factor scores or weights that served as the dependent variables in the analyses are the metric depicted along the ordinate of the graph.)

In addition, a Match × Hemisphere interaction was noted for Factor 6, $F(1,13) = 12.85$, $p < .0033$. The means for this interaction are represented graphically in Figure 6. Post hoc Scheffé procedures indicated that this interaction effect was due to left hemisphere discrimination of the match and mismatch trials as reflected in changes in the region of the P560 component, $F(1, 13) = 9.39$, $p < .0089$. This effect is represented by the differences in the means between the diagonal striped bar and the dotted bar for the left hemisphere. In addition, the left and right hemisphere recorded AERs differed from each other during the match trials, $F(1,13) = 12.14$, $p < .0042$. This last contrast is characterized by the differences in the bars for the match condition between the left and right hemispheres. The differences in means as represented by the two interactions just described can be seen in the group-averaged AER waveforms presented in Figure 3, as well as in the individual subject data (from Infants 2 and 13), which are presented in Figures 7 and 8. The region of the AER waveform that was identified in the Match × Electrode interaction of Factor 4 (and that discriminated between the match and mismatch conditions at only the frontal electrode sites) is represented by the region of the AER contained within the rectangle in both figures. The numeral 4 immediately above the rectangle above the far left AER waveform indicates that this was the region identified as Factor 4. As can be seen in both figures, the AER activity for the mismatch condition is characterized by a downward pointing peak for both the left and the right hemisphere electrode sites. For the match condition, however, the waveform shows a positive deflection. No such difference can be reliably noted across the other electrode sites for this region of the AER.

The left hemisphere lateralized effect identified by the Match × Hemisphere interaction of Factor 6 can also be seen as

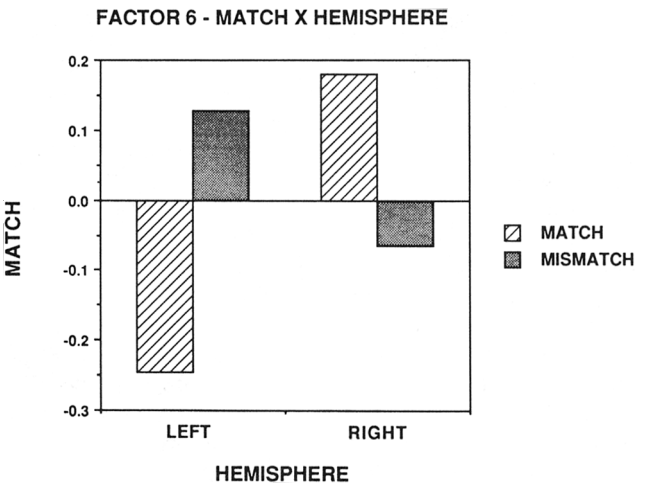

Figure 6. The graphic representation of the Match × Hemisphere interaction for Factor 6, depicting the differences in left and right hemisphere responses to the consonant–verb–consonant–verb (CVCV) bisyllables during the match and mismatch conditions. (The factor scores or weights that served as the dependent variables in the analyses are the metric depicted along the ordinate of the graph.)

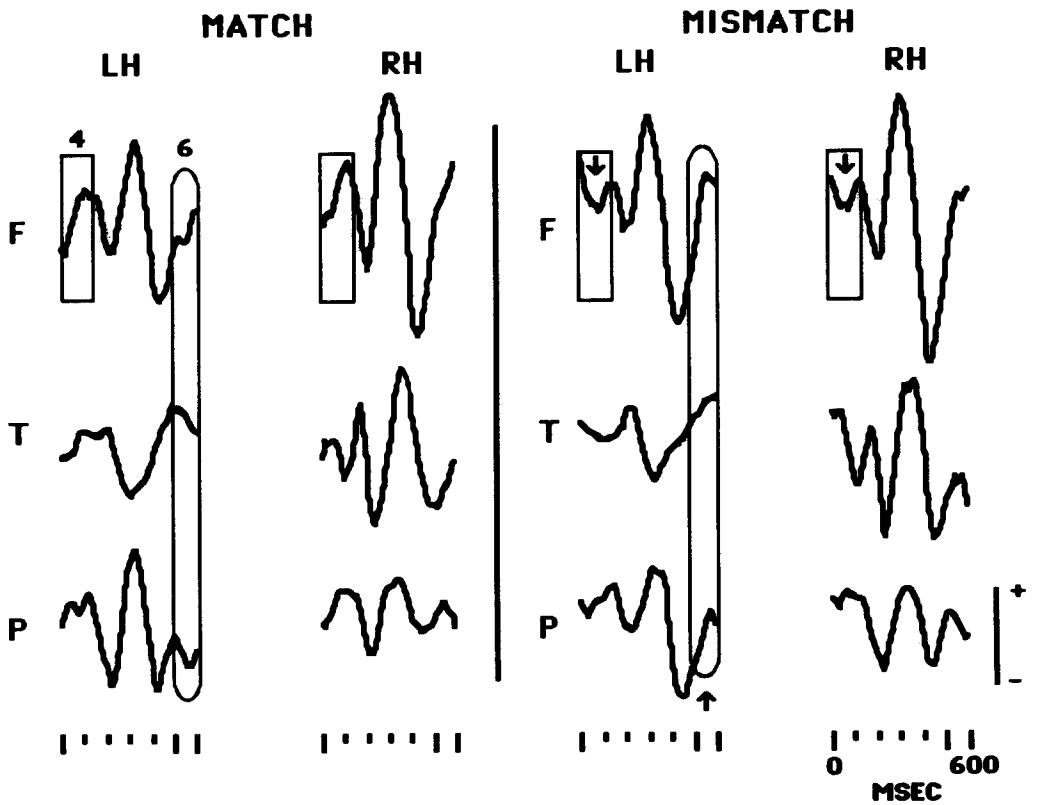

Figure 7. The averaged auditory evoked potentials (AERs) from Infant 2 recorded from the left and right frontal (F), temporal (T), and parietal (P) electrode sites in response to auditorily presented consonant–vowel–consonant–vowel (CVCV) bisyllables that either matched or failed to match (mismatch) the objects presented to the infant. (Positivity is up. The calibration marker is 10 μV. AER duration is 600 ms. LH = left hemisphere; RH = right hemisphere.)

changes in the AER waveforms for these two infants. This effect, which characterized changes in the AER waveform between 520 and 600 ms, is characterized by marked upward deflections of the AER waveform elicited during the mismatch condition. This region of variability occurred within the late portion of all of the left hemisphere waveforms contained within the oval and is identified by the numeral 6 (for Factor 6) positioned above the oval in the left, topmost AER for both infants. The match condition appeared to elicit either a downward falling wave during this time or a markedly smaller rising component than that noted for the mismatch condition.

Electrode effects. In addition to the match-related effects, a number of regions of the AER were found to vary across electrode sites within hemispheres. These included a main effect for electrodes for Factor 1, $F(2, 26) = 6.84$, $p < .0041$. Tests of this effect indicated that AER activity recorded from the frontal electrode sites differed from that recorded over the parietal, $F(1, 26) = 13.65$, $p < .0013$. No differences were noted between the frontal and temporal, $F(1, 26) = 2.91$, $p > .05$, or temporal and parietal, $F(1, 26) = 3.96$, $p > .05$, sites. A main effect for electrodes, $F(2, 26) = 20.57$, $p < .00001$, was found for Factor 3. Tests of this effect indicated that AER frontal activity differed from temporal activity, $F(1, 26) = 66.07$, $p < .00001$, frontal differed from parietal activity, $F(1, 26) = 15.48$, $p < .0008$, and temporal differed from parietal activity, $F(1, 26) = 17.59$, $p < .0005$. Finally, a main effect for electrodes, $F(2, 26) = 17.90$, $p < .00001$, was noted for Factor 7. Tests of this effect indicated that AER activity recorded from the frontal sites differed from the temporal sites, $F(1, 26) = 26.62$, $p < .00001$, but not the parietal sites, $F(1, 26) = .32$, $p > .5$. The temporal AER activity was different from the parietal activity, $F(1, 26) = 23.75$, $p < .00001$.

Split-Half Comparison

A split-half comparison was executed for the posttraining AER data set to determine whether the data set had some degree of internal consistency. To accomplish this, the digitized single trial data were divided into two sets and reaveraged to produce two averaged brain responses for each subject, condition, and electrode site. Following this procedure, each subject had one average based on one half of the data and a second average based on the other half. In this manner, each average was based on 40 repetitions of the CVCV syllables (20 from "bidu" and 20 from "gibu" when they were in the match condition and 20 from each when they were in the mismatch condition) instead of on the 80 repetitions used to construct each

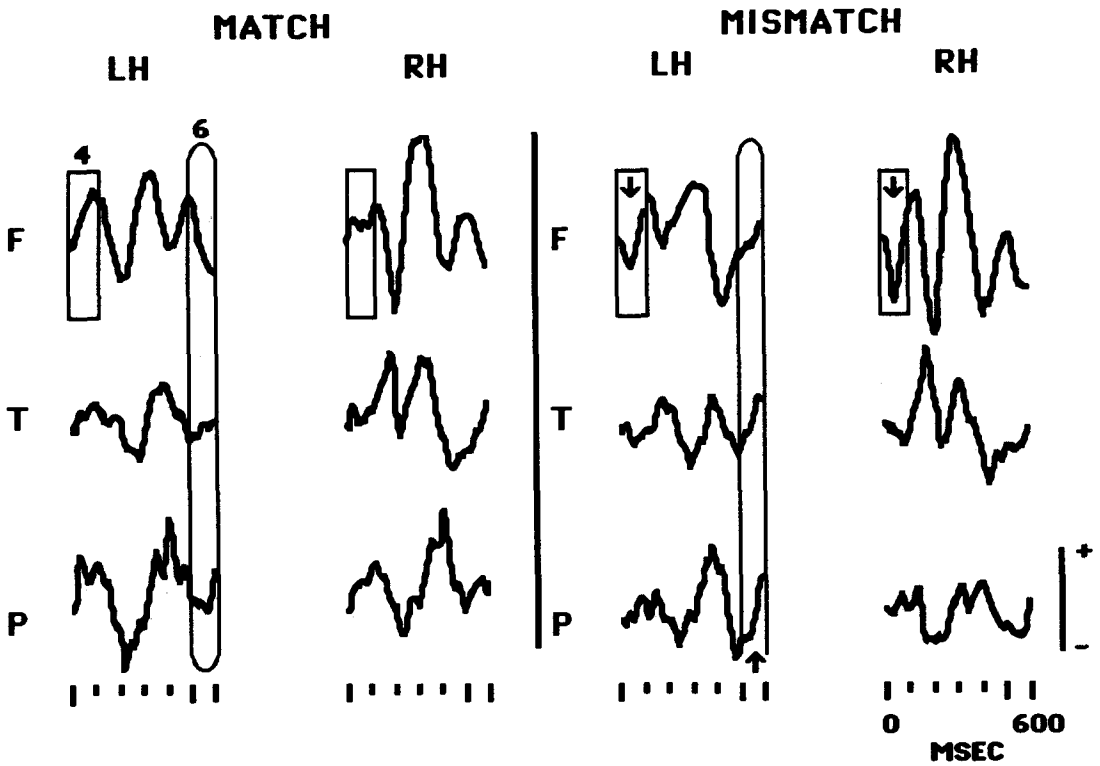

Figure 8. The averaged auditory evoked potentials (AERs) from Infant 13 recorded from the left and right frontal (F), temporal (T), and parietal (P) electrode sites in response to auditorily presented consonant–vowel–consonant–vowel (CVCV) bisyllables that either matched or failed to match (mismatch) the objects presented to the infant. (Positivity is up. The calibration marker is 10 μV. AER duration is 600 ms. LH = left hemisphere; RH = right hemisphere)

average in the original analyses. As before, averages were collapsed over the different acoustic stimuli and were developed only for the match and mismatch conditions. Because the conditions had been presented using a blocked random procedure, the contributions to each split-half average came equally from the beginning, middle, and end of the digitized trials. This data set was then included in a second series of PCA–ANOVA procedures, using the split halves as another factor in the ANOVA. The PCA procedure, to insure comparability to the original analyses, was set to generate eight sets of factors and factor scores, which were then analyzed by using an ANOVA, 14 × 2 × 2 × 3 × 2 (Subjects × Split Half × Word Match × Electrode Sites Within Hemispheres × Hemispheres). The same p level and post hoc procedures as were used in the main analyses were applied to this new analysis. This analysis accounted for approximately the same amount of the total variance as the original procedure, 54.94%. Although no main effects or interactions for the Split-Half factor were identified, Match × Electrodes, $F(2, 26) = 5.92$, $p < .01$, and Match × Hemisphere, $F(1, 13) = 11.26$, $p < .005$, interactions were statistically significant. The latencies of these factors were identical to those corresponding effects reported earlier. Furthermore, Scheffé tests of these two interactions identified comparable differences to those noted in the original analyses. This test of within-subjects consistency continued to support the main findings reported previously concerning the ability of the evoked potentials to differentiate trials on which a match had occurred between an auditory and a visual stimulus and those trials on which a mismatch occurred.

Discussion

The AER activity identified in this study during the post-training AER session discriminated between auditory stimuli that were correctly paired with named objects versus objects that were trained to different names. Because no such effects were noted in the pretraining AER session, it is clear that the AERs can detect changes that occur as a function of training and that they can detect differences between the match and mismatch conditions. Two regions of the AER waveform reliably varied during this task: an early component of the AER that changed bilaterally over the frontal regions of both hemispheres, and a late-occurring lateralized response that was restricted to only the left hemisphere electrode sites. When a correct match occurred between the auditorily presented word and the object that the infant held, both the left and right frontal regions of the brain emitted brain responses that contained an initial positive deflection or peak between 20 and 100 ms

following the auditory onset of the object name. If a mismatch occurred, however, this early positive deflection inverted 180° and became a negative deflection. Later in time, between 520 and 600 ms, just before the conclusion of the AER, a large positive-going wave occurred over only the three left hemisphere electrode sites when the infant listened to a stimulus that did not name the object that the infant held. Given the short latency of the initial changes in the AER waveshape across the frontal regions, it appears that the young infant must recognize almost immediately if there is agreement between something that it hears and something that it sees and touches. Approximately 400 ms later, the frontal, temporal, and parietal regions of the left hemisphere of the infant's brain similarly discriminated between the CVCV bisyllables when they were correctly versus incorrectly paired with objects.

Perception of Coarticulated Cues

Even though the early AER response during the first 100 ms following stimulus onset might superficially appear to have occurred before the infant could process the acoustic information of the CVCVs, such early discrimination is not without precedence. Given other behavioral and electrophysiological investigations of coarticulated speech cues (Ali, Gallagher, Goldstein, & Daniloff, 1971; Daniloff & Moll, 1968; MacNeilage & DeClerk, 1969; Molfese, 1979), it is possible that the infants used such information to discriminate between the match and mismatch conditions. Coarticulation refers to the finding that the shape of the vocal tract during the production of a speech sound will be altered by the place and manner of articulation for later speech sounds. In one such study, MacNeilage and DeClerk (1969) made cineflurograms and electromyograms of individuals producing a series of 36 CVC syllables and found that the articulation of the initial consonant sounds changed as a function of the identity of the following sounds. Ali et al. (1971) noted a perceptual counterpart of coarticulation. They constructed a series of CVC and CVVC syllables in which the final sound was either a nasal [m, n] or nonnasal consonant. After the final vowel–consonant and consonant transitions were removed, the resulting CV and CVV syllables were presented to a group of adults who were able to discriminate between the nasal and nonnasal sequences at well above chance levels. Ali et al. and others have argued that such coarticulated information allows the listener to perceive and process some or all of the utterance before it has been completely articulated. Molfese (1979), in a study with adults, recorded AERs to CVC words and nonsense syllables that differed from each other only in the final consonant sound. Adults listened to each CVC and then after a brief delay pressed one of two keys to indicate whether they had heard a word or a nonsense syllable. Three regions of the AER, including one that peaked 60 ms following stimulus onset, changed systematically as a function of the meaningfulness of the CVC syllables. Molfese interpreted this component, as well as the later occurring negative components at 260 and 400 ms, as sensitive to the coarticulated speech cues that carried information concerning the later occurring (after 650 ms) final consonant sound. In our study, given that the consonant burst and frequency transition information that discriminated one CVCV from the other occurred during the first 50 ms of each stimulus (MacNeilage & DeClerk, 1969), it is possible that the infants could have used this coarticulated information to rapidly identify and discriminate early in time between the auditory tokens that matched or failed to match the object the infant was holding throughout a block of trials. Thus, the early AER component identified as Factor 4 could reflect such a process.

AERs as Measures of "Meaning"

These results demonstrate that 14-month-old infants can learn to match and retain "arbitrary" (and relatively uninteresting) auditory–visual pairings after several days of parental training. In studies of "amodal" matching, Wagner and Sakovits (1986) provided support for a model that describes the complex changes of novel/familiar preferences in cross-modal matching as a function of increasing levels of stimulus exposure. Future studies of "arbitrary" cross-modal matching might also profit from a similar examination of the role of training/exposure experience. Furthermore, as predicted from the adult aphasia literature, this rudimentary level of semantic processing was evident in infants at all sites recorded over the left hemisphere. In addition, as suggested by the adult neuropsychological semantic processing literature, infants also exhibit a bilateral AER component to matching versus mismatching auditory–visual pairs.

The procedure outlined in our study first trained some link or association between an object and a name and then tested for the presence of neuroelectrical components of the AER that signal that such an association has occurred. Although this project appears to have succeeded, it is obvious that little information concerning the nature of these brain differences is yet known. Do they reflect general associations between the sounds and images? Can such procedures be used to assess infants' learned associations to specific auditory and visual stimuli? Although it is obviously beyond the scope of a single experiment to demonstrate convincingly that the patterns of brain activity that differentiated object–name matches from mismatches reflect specific meaning effects, this study marks an important step in this direction in that it identifies a training procedure, a testing procedure, and a method of analysis for evaluating such differences. Furthermore, such procedures may eventually allow investigators to test more directly various theories of early word development that are concerned with the nature of the early word meanings acquired by young infants (Bates, 1979; Clark, 1983; Nelson, 1972). For example, if word meanings are restricted to perceptual features of an object, one might accordingly expect changes in the AER to occur only to differences in the perceptual features trained and not to differences in the functions of the objects or to the different experiences that the infant has had with the object. These procedures could also be used to test more directly the extent of the infant's concept for a name. By presenting the object name and then pairing it in a match/mismatch condition with different objects, the investigator may be able to identify which name labels the object, simply by inspecting these two components of the AER identified in our study. If the region between 20 and 100 ms contains a negative peak and if the 520- to 600-ms region is characterized by a large positive deflection, then the object

would not be labeled with that word by the infant. In this manner, it may be possible to map out aspects of the infant's semantic space. In addition, it may be possible to extend these procedures to study the emergence of syntactic and pragmatic relationships. Although little has been accomplished to date in this area, the possibilities for applying these procedures to other areas of language development seem limitless. Moreover, these techniques with further elaboration may be applicable to other populations such as brain-damaged children and adults who are not able to respond behaviorally to questions concerning their own understanding of words or instructions (Molfese et al., 1990; Molfese, Morse, & Cornblatt, 1990).

Controls for Confounding Factors

Our results do not appear to be due to confounds related to acoustic stimulus differences or to practice effects. In this study, half of the infants were trained that one name was associated with Object A and that a second name was associated with Object B; the other half of the infants received the opposite pairings. Furthermore, averaged brain responses combined responses to both CVCV stimuli so that averages were obtained to the match and mismatch conditions across stimulus sounds for each infant and not to the specific CVCV bisyllables. Consequently, the results presented here could not have been confounded across the group of infants by simple visual or auditory perceptual differences. In addition, our results would not appear to be the result simply of multiple exposures to the auditory materials, that is, practice effects. If the AERs reflected only frequency of experience due to repeated testing with the same stimuli, the exposure would be expected to have an equal effect on all conditions. In the mismatch condition, Object B was presented while the name for Object A was being heard. If it were simply a matter of practice effects, there should have been no AER discrimination across a match or mismatch condition because the objects had been practiced an equal number of times. The results, then, must reflect some specific relationship between the acquired meaning shared by the object and its name.

Lateralized and Bilateral Brain Responses

The pattern of an early bilateral response that is followed later in time by a lateralized change in the AER waveform has been noted consistently in studies of child and adult speech perception (Molfese, 1984; Molfese & Betz, 1988; Molfese & Hess, 1978; Molfese & Molfese, 1988). Molfese had previously and consistently reported such patterns of bilateral and lateralized responses in infants, children, and adults during speech sound discrimination tasks. Our data extend this finding to semantic related studies and suggest that the processing of a variety of different types of materials during both speech perception and word discrimination involves multiple levels within the nervous system and not simply a left- or right-hemisphere response. Such an interpretation certainly appears in keeping with findings from brain-damaged populations (Gainotti et al., 1976; Pizzamiglio & Appicciafuoco, 1971; Riedel, 1981).

As in previous studies that have used electrophysiological measures and multivariate procedures to study perception, hemisphere effects were noted. However, not all match effects were restricted to the left hemisphere. Instead, at some points in time the electrophysiological activity displayed by the two hemispheres is similar and at other times dissimilar. Such effects appear inconsistent with previous arguments that general state differences between the hemispheres are activated when the infant is engaged in a cognitive or language task (Kinsbourne, 1974; Kinsbourne & Hicks, 1978). Instead, these findings serve to reinforce the notion that brain responses to speech and language materials are multidimensional and involve a variety of different processes (some of which are lateralized and some of which are not) that interact and occur both sequentially and in a parallel fashion during the processing of information. Such results obviously strain our attempts to simplify brain processes as restricted to either the left or right hemispheres.

Summary

In this study, we used auditory evoked response procedures to study the acquisition of the comprehension of names for different objects in 14-month-old infants. Changes in two portions of the AER waveforms were found to reliably occur when a name was correctly paired with the object it named versus incorrectly paired with the object it named. These procedures provide another means to study early receptive language development in infant populations.

References

Ali, L., Gallagher, T., Goldstein, J., & Daniloff, R. (1971). Perception of coarticulated nasality. *Journal of the Acoustical Society of America, 49,* 538–540.

Baker, E., Blumstein, S., & Goodglass, H. (1981). Interaction between phonological and semantic factors in auditory discrimination. *Neuropsychologia, 19,* 1–15.

Bates, E. (1979). *Emergence of symbols.* San Diego, CA: Academic Press.

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

Bates, E., Bretherton, I., Snyder, L., Shore, C., & Volterra, V. (1980). Vocal and gestural symbols at 13 months. *Merrill-Palmer Quarterly, 26,* 407–423.

Benedict, H. (1975, April). *The role of repetition in early language comprehension.* Paper presented at the meeting of the Society for Research in Child Development, Denver, CO.

Bloom, L., Lahey, M., Hood, L., Lifter, K., & Fiess, K. (1980). Complex sentences: Acquisition of syntactic connectives and the semantic relations they encode. *Journal of Child Language, 7,* 235–261.

Blume, W. T., Buza, R. C., & Okazaki, H. (1974). Anatomic correlates of the ten–twenty electrode placement system in infants. *Electroencephalography and Clinical Neurophysiology, 36,* 303–307.

Brown, W. S., Marsh, J. T., & Smith, J. C. (1979). Principal component analysis of ERP differences related to the meaning of an ambiguous word. *Journal of Electroencephalography and Clinical Neurophysiology, 46,* 706–714.

Bryden, M. P. (1982). *Laterality.* San Diego, CA: Academic Press.

Callaway, E., Tueting, P., & Koslow, S. H. (1978). *Event-related brain potentials and behavior.* San Diego, CA: Academic Press.

Cattell, R. B. (1966). The Scree Test for the number of factors. *Multivariate Behavioral Research, 1,* 245.

Chapman, R. M., McCrary, J. W., Bragdon, H. R., & Chapman, J. A. (1979). Latent components of event-related potentials functionally related to information processing. In J. E. Desmedt (Ed.), *Progress in clinical neuropsychology: Vol. 6. Cognitive components in cerebral event-related potentials and selective attention* (pp. 36–47). Basel, Switzerland: Karger.

Clark, E. V. (1983). Meanings and concepts. In Paul H. Mussen (Ed.), *Handbook of child psychology* (Vol. 3, 4th ed., pp. 787–840). New York: Wiley.

Daniloff, R., & Moll, K. (1968). Coarticulation of lip rounding. *Journal of Speech and Hearing Research, 11,* 707–721.

Dixon, W. J. (Ed.). (1987). *BMDP Statistical Software 1986.* Berkeley: University of California Press.

Donchin, E., Tueting, P., Ritter, W., Kutas, M., & Heffley, E. (1975). On the independence of the CNV and the P300 components of the human averaged evoked potential. *Journal of Electroencephalography and Clinical Neurophysiology, 38,* 449–461.

Eimas, P. D. (1974). Auditory and linguistic processing of cues for place of articulation by infants. *Perception & Psychophysics, 16,* 513–521.

Eimas, P. D., Miller, J. L., & Jusczyk, P. W. (1987). On infant speech perception and the acquisition of language. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 161–195). New York: Cambridge University Press.

Gainotti, G., Caltagirone, C., and Ibba, A. (1976). Semantic and phonemic aspects of auditory language comprehension in aphasia. *Linguistics, 154,* 15–28.

Gelfer, M. P. (1987). An AER study of stop-consonant discrimination. *Perception & Psychophysics, 42,* 318–327.

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.

Goodsitt, J., Morse, P., Ver Hoeve, J., & Cowan, N. (1984). Infant speech perception in a multisyllabic world. *Child Development, 55,* 903–910.

Harris, L. J. (1988). Pathological left-handedness: An analysis of theories and evidence. In D. L. Molfese & S. J. Segalowitz (Eds.), *Brain lateralization in children: Developmental implications* (pp. 289–372). New York: Guilford Press.

Ingram, D. (1978). Sensorimotor intelligence and language development. In A. Lock (Ed.), *Action, gesture, and symbol: The emergence of language.* San Diego, CA: Academic Press.

Jasper, H. H. (1958). The Ten–Twenty Electrode System of the International Federation of Societies for Electroencephalography: Appendix to Report of the Committee on Methods of Clinical Examination in Electroencephalography. *Journal of Electroencephalography and Clinical Neurophysiology, 10,* 371–375.

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

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika, 23,* 187–200.

Kamhi, A. G. (1986). The elusive first word: The importance of the naming insight for the development of referential speech. *Journal of Child Language, 13,* 155–161.

Keppel, G. (1982). *Design and analysis: A researcher's handbook* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Kinsbourne, M. (1974). Direction of gaze and distribution of cerebral thought processes. *Neuropsychologia, 12,* 279–281.

Kinsbourne, M., & Hicks, R. E. (1978). Mapping cerebral functional space: Competition and collaboration in human performance. In M. Kinsbourne (Ed.), *Asymmetrical function of the brain.* New York: Cambridge University Press.

Kuhl, P. K. (1985). Constancy, categorization, and perceptual organization for speech and sound in early infancy. In J. Mehler & R. Rox (Eds.), *Neonate cognition: Beyond the blooming, buzzing confusion.* Hillsdale, NJ: Erlbaum.

Lewkowicz, D., & Turkewitz, G. (1980). Cross-modal equivalence in early infancy: Auditory–visual intensity matching. *Developmental Psychology, 16,* 597–607.

Luria, A. (1973). *The working brain: An introduction to neuropsychology.* New York: Basic Books.

MacNeilage, P. F., & DeClerk, J. L. (1969). On the motor control of coarticulation in CVC monosyllables. *Journal of the Acoustical Society of America, 45,* 1217–1233.

Meltzoff, A., & Bornton, R. (1979). Intermodal matching by human neonates. *Nature, 282,* 403–404.

Miller, J. F., & Chapman, R. S. (1981). The relation between age and mean length of utterance in morphemes. *Journal of Speech and Hearing Research, 24,* 154–161.

Molfese, D. L. (1972). *Cerebral asymmetry in infants, children and adults: Auditory evoked responses to speech and music stimuli.* Unpublished doctoral dissertation, Pennsylvania State University.

Molfese, D. L. (1978a). Electrophysiological correlates of categorical speech perception in adults. *Brain and Language, 5,* 25–35.

Molfese, D. L. (1978b). Left and right hemispheric involvement in speech perception: Electrophysiological correlates. *Perception & Psychophysics, 23,* 237–243.

Molfese, D. L. (1979). Cortical involvement in the semantic processing of coarticulated speech cues. *Brain and Language, 7,* 86–100.

Molfese, D. L. (1980). The phoneme and the engram: Electrophysiological evidence for the acoustic invariant in stop consonants. *Brain and Language, 9,* 372–376.

Molfese, D. L. (1983). Event related potentials and language processes. In A. W. K. Gaillard & W. Ritter (Eds.), *Tutorials in ERP research: Endogenous components* (pp. 345–368). Amsterdam: North-Holland.

Molfese, D. L. (1984). Left hemisphere sensitivity to consonant sounds not displayed by the right hemisphere: Electrophysiological correlates. *Brain and Language, 22,* 109–127.

Molfese, D. L. (1988). *Evoked potential analysis and collection system* [EPACS©; computer program]. Carbondale, IL: Author.

Molfese, D. L. (1989). Electrophysiological correlates of word meanings in 14-month-old human infants. *Developmental Neuropsychology, 5,* 79–103.

Molfese, D. L. (1990). Auditory evoked responses recorded from 16-month-old human infants to words they did and did not know. *Brain and Language, 38,* 345–363.

Molfese, D. L., & Betz, J. C. (1988). Electrophysiological indices of the early development of lateralization for language and cognition and their implications for predicting later development. In D. L. Molfese and S. J. Segalowitz (Eds.), *Brain lateralization in children: Developmental implications* (pp. 171–190). New York: Guilford Press.

Molfese, D. L., & Hess, R. M. (1978). Speech perception in nursery school age children: Sex and hemispheric differences. *Journal of Experimental Child Psychology, 26,* 71–84.

Molfese, D. L., & Molfese, V. J. (1979a). Hemisphere and stimulus differences as reflected in the cortical responses of newborn infants to speech stimuli. *Developmental Psychology, 15,* 505–511.

Molfese, D. L., & Molfese, V. J. (1979b). Infant speech perception: Learned or innate. In H. A. Whitaker and H. Whitaker (Eds.), *Advances in neurolinguistics* (Vol. 4). San Diego, CA: Academic Press.

Molfese, D. L., & Molfese, V. J. (1980). Cortical responses of preterm infants to phonetic and nonphonetic speech stimuli. *Developmental Psychology, 16,* 574–581.

Molfese, D. L., & Molfese, V. J. (1985). Electrophysiological indices of auditory discrimination in newborn infants: The basis for predicting later language performance? *Infant Behavior and Development, 8,* 197–211.

Molfese, D. L., & Molfese, V. J. (1988). Right hemisphere responses from preschool children to temporal cues to speech and nonspeech materials: Electrophysiological correlates. *Brain and Language, 33,* 245–259.

Molfese, D. L., Morris, R. D., & Romski, M. A. (1990). Semantic discrimination in nonspeaking youngsters with moderate or severe retardation: Electrophysiological correlates. *Brain and Language, 38,* 61–74.

Molfese, D. L., Morse, P. A., & Cornblatt, R. (1990, February). *Comprehending what's in a name: Auditory evoked responses of receptive naming in traumatic brain injured adults.* Paper presented at the meetings of the International Neuropsychological Society, Orlando, FL.

Molfese, D. L., & Schmidt, A. (1983). An Auditory evoked potential study of consonant perception in different vowel environments. *Brain and Language, 18,* 57–70.

Morse, P. A. (1974). Infant speech perception: A preliminary model and review of the literature. In R. Schiefelbusch and L. Lloyd (Eds.), *Language perspectives: Acquisition, retardation, and intervention* (pp. 19–53). Baltimore: University Park Press.

Morse, P. A. (1979). The infancy of infant speech perception: The first decade of research. *Brain, Behavior, and Evolution, 16,* 351–373.

Morse, P. A., & Cowan, N. (1982). Infant auditory and speech perception. In T. Field (Eds.), *Review of developmental psychology* (pp. 32–61). New York: Wiley.

Murphy, C. (1978). Pointing in the context of shared activity. *Child Development, 49,* 371–380.

Nelson, C. A., & Salapatek, P. (1986). Electrophysiological correlates of infant recognition memory. *Child Development, 57,* 1483–1497.

Nelson, K. (1972). The relation of form recognition to concept development. *Child Development, 43,* 67–74.

Oldfield, R. L. (1971). The assessment of handedness: The Edinburgh Inventory. *Neuropsychologia, 9,* 97–113.

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

Pizzamiglio, L., & Appicciafuoco, A. (1971). Semantic comprehension in aphasia. *Journal of Communication Disorders, 3,* 280–288.

Ramsey, D., & Campos, J. (1978). The onset of representation and entry into Stage 6 of object permanence development. *Developmental Psychology, 14,* 79–86.

Retherford, K. S., Schwartz, B. C., & Chapman, R. S. (1981). Semantic roles and residual grammatical categories in mother and child speech. *Journal of Child Language, 8,* 583–608.

Riedel, A. (1981). Auditory comprehension in aphasia. In M. Sarno (Ed.), *Acquired aphasia* (pp. 215–279). San Diego, CA: Academic Press.

Rockstroh, B., Elbert, T., Birbaumer, N., & Lutzenberger, W. (1982). *Slow brain potentials and behavior.* Baltimore, MD: Urban & Schwarzenberg.

Ruchkin, D., Sutton, S., Munson, R., Silver, K., & Macar, F. (1981). P300 and feedback provided by absence of the stimulus. *Psychophysiology, 18,* 271–282.

Scheffé, H. (1959). *The analysis of variance.* New York: Wiley.

Segalowitz, S. J., & Cohen, H. (1989). Right hemisphere EEG sensitivity to speech. *Brain and Language, 37,* 220–231.

Spelke, E. (1976). Infants' intermodal perception of events. *Cognitive Psychology, 8,* 553–560.

Snyder, L., Bates, E., & Bretherton, I. (1981). Content and context in early lexical development. *Journal of Child Language, 8,* 565–582.

Stevenson, M., Leavitt, L. A., Roach, M. A., & Chapman, R. S. (1986). Mothers' speech to their 1-year-old infants in home and laboratory settings. *Journal of Psycholinguistic Research, 15,* 451–461.

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

Trehub, S. (1973). *Auditory-linguistic sensitivity in infants.* Unpublished doctoral dissertation, McGill University, Montreal, Quebec, Canada.

Wagner, S., & Sakovits, L. (1986). A process analysis of infant visual and cross-modal recognition memory: Implications for an amodal code. In Lewis P. Lipsitt (Ed.), *Advances in infancy research* (Vol. 4). Norwood, NJ: Ablex.

Wagner, S., Winner, E., Cicchetti, D., & Gardner, H. (1981). "Metaphorical" mapping in human infants. *Child Development, 52,* 728–731.

Received April 6, 1988
Revision received February 16, 1990
Accepted February 21, 1990 ■