INFANT BEHAVIOR AND DEVELOPMENT **13,** 343-354 (1990)

Infants Use Multimodal Information to Imitate Speech Sounds

MARIA LEGERSTEE
Dalhousie University

The role of audition and vision in eliciting early imitation of speech sounds was examined in this study. Thirty infants, 3 to 4 months of age, were presented with the vowel sounds /a/ and /u/. For one half of the infants, these sounds were paired with an adult who silently articulated the same vowel; for the other half, the adult articulated the opposite one. Only the infants who were exposed to matched auditory and visual information were observed to imitate the vowels. The results suggested that multimodal information is useful for the acquisition of speech sounds.

infant imitation intermodal vowels speech

In the beginning, to produce their native language, infants must imitate the speech sounds they hear around them. Although newborns have the potential to learn any human language (Eimas, 1978; Lieberman, 1984; Werker & Tees, 1984), not only are their first words part of the vocabulary of the target language, but even prespeech vocalizations have characteristics that resemble that language (de Boysson-Bardies, Sagart, & Durand, 1984; Stark, 1980; Weir, 1962).

In order to imitate, the infant must perceive and then reproduce the vocal sounds of the model. The speech young infants hear has multimodal qualities. When an adult talks, not only does the voice emanate directly from the face, but the lip movements match precisely the pattern of the spoken language (Sullivan & Horowitz, 1983). The perception of speech sounds by adults and infants has previously been shown to be influenced by both the auditory and the visual properties of the vocal act (Dodd, 1979; Grant, Ardell, Kuhl, & Sparks, 1986; Kuhl & Meltzoff, 1988; MacKain, Studdert-Kennedy, Spieker, & Stern, 1983; Summerfield, 1979). For example, in certain linguistic contexts, visual information correspond-

This research was funded by a research grant of the Social Sciences and Humanities Research Council of Canada (455-89-0117) and the Dalhousie Research Development Fund. I would like to thank Bruce Earhard and John Fentress for their comments and advice concerning the manuscript, Edward Yang of the Department of Human Communication Disorders for technical advice, Doug Williams for help with statistical analyses, Sabina Fells for coding and phonetical assistance, and Johanna Robertshaw and Elisabeth Walsh for assistance with the various phases of this research. Special thanks also to the mothers and babies who made this work possible and to the staff of the Grace Maternity Hospital who helped with the recruitment of participants.

Correspondence and requests for reprints should be addressed to Maria Legerstee, Department of Psychology, Dalhousie University, Halifax, Nova Scotia B3H 4J1.

ing to one phoneme combined with auditory information of another may lead to the perception of a third in children and adults (McGurk & MacDonald, 1976). In addition, 4- to 5-month-old infants seem sensitive to the visual and acoustic properties of speech (Kuhl & Meltzoff, 1982). In the latter study, infants 18 to 20 weeks old were tested in two experiments on the ability to detect correspondences involving two vowel sounds. In the first experiment, the infants were shown two filmed female faces, one would produce the visual component of the vowel sound /a/, and the other, the vowel sound /i/. A loudspeaker, located centrally, produced the auditory component of one of the vowel sounds. The results showed that infants looked significantly longer at the face that matched the sound. In the second experiment, the spectral information that identified the nature of the vowel was removed, while the amplitude and duration of the vowel sound remained. This time the performance of the infants fell to chance level. Thus, the general temporal synchrony of the visual and auditory components of the vowel was not sufficient to produce a match in infants.

With respect to the reproduction of the vocal act, matching of pitch prosody and phonetic structure has been previously documented in young infants (Kessen, Levine, & Wendrich, 1979; Kuhl & Meltzoff, 1982; M. Papousek, Papousek, & Bornstein, 1985). Furthermore, 2-month-olds are capable of prespeech mouth movements (Trevarthen, 1979) and are known to imitate other mouth movements such as tongue protrusion, mouth pursed, and mouth opening (Abravanel & Sigafoos, 1984; Fontaine, 1984; Jacobson, 1979; Legerstee, in press; Meltzoff & Moore, 1977, 1983; Reissland, 1988; Vinter, 1986).

Combined, the above evidence would suggest that by 3 months of age, young infants perceive various vowel sounds and are able to reproduce them. Although there have been previous suggestions of infant matching of vocalizations at 3 months of age (H. Papousek & Papousek, 1981; Pawlby, 1977; Piaget, 1962; Uzgiris, 1973, 1984), most have been based on observations of natural infant–mother interactions, where it is virtually impossible to determine whether true imitation has actually occurred (see Kuhl & Meltzoff, 1988; Meltzoff & Moore, 1977). In studies of imitation, it is essential to show that infants selectively reproduce the particular act modeled, rather than merely appear to have imitated as a result of general arousal or adult feedback. Indeed, Piaget (1962), who developed the most complete theory of the ontogeny of imitation, suggested that although early during the sensorimotor period (between 1 and 4 months of age) the infant may readily reproduce sounds previously imitated by a model, this should not be confused with imitation. Rather, the infant, unable to distinguish between self and other, perceives the vocalizations of the model as his/her own and is stimulated to repeat the same sound. Thus, the infant's imitation is in a way illusory; the adult may say /a/ and the infant may respond by saying /a/. However, the infant is not capable of switching from /a/ to /u/ under the influence of selective imitation.

Uzgiris (1973) supported Piaget's theory on imitation. Studying 12 infants longitudinally during natural face-to-face interactions, she found that all engaged in vocal contagion of cooing sounds (*aah* and *uuh*) by the age of 3 months. She stated that "though imitation as such does not occur, there is the pattern of acting in turn with the experimenter" (p. 602). Thus, when assessing for infant vocal imitation, controlled studies should be designed such that the sounds are in no way contingent on the infants' vocal responses.

Recently, Kuhl and Meltzoff (1982) indicated that they serendipitously observed vocal imitation while studying intermodal perception in 18- to 20-week-old infants in the study referred to earlier. Not only did the infants look longer at the face that articulated the nonvocal component of the speech sound they heard, but they imitated the vowel as well. Thus, infants who heard /a/ vowels responded with /a/ vowels, and those who heard the /i/ vowels responded with /i/ vowels.

However, Kuhl and Meltzoff (1982) did not vary independently the visual and auditory components involved in reproduction of the speech signal, and therefore it is not clear whether the infants were imitating the sound produced by the filmed faces or the mouth movements. The present study was designed to investigate this aspect. Consequently, we assessed whether 3- to 4-month-old infants were able to selectively imitate the vowel sounds /a/ and /u/ in a cross-target design. In order to examine the effects of vision and audition on the imitative responses of the infant, the auditory and visual components were presented to the infants separately in a matched and mismatched condition. In the first condition, the vowels were presented to one group of infants through headphones, while the experimenter mouthed the matching visual component of the sound in synchrony. In the second condition, another group of infants would hear either /a/ or /u/, but the experimenter would mouth in synchrony the mismatched visual component.

METHOD

Subjects

Of the 54 infants recruited, 39 infants achieved a quiet alert state (State 4; Wolff, 1966) and 30 infants completed the study (15 females). Data for 9 infants were lost due to experimental error. All infants were healthy, born at term (40 ± 2 weeks) and experienced vaginal deliveries. Their postpartum ages ranged from 3 to 4 months (*M* age = 101 days) and their birth weight ranged from 2620 to 4320 gms (*M* = 3538 gms). Parents were contacted through birth announcements in the local newspapers and from a maternity hospital in town. The infants came from middle-income families, and some had one or two siblings.

Stimuli

Because we were interested in examining the effects of sound and face separately on the ability of infants to imitate vowel sounds, we could not use the same adult to voice all the vowels to the infants. First, it would be impossible for the adult to execute the mismatched condition. Second, using an experimenter who would voice the vowels in front of the infant would mean the voice would be recorded simultaneously on the audiotape with the voice of the infant. This would make it impossible for the coders to remain unbiased as to what sound was being modeled. Therefore, we produced an audiotape on which the voices of four English-speaking women were recorded with the use of a high-quality tape recorder (Sony 5000A). Two female experimenters were used to produce the visual component of the vowel sounds in the matched and mismatched conditions. The reason we used more than one voice and more than one face was to introduce slight variations in the pitch and loudness of the production of the stimulus. This method would not allow infants to match face and voice for characteristics other than the vowel sound. Kuhl and Meltzoff (1988, p. 245) also used a series of facial and vocal productions "so that no idiosyncratic feature of a single articulation could influence the detection of the match."

All vowels were of the same duration. They lasted 1 s and were separated by a 2-s interval. Each sequence of 5 /a/'s or 5 /u/'s lasted 15 s and was followed by a 15-s response interval. Three sequences of 5 /a/'s (modeling and response intervals) and three sequences of 5 /u/'s (modeling and response intervals) were recorded on the audiotape. Thus, total duration of one vowel condition was 90 s. The sounds and response times were recorded on a tape recorder, and the sounds were played to the infant with the use of an amplifier (Sony Integrated Amplifier TA-88), at a hearing level of approximately 60 dB. The sounds were presented simultaneously to the experimenter and to the infant through high-quality earphones (Telephonics, TdH-39, Supra-aural). Prior to presentation of the auditory and visual stimulus to the infants, the experimenters had practiced mouthing the visual signal in synchrony with the auditory signal for the matched and mismatched conditions until they had achieved 100% agreement (agreement/agreement + disagreement) over at least six consecutive matched and mismatched trials. The data of four infants were lost during testing because of discrepancies between the onset of the visual and auditory stimuli.

Procedure

All infants were observed in an infancy laboratory at the university. Each infant was seated in an infant seat facing one of the two experimenters who wore headphones and had a video monitor in her visual field. She sat approximately 30 cm or less from the infant's face. To prevent elicitation of a rooting response in the infant, she held the earphones close but not against the ears of the infant.

Following Kuhl and Meltzoff (1982), the experimental procedure included two phases, a familiarization phase and an experimental phase. During the familiarization phase, all infants saw the visual component of the /a/ and /u/ sound two times. This was done to familiarize the infants with the experimental setting. During the subsequent experimental phase, one half of the infants received the matched condition and the experimenter moved her lips in synchrony and matched the visual component of the sound being presented. The other half of the infants received the mismatched condition and the experimenter presented the visual component of one vowel (e.g., /u/) while the infant heard the other vowel (e.g., /a/), and vice versa. The experimenter maintained the same position during the modeling and response phases of the condition. Throughout the sessions, the experimenter kept a pleasant expression on her face. The experimenter was surrounded by white walls and a white curtain that blocked the entrance, and her face was illuminated by a 40-watt spotlight 20 cm above and 50 cm behind the infant. During testing, the other lights were turned down. A video camera (Hitachi, HV 730 BW) with a zoom lens (V 12.5-75 M3) was used to film the infants. This white-colored camera was mounted on the wall, behind and above the left shoulder of the experimenter and out of the visual field of the infant. It was remotely controlled by a research assistant from the adjacent operating room. This allowed for continuous filming of the infants when changes in head position occurred.

All infants were observed shortly after feeding to avoid behaviors attributed to hunger. They were presented with the stimuli when they had achieved a state of alert inactivity. The sound, the face, the condition (matched or mismatched), and the order of the visual component of the sound (either mouth opening or protruding mouth) shown during the familiarization period were counterbalanced.

Scoring

Six infant dependent variables were coded: for sounds—/a/-like sounds, /u/-like sounds, and "other" sounds; for mouth movements—wide mouth open, pursed mouth open, and "other mouth open" movements.

The frequencies of the wide mouth open movements from a closed or semi-closed position, of the pursed mouth open movements from a closed or semi-open position, and of other mouth open movements that did not fall into the two previous categories were scored. Mouth opening associated with yawning, tonguing, and spitting were not scored.

To achieve reliable coding for vocal sounds, the perceived qualities of /u/-like sounds and /a/-like sounds were used. Thus, the sounds were identified from a social-perceptual perspective. Because pitch and loudness of the vowel sounds /a/ and /u/ can be identical, spectrographic analyses of these sounds are not possible. Instead, the coder was trained by a phonetically trained experimenter to identify those sounds that sounded like /u/'s and those that sounded like /a/'s. These vocal categories have been found to be part of the 3-month-old infant's repertoire (Lieberman, 1980; Oller, 1981; Stark, 1980), and they can be perceived reliably by phonetically trained coders (Kuhl & Meltzoff, 1988; Oller, 1981; Stark, 1980; Uzgiris, 1973). Sounds not belonging to the /a/ and /u/ categories were also transcribed and labelled "other" sounds. Physiological sounds (fussing, sneezing, coughing, smacking, tongue sucking, etc.) were not coded.

Transcription of Videotapes

Trained assistants coded in real-time the frequencies of the infant behaviors from the videotapes, using a 54-cm video monitor. Before beginning the coding process, the six stimulus conditions were assigned a segment number and then copied to a new tape in a random order. Thus, the coders were rendered blind to the type of condition (/a/ or /u/ sound, matched or mismatched). Because the sounds had been presented to the infant through headphones, the coder was unable to hear the modeled vowels. One coder was trained to code the visual responses for all infants in each group, and another coder was trained to transcribe the vocalizations. During coding of the vowel categories, the screen was covered, and during scoring of the visual components the sound was turned off. A phonetically trained experimenter coded the same behaviors in 20% of the infants. Cohen's kappa was .81 for the /a/ sounds. .80 for the /u/ sounds, .81 for other sounds, .87 for wide mouth open, .81 for pursed mouth open, and .80 for other mouth open movements.

RESULTS

A two-way analysis of variance (ANOVA) was conducted on each behavior dependent variable: /a/-like sounds, /u/-like sounds, other vowel sounds, wide mouth open, pursed mouth open, and other mouth open movements (see Huberty & Morris, 1989). Sound (/a/-like or /u/-like) was a within-subject variable, and condition (matched or mismatched) was a between-subjects variable. The results of these analyses are summarized below, and the mean frequencies of the responses are presented in Figure 1.

Infant /a/ Sounds

The two-way ANOVA (σ = .05) for the /a/ sound dependent variable revealed a main effect for Condition, $F(1,28)=13.27$, $p<.001$. Infants, in general, produced more infant /a/ sounds in the matched condition than in the mismatched condition. There was a main effect for Sound,

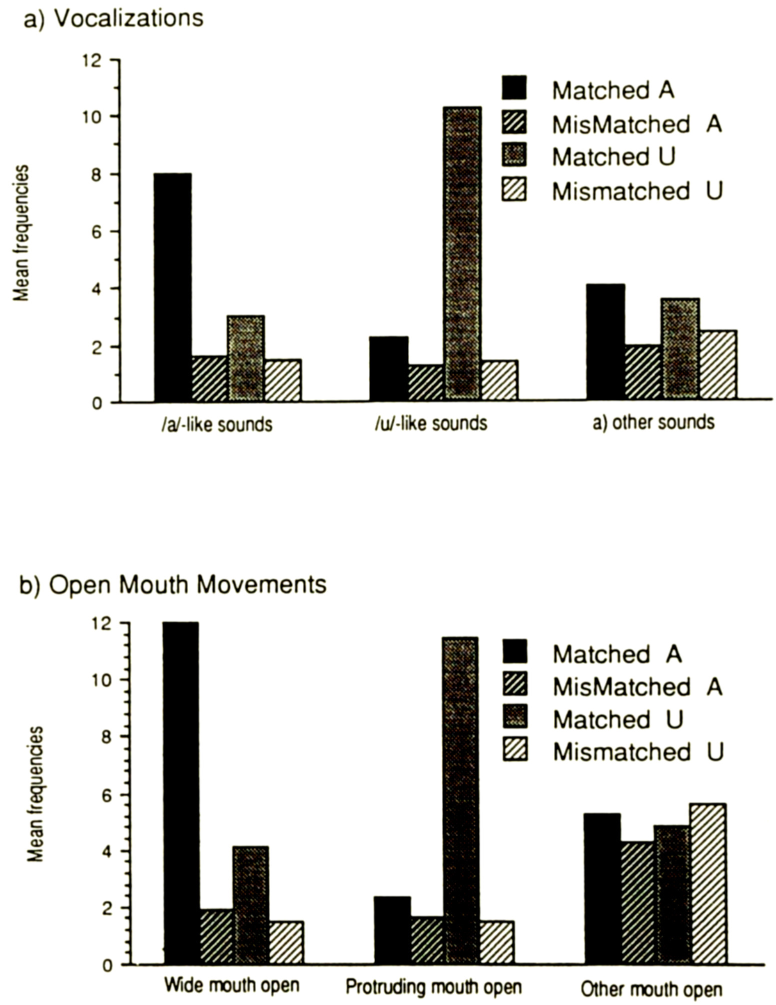

Figure 1. Mean frequency of (a) vocalizations and (b) mouth open movements produced in matched and mismatched conditions.

$F(1,28)=14.48$, $p<.001$. The infants produced more /a/ sounds when the /a/ sound was modeled than when the /u/ sound was modeled, and there was a significant Condition × Sound interaction, $F(1,28)=12.90$, $p<.001$. The interaction was a consequence of significantly more /a/ sounds produced when the /a/ sound was modeled in the matched condition than in the mismatched condition.

Infant /u/ Sounds

The two-way ANOVA for the /u/ sound dependent variable gave significant main effects for Condition, $F(1,28)=17.99$, $p<.001$, and Sound,

$F(1,28)=20.89$, $p<.001$, and a significant Sound × Condition interaction, $F(1,28)=19.45$, $p<.001$. Infants, in general, produced more /u/ sounds in the matched condition than in the mismatched condition, and more /u/ sounds when they heard the /u/ sound than when they heard the /a/ sound. The interaction was a consequence of significantly more /u/ sounds produced when the /u/ sound was modeled during the matched condition.

Other Infant Sounds

No significant effects resulted from the three-way univariate ANOVA.

Wide Mouth Open

The two-way ANOVA revealed significant main effects for Condition, $F(1,28)=25.66$, $p<.001$, and for Sound, $F(1,28)=14.13$, $p<.001$. In general, the infants produced more wide mouth open gestures when the /a/ sound was modeled than when the /u/ sound was modeled, and they produced more /a/ sounds during matched than mismatched conditions. A significant Condition × Sound interaction, $F(1,28)=18.54$, $p<.001$, indicated that the infants produced significantly more wide mouth open gestures when the /a/ sound was modeled during the matched condition.

Pursed Mouth Open

The two-way ANOVA revealed significant main effects for Condition, $F(1,28)=21.26$, $p<.001$, with more pursed mouth gestures produced during the matched condition, and a main effect for Sound, $F(1,28)=13.58$, $p<.001$, with more pursed mouth open gestures produced when the /u/ sound was heard than when the /a/ sound was heard. The interaction between Condition and Sound was also significant, $F(1,28)=14.79$, $p<.001$, indicating that for the modeled /u/ sound, infants produced more pursed mouth open gestures in the matched condition than in the mismatched condition.

Other Mouth Open Movements

No significant effects resulted from the three-way univariate ANOVA.

DISCUSSION

The results of the present study indicate that infants as young as 3 months of age selectively imitate the vowel sounds /a/ and /u/ produced by a human model. The data confirm and extend the findings by Kuhl and Meltzoff (1982), who found that 4.5-month-old infants were more prone to produce /a/ sounds when they heard /a/ sounds and /i/ sounds when they heard /i/ sounds. The current study used a different procedure. Instead of films, real faces were used, and instead of /a/ and /i/ vowels, /a/ and /u/ sounds were modeled.

Our results provide new evidence on the mechanism by which reproduction of vowel sounds is achieved in 3-month-old infants. Whereas in the Kuhl and Meltzoff (1982) study the infants saw the matched and mismatched faces simultaneously, the infants in this study were presented with the two conditions separately. This allowed for the examination of the effects of the auditory and visual components of the stimulus on the vocal imitations of infants. Our study showed that when presented with speech sounds by an adult, 3- to 4-month-old infants use both the auditory signal and its visual identifier to imitate. Thus, multimodal information is useful in the production of speech sounds in young infants.

It should be noted that although we found that the infants produced more /a/-like and /u/-like vowels in the matched condition than in the mismatched condition, this only meant that if both the visual and auditory components of the speech signal were present and matched, they would aid the infants to imitate; if they were incongruent, they interfered. This lends support to the suggestion that the infant has come into the world with "a set of flexible strategies for discovering the multimodal properties of objects and events" (Spelke, 1976, p. 554). The present paradigm did not allow us to infer whether the infants needed both the auditory and visual input to reproduce the sounds, or whether only the auditory stimulus would be sufficient for reproduction to occur. These questions should be studied in an experimental paradigm in which separate groups of infants are presented with either an auditory stimulus or its visual identifier, as well as with the matched and mismatched conditions. Although a recent study (Lefkowitz, 1988) has shown auditory dominance in 6-month-old infants presented with abstract novel multisensory stimulus compounds in a habituation paradigm, it still needs to be determined whether this sensory dominance holds when infants are presented with more natural stimuli, such as human speech, to which they have been exposed both visually and auditorily from birth. It is possible that for 3- to 4-month-old infants, there is primitive sensory unity for objects and events (cf. Gibson, 1969), but that with age, and depending on the stimulus, one modality may dominate over another in multimodal situations (Trehub & Thorpe, 1989; Welch & Warren, 1980).

The infants' gestural imitation was also aided by a sound that matched the visual component. The infants reproduced more wide mouth open gestures and mouth open pursed gestures when sound matched the shape of the mouth than when it did not. Meltzoff and Moore (1977, 1983) demonstrated imitation of oral gestures in infants less than 1 month old. Because their mouthing gestures (mouth open and mouth pursed) produced silently by the model were similar to the articulatory posture adopted for the production of the /a/ and /u/ sounds, respectively, these data support the hypothesis that infants are capable of imitating some speechlike gestures produced in the absence of vocalizations.

In summary, two main conclusions can be drawn from these results. First, through employing a cross-target design, we have strong evidence that infants are able to imitate vowel sounds that they themselves can produce almost as early as they begin to produce sounds resembling those of speech. Second, imitation of speech sounds makes use of both auditory and visual information. Thus, the multimodal nature of speech sound is significant not only for speech sound perception but also for production of speech sounds. This evidence supports earlier studies on cross-modal speech perception in suggesting that speech is represented intermodally in the first few months of life (Kuhl & Meltzoff, 1982, 1988).

Furthermore, it is possible that the use of multimodal information in imitation of speech plays an important role in the acquisition of the native language. As suggested by Gibson (1969), the infants face a formidable task in learning meaningful language. Before they learn to map meaning onto objects in the environment they have to discriminate the various segments of the speech signal. Intermodal perception, the ability of the infant to perceive the environment in an integrated way, along several perceptual modalities (e.g., audition and vision) should facilitate the learning of the phonetic contrasts that exist in the native language (cf. Sullivan & Horowitz, 1983). In this way, the infant would have more than one source of information to aid in the production of a better match between what is heard and what is said. If so, and if the results generalize to other speech sounds, then these findings have important implications for the acquisition of language. Certainly, infants who lack visual information about speech will be at a disadvantage with respect to language acquisition, while for those infants who have impoverished auditory information, emphasis on the visual modality should aid in the acquisition of productive speech.

REFERENCES

Abravanel, E., & Sigafoos, A. (1984). Exploring the presence of imitation during early infancy. *Child Development, 55,* 381–392.

de Boysson-Bardies, B., Sagart, L., & Durand, C. (1984). Discernable differences in the babbling of infants according to target language. *Journal of Child Language, 11,* 1–5.

Dodd, B. (1979). Lip reading in infants: Attention to speech present in- and out-of-synchrony. *Cognitive Psychology, 11,* 478–484.

Eimas, P.D. (1978). Speech perception in infancy. In L. Cohen & P. Salapatek (Eds.), *Infant perception: From sensation to cognition* (Vol. II). New York: Academic.

Fontaine, R. (1984). Imitation skills between birth and six months. *Infant Behavior and Development, 7,* 323–333.

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

Grant, K.W., Ardell, L.H., Kuhl, P.K., & Sparks, D.W. (1986). The transmission of prosodic information via an electrotactile speedreading aid. *Ear and Hearing, 7,* 243–251.

Huberty, C.J., & Morris, J.D. (1989). Multivariate analysis versus multiple univariate analysis. *Psychological Bulletin, 105*, 302–308.

Jacobson, S.W. (1979). Matching behavior in the young infant. *Child Development, 50*, 425–430.

Kessen, W., Levine, J., & Wendrich, K.A. (1979). The imitation of pitch in infants. *Infant Behavior and Development, 2*, 93–99.

Kuhl, P.K., & Meltzoff, A.N. (1982). The bimodal perception of speech in infancy. *Science, 216*, 1138–1141.

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

Lefkowitz, D.J. (1988). Sensory dominance in infants: 1. Six-month-old infants' response to auditory-visual compounds. *Developmental Psychology, 24*, 155–171.

Legerstee, M. (in press). The role of person and object in eliciting early imitation. *Journal of Experimental Child Psychology.*

Lieberman, P. (1980). On the development of vowel production in young children. In G.H. Yeni-Komshian, J.F. Kavanagh, & C.A. Ferguson (Eds.), *Child phonology: Vol. 1. Production.* New York: Academic.

Lieberman, P. (1984). *The biology and evolution of language.* Cambridge, MA: Harvard University Press.

MacKain, K., Studdert-Kennedy, M., Spieker, S., & Stern, D. (1983). Infant intermodal speech perception is a left-hemisphere function. *Science, 218*, 1138–1141.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746–748.

Meltzoff, A.N., & Moore, M.K. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198*, 75–78.

Meltzoff, A.N., & Moore, M.K. (1983). Newborn infants imitate adult facial gestures. *Child Development, 54*, 702–709.

Oller, D.K. (1981). Infant vocalizations: Exploration and reflexivity. In R.E. Stark (Ed.), *Language behavior and early childhood.* Amsterdam: Elsevier North Holland.

Papousek, H., & Papousek, M. (1981). Musical elements in the infant's vocalization: Their significance for communication, cognition, and creativity. In L.P. Lipsitt & C.K. Rovee–Collier (Eds.), *Advances in infancy research* (Vol. 1). Norwood, NJ: Ablex.

Papousek, M., Papousek, H., & Bornstein, M.H. (1985). The naturalistic environment of young infants. In T.M. Field & A. Fox (Eds.), *Social perception in infants.* Norwood, NJ: Ablex.

Pawlby, S.J. (1977). Imitative interaction. In H.R. Schaffer (Ed.), *Studies in mother–infant interaction.* New York: Academic.

Piaget, J. (1962). *Play, dreams, and imitation in childhood.* New York: W.W. Norton.

Reissland, B. (1988). Neonatal imitation in the first hour of life: Observations in rural Nepal. *Developmental Psychology, 24*, 464–469.

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

Stark, R.E. (1980). Stages of speech development in the first year of life. In G.H. Yeni-Komshian, J.F. Kavanagh, & C.A. Ferguson (Eds.), *Child phonology: Vol. 1. Production.* New York: Academic.

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

Summerfield, Q. (1979). Use of visual information for phonetic perception. *Phonetica, 36*, 314–331.

Trehub, S.E., & Thorpe, L.A. (1989). Infants' perception of rhythm. *Canadian Journal of Psychology, 43*, 217–229.

Trevarthen, C. (1979). Communication and cooperation in early infancy: Description of primary intersubjectivity. In M. Bullowa (Ed.), *Before speech.* London: Cambridge University Press.

Uzgiris, I.C. (1973). Patterns of vocal and gestural imitation in infants. In L.J. Stone, H.T. Smith, & L.B. Murphy (Eds.), *The competent infant: Research and commentary.* New York: Basic Books.

Uzgiris, I.C. (1984). Imitation in infancy: Its interpersonal aspects. In M. Perlmutter (Ed.), *Minnesota Symposia on Child Psychology* (Vol. 17). Hillsdale, NJ: Erlbaum.

Vinter, A. (1986). The role of movement in eliciting early imitations. *Child Development, 57,* 66–71.

Weir, R.H. (1962). *Language in the crib.* The Hague: Mouton.

Welch, R.B., & Warren, D.H. (1980). Immediate perceptual responses to intersensory discrepancy. *Psychological Bulletin, 88,* 638–667.

Werker, J.F., & Tees, R.C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7,* 49–63.

Wolff, P. (1966). The causes, controls and organization of behavior in the neonate. *Psychological Issues, 5* (Monograph No. 17).

7 September 1989; Revised 15 January 1990 ■