To appear in M. E. Lamb and L. R. Sherrod (Eds.)
Infant Social Cognition. Erlbaum, 1980

BLANK
PP. 60

12

# 4 Perceptual Aspects of Social Knowing: Looking and Listening in Infancy

Elizabeth S. Spelke
Alexandra Cortelyou
*University of Pennsylvania*

An infant encounters human beings through rich and complex patterns of sensory information. He or she comes to recognize individual persons, and to perceive their actions and expressions, by looking, by listening, and perhaps even by touching, tasting, and smelling. More importantly, the infant comes to relate these sources of information to each other. He or she eventually perceives each audible, visible, and tangible person as a unitary object. There comes a time when a child knows, for example, that a certain voice belongs to a certain person. This chapter focuses on the development of auditory-visual perception of people. The discussion centers on three capacities: (1) the ability to attend to a person by looking and listening at the same time; (2) the ability to perceive that the person one sees is (or is not) the source of a concurrent pattern of speech; (3) the ability to learn that certain patterns of light and patterns of sound specify the same individual.

In this discussion, two themes are highlighted. The first theme is substantive. Infants appear to possess innately, or to develop quickly, remarkable abilities to perceive the actions and expressions of other persons. These abilities may facilitate infants' learning about people and social encounters. The second theme is methodological. The study of infants' capacities to perceive and learn about people has only begun. This enterprise might benefit greatly from a closer collaboration between students of perception and students of social interaction. A collaborative venture could enrich both fields. Perceptual psychologists might better appreciate the earliest human capacities if they observe infants in contexts that are social. When infants confront other people, they may reveal perceptual abilities that rarely come to light in nonsocial settings. In addition, social psychologists

might come to understand more about the growth of social and communicative competence if they attempt to characterize the stimulus information that is available to an infant in a communicative setting, the infant's sensitivity to that information, and the infant's consequent perception of people, actions, and relationships. This collaboration, difficult though it will be to achieve, might permit psychologists to turn with new force toward questions such as those we raise.

## PRELIMINARIES: PERCEIVING FACES AND VOICES

Before considering the infant's ability to relate visual and auditory information about a person, we first ask about the developing ability to perceive a face or a voice that is presented by itself. Although the evidence is somewhat contradictory (see Sherrod, this volume), it appears that infants attend to human faces and voices, detect some of their properties, and learn about them. In this section, we consider studies of attention, discrimination, and learning.

Infants attend to voices and faces from the beginning of life. Voices and voice-like sounds are among the most effective audible events for attracting the interest of newborn infants (Church, 1970; Eisenberg, 1976; Hutt, Lenard, & Prechtl, 1969; Turnure, 1971; Wolff, 1963). Similarly, a drawing of a face attracts the eye of a newborn infant in preference to a variety of other visual patterns (Fantz, 1963; Goren, 1975), whereas a real, animated face elicits strong interest and apparent attempts at communication in the first 2 months (e.g., Trevarthen, 1977; Wolff, 1963). Further evidence for the infant's interest in faces comes from studies of conditioning. Both faces and voices provide effective reinforcers for young infants. Infants of 2 months will learn to suck or turn their heads in order to see a person (e.g., Bower, 1966; Koch, 1967), whereas infants of various ages will learn to suck or press a lever in order to hear speech (e.g., Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Friedlander, 1968).

Infants are able to discriminate a number of the properties of a person's face or voice. Studies of speech perception reveal that infants of 4 months or less are sensitive to some of the distinctive sound patterns of human languages (cf. Eimas, 1975) and of inanimate sounds with certain characteristics of speech (Jusczyk, Rosner, Cutting, Foard, & Smith, 1977; Jusczyk, Walley, & Pisoni, 1979). Furthermore, infants of 3 to 6 months are sensitive to the pitch of a speech sequence (Kessen, Levine, & Wendrich, 1979), 1-month olds respond to its intonation (Mehler, Bertoncini, Barriere, & Jassik-Gerschenfeld, 1978), infants of 9 months, and perhaps younger, perceive its affective tone (Lieberman, 1967), and 2-month olds appear to detect its time-spanning structure of rhythm or rhyme (Horowitz, 1974).



Studies of face-to-face interaction suggest that infants are also sensitive to some actions and expressions of the human face. For example, infants as young as 2 months evidently anticipate that a visible person will communicate with them, for they are disturbed by an immobile face (Tronick, Adamson, Wise, Als, & Brazelton, 1975; see also Bloom, 1977; Field, 1979), or by a person who suddenly becomes unresponsive (Brazelton, Koslowski, & Main, 1975; Fogel, Diamond, Langhorst, & Demos, 1979). Even when the mother is actively communicating, the infant seems to be sensitive to a lack of contingency in the exchange. Signs of distress have been observed in 2-month olds whose mothers switched their attention to another adult while continuing to face the infant (Trevarthen, 1977). Infants may also discriminate and imitate certain specific actions of the lips, tongue, and mouth (Church, 1970; Meltzoff & Moore, 1977), although the existence of early imitation has been questioned (Hamm, Russell, & Koepke, 1979; Hayes & Watson, 1979). Finally, studies of infants' responsiveness to an active adult or to a still photograph suggest an early appreciation of some of the affective expressions of another person (Barrera, 1979; Kreutzer & Charlesworth, 1973; LaBarbera, Izard, Vietze, & Parisi, 1976; Young-Browne, Rosenfeld, & Horowitz, 1975).

These investigations suggest that infants are sensitive to some of the temporal and spatial information that characterizes human utterances and expressions. Infants also have a further ability assuring that their capacity for social perception will grow: They are able to learn rapidly about human faces and voices. Infants are capable of at least two kinds of perceptual learning. First, they learn about some of the distinctive characteristics of human speech and human physiognomy. Infants learn quickly about some of the sound patterns of their own language (Eilers, Gavin, & Wilson, 1979; Streeter, 1976). They may also learn about some of the characteristics and actions of people (e.g., Caron, Caron, Caldwell, & Weiss, 1973; Gibson, 1969), although this ability is in need of further study with more natural displays. Second, infants come to recognize certain individuals by their voices and by their faces. One-month-old infants can recognize the voice of a parent (Mehler et al., 1978; see also Turnure, 1971), and they quickly learn to recognize the voice of an unfamiliar person (Horowitz, 1974). Recognition of tape-recorded voices can, nevertheless, cause some difficulty well into the school years (Mann, Diamond, & Carey, 1979). Infants also recognize the face of a parent by 1 month of age according to some criteria (Maurer & Salapatek, 1976), and by 2 to 4 months according to other criteria (Fitzgerald, 1968; Fogel, 1979). Infants learn rapidly to recognize the face of an unfamiliar person by 4 months of age, if the person moves expressively (Spelke, 1975). Until 5 or 6 months, however, they appear not to discriminate and recognize static photographs of people they do not know, even two people of opposite sex (Cornell, 1974; Fagan, 1972). Again, recognizing faces in photographs can cause some difficulties throughout childhood (Carey & Diamond, 1977).

In summary, infants explore, perceive, and learn to distinguish complex patterns of auditory and visual information about other human beings and their actions. Young infants appear especially sensitive to people whose actions and expressions change over time. They may extract visual and auditory information about a person who is communicating with them more easily than they can analyze the static, configural information in a photograph or a schematic drawing. The ability to perceive actions has, at least, been easier to document at young ages. Infants may also be more sensitive to the actions and the emotional expressions of a person than they are to a person's static features. For example, infants appear to discriminate between photographs of a face displaying different emotional expressions (e.g., Barrera, 1979) before they can discriminate between photographs of two different faces with neutral expressions (e.g., Fagan, 1972).

The sensitivity of young infants to some human actions and expressions raises an important possibility. Many aspects of a communicative act can be perceived either by looking or by listening. The same pattern of timing may characterize both a visual and a vocal expression of surprise; the same interactional contingencies may underlie both a visual and a vocal social exchange. Infants might be able to detect such invariances and perceive a relationship between the speech and the visible appearance of a person. We return to this possibility.

With this background, the chapter now focuses on the early development of coordinated looking and listening to people. Infants' exploration, perception, and knowledge of the people whom they see and hear are considered in turn.

## Exploring People by Listening and Looking

It is commonly observed that babies tend to look at a person when he or she speaks to them. And it is also widely believed, if less often tested by parents, that babies will lose interest in a person who looks at them silently and impassively. These observations suggest that infants attempt to explore people by looking and listening at the same time.

What exploratory patterns underlie the infant's looking and listening to a person? Some of these patterns may be quite general. Infants are known to look attentively at times and in places in which the sounds of inanimate objects are played. When young infants are presented with a sound, they often open their eyes (Haith, 1973), increase their attention to a concurrent visual display (Horowitz, 1974), and turn in the sound's direction (e.g., Muir & Field, 1979; Wertheimer, 1961). If looking is similarly affected by animate sounds, then these exploratory patterns will tend to increase the infant's looking to a person who speaks. But a further, more interesting exploratory pattern may also underlie the infant's visual attention to a speaking person. When an infant detects a voice, he or she may tend to look specifically for a face, in preference to other displays that are visible at the same time and place.



This exploratory pattern would depend on an ability to relate auditory and visual information about a person in a special way. An infant who looked specifically for a face when hearing a voice would seem implicitly to know that voices are the kinds of sounds produced by people. We now consider evidence that an infant who hears a voice looks for: (1) an object occurring at the same time as the voice; (2) an object occurring in the same spatial direction as the voice; and/or (3) an object that looks like a face.

Do infants tend to look attentively to a face at the time that a voice is presented? A recent experiment indicates that they do. Haith, Bergman, and Moore (1977) presented infants from 3 to 11 weeks of age with a mirror reflection of the face of an adult—the mother or a stranger—under three conditions: still, moving from side to side, and moving while talking. The infant's point of fixation was calculated every half second, and scanning patterns were assessed. Infants of 7 weeks and older tended to look on or near the person's eyes. In the presence of a voice, infants slightly increased their looking at the eyes, and their scanning of the face became distinctly more concentrated. The locations of different eye fixations varied less when the adult spoke than when she was silent. These observations suggest that the presence of the voice increased visual attention to a face.

There is some doubt about whether infants tend to look at a face in the spatial direction of a voice. In an experiment by Field, DiFranco, Dodwell, and Muir (1979), infants' looking to a peripherally presented face was unaffected by the spatial direction of a voice accompaniment. While 2 ½-month-old infants looked straight ahead, a still photograph of a person was presented to one side. At the same time, a voice was heard on the same or on the opposite side. Infants looked to the face in both conditions. They looked to the face equally quickly, regardless of whether the voice was played in the appropriate direction or on the opposite side. In contrast, looking at a face was affected by the direction of a voice in a different experiment (Spelke, 1978). Four-month-old infants were presented with two films, projected side by side. One film depicted a person speaking, whereas the other depicted moving, inanimate objects. A voice sound track, synchronized with the speaking person, was played either through a speaker centered between the films (at a 20° separation from the face), or through a speaker next to the film of the inanimate objects (at an 80° separation from the face). Infants looked at the person more when the spatial separation of voice and face was small than when it was large. The discrepancy between the findings of these two experiments could derive from a number of factors. For example, Spelke's infants were older, they viewed a moving, voice-synchronized face rather than a photograph, the moving face was continuously present, and the duration of infant's looking, rather than its latency, was the critical response measure.

In brief, young infants tend to look attentively at a face at the time that a voice is played. By 4 months, they also appear to look more at a face presented in the spatial direction of a voice. Do infants also tend to look specifically for

a face when they hear a voice? Two kinds of comparison are needed to answer this question. Experiments must compare looking at a face in the presence of a voice with looking at a face in the presence of some other auditory pattern. Similarly, experiments must compare the effect of a voice on looking at a face with its effect on looking at other visual displays. Several such studies have been conducted. None provides evidence for a specific effect of voices on looking at faces.

Hainline (1978) studied 1- to 3-month-old infants' scanning of a photograph of a face that was accompanied by a voice, by a tone varying in frequency, or by no auditory presentation. Unlike Haith et al. (1977), she found no effect of the voice on the amount of fixation on the eye region of the face. Hainline did report that eye fixations were more concentrated in the presence of a voice, but this effect was not specific to the voice. Scanning was even more concentrated in the presence of the tone. Inanimate sounds thus may influence infants' scanning of faces in much the way that voices do. Similarly, voices may affect infants' scanning of inanimate displays much as they affect scanning of faces. Mendelson and Haith (1976) observed that newborn infants scanned simple inanimate displays with more concentrated fixations when a display was accompanied by a voice than when it was presented silently. Furthermore, the infants tended to look at the central region of the display when the voice was presented from behind, as it had been in Haith et al.'s (1977) study. This central region would seem to correspond to the location of the eyes of the person in Haith et al.'s experiment.

Although young infants tend to look at a person when he or she speaks, these studies suggest that they do not do this because of any specific effect voices have on looking at faces. The presence of any interesting sound may lead to more concentrated looking at any interesting visual display. It might be argued, however, that these experiments do not provide a sensitive test of the effect of voices on looking at faces. In each study, only one visual display and one sound were presented to the infant at any given time. Although sounds may not affect visual attention differentially under these conditions, they may have more specific effects under other conditions. For example, if the infant were given a choice between two visual displays, one of them a face, he or she might attend to that display preferentially during times in which a voice was played. The same preference might not be observed during times in which the displays were accompanied by an inanimate sound.

At first glance, an experiment by one of us would seem to provide just such evidence (Spelke, 1976). Four-month-old infants were shown a film of a person speaking, side by side with a film of inanimate objects mving into contact. Different sounds, played through a central speaker, accompanied these events. When the films were accompanied by a voice synchronized with the speaking person, infants exhibited a reliable visual preference for that person. When the films were accompanied by percussive sounds synchronized with the impacts of the inanimate objects, infants no longer preferred the film of the speaking person. In fact, they tended to look more to the film of the inanimate objects.



Although this experiment would seem to show that voices have specific effects on looking at faces, there is an alternative explanation for the infants' looking patterns. The voice sound track was temporally synchronized with the movements of the speaking person, and the percussion sound track was synchronized with the movements of the inanimate objects. Thus, the infants might have looked at the face when they heard the voice because they detected this temporal synchrony, not because they related faces to voices in general. Infants of this age are known to detect the synchrony of an inanimate sound with simple, translatory movements of an object (Spelke, 1979b). In order to investigate whether they tend in general to look at a face when they hear a voice, one must present infants with a face and voice that are not temporally synchronized. We have attempted to do this in a preliminary experiment.

*Experiment 1: The Effects of a Voice on Looking at a Face.* The study used a preferential looking technique. Infants were presented with a film of a face, side by side with a film of inanimate objects. In this study, as in the study just described, the face film depicted the head and shoulders of a young adult woman playing "peekaboo." The other film depicted a hand holding a baton that struck two different toy percussion instruments in an irregular rhythm. Each event was filmed with a synchronized sound track, but the films and sounds were not presented at the same time. Infants first watched the two films with no sound. Then they heard the first sound track, unaccompanied by any film. The sound track was followed again by the two silent films, which were followed by the second sound track. The session closed with a final viewing of the two films. Because the sound tracks were not concurrent with the two films, sounds and visible movements were of course not temporally synchronized, although they may have been related in more subtle ways.[1] If infants tend specifically to look at a face when they hear a voice, then we

[1]An ideal test of the general effect of a voice on looking at a face would present a face and voice that were not temporally related in any way. This goal is not fully achieved by presenting a voice and a moving face that are shown out of synchrony. A person tends to speak in a characteristic rhythm, and a baby might detect this rhythm by looking and listening even if the person's voice and face were not synchronized. Thus, our study can provide no conclusive test of infants' knowledge about the specific relationship between a face and voice. We hoped nevertheless that its results would be suggestive.

expected them to look at the film of the person after the voice was played. We expected this preference to be diminished after the percussive sounds were played.[2]

Sixteen healthy, full-term infants, aged 3 months, 22 days to 4 months, 19 days (mean, 4 months, 3 days), participated in the experiment. They took part in five episodes. During the first, third, and fifth episode, they were presented with the peekaboo and percussion films, side by side, projected silently for 30 seconds per session. During the second and fourth sessions, infants were presented with the two sound tracks—one per session—for 50 seconds each. The films were projected on the left and right halves of a divided screen. For each infant, the lateral positions of the films remained constant throughout the experiment. The sound tracks were played through a speaker centered between the films. The left–right positions of the films and the order of the sound tracks were counterbalanced across infants. Looking was recorded throughout the session by two observers who were unaware of the position of each sound-related film. Reliabilities averaged 87%. The proportion of looking to each side of the screen was recorded for the three preference episodes.

The experiment provided only equivocal evidence that babies look at faces when they hear voices and not when they hear inanimate sounds. Infants tended to look longer to the person playing peekaboo after presentation of the voice than after presentation of the percussion sound, but this tendency was not significant. Preference for the peekaboo film averaged .58 following the voice presentation and .51 following the percussion sound, $t(15) = .88$. These results are complicated by what can only be a spurious effect. During the initial preference test, infants looked more to whichever film did not go with the sound that was about to be played. As Table 4.1(a) indicates, infants looked about equally to the two screens during the subsequent preference episodes, but these comparisons do not take account of the initial preferences. Table 4.1(b) accordingly presents the change in the proportion of looking to each film, from the first preference episode to the subsequent episodes. It is evident from Table 4.1(b) that the presence of the sounds affected looking during the preference test that followed the first sound track, but not during the test following the second sound track. Infants reliably increased their

[2]This prediction requires some elaboration. In different experiments, the presence of auditory or tactile information about an object has been found to influence subsequent preferential looking to that object in two mutually exclusive ways. First, infants have sometimes looked more to an object that they had just heard (Bahrick, 1979) or felt (Meltzoff & Borton, 1979). Second, infants have sometimes exhibited a "novelty preference," looking more to an object that they had *not* heard (Spelke, 1979a) or touched (Gottfried, Rose, & Bridger, 1977). The latter pattern is more often observed after prolonged familiarization with an object, and it is more commonly exhibited by older infants. Because we were presenting young infants with brief, interesting displays, we expected infants to prefer the more familiar object.



TABLE 4.1
The Effects of Sounds on Subsequent Visual Preferences

| | *Looking to first sound film* | *Looking to second sound film* | *Preference* | *t (15)* |
|---|---|---|---|---|
| (a) | | | | |
| Pretest | 9.1 | 15.5 | .36 | 2.33[a] |
| Post-Test 1 | 13.7 | 10.5 | .56 | 1.00 |
| Post-Test 2 | 12.4 | 11.8 | .49 | 0.13 |
| | *Change in Preference for Sound Film* | | *t (15)* | |
| (b) | | | | |
| Post-Test 1 | .18 | | 4.32[b] | |
| Post-Test 2 | −.12 | | −1.62 | |

[a] $p < .05$, two-tailed
[b] $p < .01$, two-tailed

looking to the film whose sound was the first to be played. This tendency was unaffected by the particular sound track presented.

Infants in the first study thus tended to look at a face after hearing a voice and to look at inanimate objects after hearing their sounds, but these tendencies were reliable in only one of several comparisons. Experiment 1 provides suggestive, but not conclusive, evidence for an ability to appreciate that faces belong with voices and that inanimate sounds belong with inanimate objects.

In summary, infants do appear to look at a person when they hear that person's voice, but this tendency seems to reflect primarily the tendency to look at an object when and where a sound occurs. There is no firm evidence that infants tend to look specifically at faces when they hear voices. It may be that infants do not appreciate that voices are the kind of sounds that are produced by objects with human faces. Alternatively, infants may implicitly know that faces and voices belong together, but they may fail to use this knowledge to direct their exploration in the situations that have been studied. New experimental procedures may uncover this knowledge. To date, however, knowledge in infancy of the general relation between faces and voices has not been conclusively demonstrated.

The tendency to look at a person when he or she speaks, whatever its perceptual or conceptual basis, is likely to aid the infant in learning about people and social exchanges. Under normal circumstances, young infants are likely to be faced directly by the persons who are talking to them, usually at close range. Infants will naturally look while listening, and they will thereby have the opportunity to learn about those persons through the interaction. But do infants realize that the person they see is the source of the sounds they hear? Do they perceive that a speaking face and its synchronized voice are

aspects of a single, unitary object? This is the question to which we now turn our attention.

## Perceiving People as Audible and Visible Objects

Adults usually perceive with little difficulty whether or not a person that they see is the source of a voice that they hear. This ability reflects a knowledge, perhaps a tacit knowledge, of certain auditory-visual relationships. There are at least three components to our stock of knowledge. First, we know that people with certain visible characteristics have voices with certain audible characteritics. Some of these characteristics are transient. Cheerful voices usually go with alert, smiling faces; sobbing voices usually accompany mournful faces. Other characteristics are persistent. Women have higher speech registers than men, and the old have less clear voices than the young. Thus, we more readily attribute a raspy, low-pitched, angry voice to a frowning elderly man than to the 4-year-old girl who smiles beside him. Second, we know that audible speech and visible speaking movements tend to be synchronized in certain regular ways. Visual and auditory information for speech is to some extent redundant. Because of this redundancy, adults find it easier to understand speech in a noisy environment if they can look at the speaker (Dodd, 1977; Sumby & Pollack, 1954) and some deaf people are able to understand speech by lip-reading. Our normal reliance on this redundancy may also explain the occasional discomfort caused by dubbed or poorly synchronized motion pictures. As adults, we may determine which of several people is speaking by detecting the synchrony of speech and speaking movements. Third, a speaking person is seen and heard in the same position in space. We may sometimes determine which visible person is speaking by localizing the voice. As every ventriloquist knows, however, adults can be fooled about a voice's location.

The sensitivity of infants to these auditory-visual relationships has received very little study. We know of no research asking whether infants can discover that a voice specifies a particular visible person by detecting the common spatial direction of the face and voice.[3] We also know of only one investigation of infants' ability to perceive a face-voice relationship by detecting a characteristic of a person by looking and listening: Walker (1980) has investigated infants' visual and auditory perception of a person's expressions of emotion. Finally, at the same time we began this chapter, we knew of no studies of infants' ability to determine whether a voice comes from a particular speaking person by detecting the temporal synchrony of speech

[3]The demonstration that infants will look at a face in the spatial direction of a voice does not, in itself, indicate whether they perceive the face they encounter to be related to that voice. Infants may simply localize a sound and turn to look at anything in its direction.



and speaking movements. One such study has since appeared (Dodd, 1979). We have, accordingly, conducted a second experiment. This section first descibes our investigation of infants' sensitivity to the synchrony of audible speech and visible expressive movements, and it briefly relates our own work to that of Dodd (1979). Next, it describes Walker's (1980) studies of infants' sensitivity to the common affective tone of a voice and a visually expressive face. All these investigations indicate that infants can sometimes determine which of two visibly speaking people is the source of an accompanying voice.

*Experiment 2: Detecting the Synchrony of Speech and Expressive Movement.* We confess that we approached this experiment with some trepidation. Detecting the synchrony of speech and expressive movements would seem to be a formidable task for a young infant. The temporal relationships between speech sounds and speaking movements are subtle and complex. Most articulatory movements are fully or partly invisible. Many of the movements that are visible could correspond to any of a class of speech sounds (see, for example, MacDonald & McGurk, 1978). Furthermore, only some movements of the face may correspond predictably to a stream of speech. A raising of the eyebrows, for example, may accompany any number of different sounds. In order to detect the synchrony of speech and facial movements, an infant would seem to need to attend selectively to certain facial movements and to certain aspects of the speech stream. The infant would also need to grasp the relationship between the relevant movements and sounds. Young perceivers would seem to be ill-equipped to accomplish these tasks. Their discrimination of speech sounds and facial movements is unlikely to be as good as the discrimination by adults. And infants, like adults, have been reported to look at the eyes of a speaker, not at the mouth (Haith et al., 1977). Thus, they would seem likely to miss any voice-synchronized movements that a speaking face exhibits.

Two considerations reduced our pessimism. First, the synchrony between speech and expressive movements would seem to be especially compelling in the speech of adults to infants. When a person talks to a baby, he or she tends to exaggerate facial expressions, to extend them in time and space, and to create especially full and redundant expressive displays (see Stern, Beebe, Jaffe, & Bennett, 1977). Many movements besides those of the articulators are probably synchronized with such a person's voice. Second, as our initial review indicated, the tendency to look at the eyes of a person evidently does not prevent an infant from detecting information about the person's actions and expressions. And it is just such actions that are likely to be detectable both by eye and by ear. The second experiment therefore tested infants' sensitivity to the synchrony of a person's speech and facial movements.

Four-month-old infants were presented with films of the faces of two young adult women. The women were not known to the infants. Each spoke

spontaneously, while facing the camera, as if she were addressing a baby. She greeted the baby, asked about his or her daily routines, attempted to elicit a smile, and so forth. Both women exhibited the temporally extended, exaggerated expressions characteristic of speech to infants. Each spoke in a bright, cheerful manner.

Throughout the experiment, the films of the two speaking people were presented side by side. One sound track was played at a time, through a centrally placed speaker. This display arrangement insured that only the temporal synchrony of speech and facial movements could tie each voice to one person. The voices and faces were not spatially related, each voice was played concurrently with both faces, and the particular face-voice pairings were not previously known to the infants.

The experiment made use of a procedure that one of us has used before, with inanimate objects (Spelke, 1979b). It consisted of two episodes within a single session. During the "preference episode," infants were presented with the films of both women projected side by side. They heard the synchronized voice of each woman in turn. Each voice has heard for one 100-second session through the central speaker. During the subsequent "search episode," the two films were projected silently while a light was flashed between them to attract the infant's attention. One voice, synchronized with the appropriate face and coming through the central speaker, was then played for 5 seconds. Eight such trials were given with each voice, the voices occurring in a random order. Throughout the experiment, each film was continuously projected on one side of the screen. The lateral position of the films and the order of the sound tracks were counterbalanced across infants.

Looking time to each film was continuously record by two observers, blind to the sound-object relationships. Their reliabilities averaged 89%. For each preference session, we calculated the proportion of total looking time that was devoted to the voice-synchronized face. For the search episode, looking was scored if the infant was not looking at either film at the time that the voice began. Seven scorable trials, on average, were given to each infant. We calculated four measures of the tendency to look to the person whose voice was played: the number of trials that the infant looked first to the appropriate and to the inappropriate person (*first look*), the number of trials on which he or she looked at all—first or second—to each person within the 5-second trial duration (*eventual look*), and the mean *latency* and *duration* of looking to each person. Trials in which an infant did not look to a given person received a latency score of 5 seconds and a duration score of 0 seconds. Further information about these procedures is given in Spelke (1979b).

Twenty infants from the Philadelphia area participated in the experiment. The infants ranged in age from 3 months, 21 days to 4 months, 21 days, and averaged 4 months, 8 days. They were healthy and full-term.



TABLE 4.2
Looking to Voice-Synchronized Faces

| | *Synchronized Face* | *Nonsynchronized Face* | *Preference* | *t (19)* |
|---|---|---|---|---|
| (a) Preference episode | | | | |
| Session 1 | 49.94 | 37.71 | .57 | 1.49 |
| Session 2 | 44.40 | 32.56 | .58 | 1.76[a] |
| Total | 94.34 | 70.27 | .58 | 3.90[b] |

| | *Synchronized Face* | *Nonsynchronized Face* | *t (19)* |
|---|---|---|---|
| (b) Search episode | | | |
| First Look | 4.15 | 2.75 | 2.19[a] |
| Eventual Look | 5.30 | 4.35 | 2.76[b] |
| Latency | 2.09 | 2.45 | 1.74[a] |
| Duration | 1.96 | 1.66 | 1.30 |

[a] $p < .05$, one-tailed
[b] $p < .01$, one-tailed

The results of this experiment appear in Table 4.2. Infants looked primarily to whichever woman they heard. They tended to prefer the "speaking" woman during each preference session. This tendency was reliable for the second session and for both sessions combined. During the search test, infants looked first, and looked eventually, more often toward the woman whose voice they heard. They also looked at her more quickly. They did not look at her for a significantly longer duration.

These results agree well with those of studies with inanimate objects (Spelke, 1979b). When infants are presented with two moving inanimate objects, accompanied by synchronized sounds during a preference episode and a search episode, they respond with looking patterns that are very similar to those observed in the present experiment. In particular, they tend to look for a sounding object on every measure of the search test except the duration measure.

The results of this experiment suggest that 4-month-old infants are sensitive to the synchrony of speech with some of the visible movements of the face. Infants can perceive a temporal relationship between the sight of a speaking person and the sound of his or her voice, even when the spatial location of the voice cannot guide their discovery of the face-voice relationship. By detecting the synchronizing of speech and facial movement, infants can determine whether a particular voice belongs to a particular visible person.

The findings of this experiment have been corroborated in an experiment by Dodd (1979) and in ongoing research by Walker (1980). Dodd presented

an unfamiliar, speaking person to infants who ranged in age from 10 to 16 weeks. Each infant viewed the person reciting nursery rhymes for four or more 60-second periods. During half of these periods, the persons' voice was presented in synchrony with the speaking movements of the face. During the remaining periods, the voice lagged behind the facial movements by 400 milliseconds. In both conditions, the voice was played through a loudspeaker presented behind the face. Visual attention to the face was scored in each of these conditions. As in our experiment, the presence or absence of synchrony had a pronounced effect on visual attention. Infants looked at the face more when it moved in synchrony with the voice.

Walker (1980) investigated the infant's sensitivity to the affective tenor of a communication, using a method very similar to ours. Her research focused on the ability to detect the common emotional tone of a voice and a visually expressive face. In two experiments, 7-month-old infants viewed films of a person engaging in a "happy" monologue and a person engaged in a "sad" monologue. The films were projected side by side and were accompanied by a centrally located, happy or sad voice. In one experiment, each voice was synchronized with the movements of the appropriate face. In a second experiment, the voice sound tracks were delayed by 1 second, such that the faces and voices were not temporally synchronized. Infants viewed both films at once while they heard each voice in turn through a central speaker. Infants in the first experiment looked longer to the face that corresponded to each voice during both voice episodes. Infants in the second experiment—in which the faces and voices were nonsynchronized—showed no preferences during the first voice episode, and then exhibited the appropriate preference during the second voice episode. Walker concluded that infants were sensitive both to the synchrony and to the emotional tenor of a voice and a face. When faces and voices are presented out of synchrony, infants may initially be confused by this discordance and, therefore, show no preferences during the first testing period. During the second period, they may become accustomed to the lack of synchrony and so begin to exhibit the preference for the face with the same emotional tone as the voice.

The findings of these experiments raise a question that bears special emphasis. What perceptual information does an infant use in detecting a relationship between speech and expressive movements? This question is not easily answered, for the speech and expressions of a person are richly redundant. In Walker's second experiment, for example, infants might have responded to expressions of emotion per se, or they might have responded to the common tempo of sound and movement when a person is happy or sad. In our second experiment and in Dodd's research, infants could have responded to temporal relationships of many kinds. For example, infants might have detected the synchrony of particular speech sounds and articulatory movements. Alternatively, they might have responded to the pattern of speech and pauses that characterizes any natural communication. Infants might have detected the simultaneous occurrence of speaking and moving, and of pausing and resting, in the people whom they saw and heard. As a third possibility, infants in both experiments might have responded to changes with time in the affective quality of a face and voice. Although the overall emotional tones of the speakers in our experiment were not noticeably different in our judgment, every person's speech is characterized by moment-to-moment changes in affective quality. These changes may be both audible and visible, and they may create a further temporal relationship between speech and facial movement. Infants might detect both auditory and visual information for a sudden expression of surprise or glee, and they may perceive the expression they see as related to the expression they hear. It will be difficult to distinguish experimentally between these possibilities. Psychologists have neither descriptive tools with which to characterize most of these relationships, nor adequate procedures for manipulating them. We think, however, that these problems must, and ultimately can, be faced. Gross descriptions of "face-voice synchrony" or "emotional expression" can yield to more specific descriptions of the auditory-visual relationships to which an infant responds.

Despite their limitations, the above studies have uncovered a remarkable perceptual ability. When two people are seen to speak, and a voice is centered between them, infants can discover which of the people is the source of that voice. However they do this, it is clear that they are sensitive to subtle and complex relationships in stimulation to the eye and ear. Infants have revealed a striking ability to perceive auditory-visual relationships in animate events.

## Learning About Specific Face-Voice Relationships

Thus far, we have considered the infant's exploration of a person while looking and listening at the same time, and the infant's perception of a relationship between a person's face and voice. Young babies appear both to explore people intermodally and to perceive face-voice relationships. They are therefore in a position to learn something new. They should be able to learn about further correspondences between the visible and audible characteristics of a person.

This last section focuses on one aspect of infants' learning. To an adult, the sight and sound of a familiar person are in some sense equivalent: They specify the same individual. Thus, one may determine that a certain acquantance has entered a house by watching her cross the room, listening to her call, or even by identifying her distinctive pattern of footsteps. One has obviously had to learn that these diverse visual and auditory patterns specify a single individual. How early in life does such learning begin? We discuss infants' learning about the audible and visible characteristics of the best-

known people in their lives: their parents. We ask if babies appreciate that the sound of a parent's voice goes with the sight of that parent's face.

A number of investigations have focused on the infant's knowledge of the relationship between the mother's face and her voice. Most of these use a "conflict" procedure. While the mother speaks to the infant, auditory and visual information are rearranged so that the mother's face and voice are spatially separated, or so that the mother's face is paired with a different person's voice. The infant's response to these rearrangements is observed. If infants are surprised or distressed, it is inferred that they know that the mother's face and voice normally go together. The findings of conflict studies have not always agreed, but evidence is now accumulating that young infants do expect the mother's voice to emanate from her face.

Infants studied by Aronson and Rosenbloom (1971) looked at the mother while listening to her synchronized voice first in her direction (straight ahead) and then 90° to one side. The 6-week-old infants were reported to be upset—in particular, to show marked "tonguing"—when the sound was spatially displaced. Two careful studies failed to replicate this effect (Condry, Haltom, & Neisser, 1977; McGurk & Lewis, 1974). McGurk and Lewis did, however, report one intriguing finding: Infants tended to look away from the mother's face more often in the spatial displacement condition than in the nondisplacement condition. Although this looking pattern could have reflected a simple tendency to look in the direction of the sound, a follow-up study suggested a more interesting interpretation. Lewis and Hurowitz (1977) presented 1- and 4-month-old infants with the faces and voices of the mother and of a female stranger. Infants saw one face and heard one voice at a time; the face and voice either came from the same person or from different people. In these conditions, the voice was always played from the direction of the face. Infants nevertheless looked around the room and away from a face more often if the face and voice were mismatched. Looking away may reflect a conflict reaction to the inappropriate face-voice pairing.

In the aforementioned studies, the faces and voices were synchronized when they were paired appropriately. In Lewis and Hurowitz's (1977) study, faces and voices were not synchronized when they were paired inappropriately. If infants in the latter study detected the incongruous arrangements, their detection could depend either on knowledge that the mother's face and voice go together or on the perception of face-voice synchrony. Three further experiments eliminated the possibility of a response to auditory-visual synchrony alone. Carpenter (cited in Bower, 1979) presented 2-week-old infants with the face and voice of the mother or a stranger. Faces and voices were paired appropriately or inappropriately. In all cases, the face moved in synchrony with its paired (tape-recorded) voice. Infants looked away from a face least when the mother spoke with her own voice. They were reported to avoid looking at either person's face when it was presented with the wrong person's voice. This gaze avoidance was taken to reflect a conflict reaction to a perceived discrepancy. The presence of confict would suggest that infants know about the normal relationship of the mother's voice and face.

Cohen (1974) presented 5- and 8-month-old infants with the mother and a stranger seated facing the infant, several feet apart. The tape-recorded voice of one of these adults was played through a loudspeaker. The loudspeaker was placed next to one of the people, who moved her lips in synchrony with the voice. These face-voice pairings could be either appropriate or inappropriate. Cohen recorded the latency and duration of the infant's looking to the person in the direction of the sound. She found such looking to be reduced, at 8 months of age, if either person appeared in the direction of, and moved in synchrony with, the other person's voice. There was no effect of matching or mismatching faces and voices at 5 months. Although Cohen discussed the infant's looking away from the location of the voice as evidence for conflict or discrepancy, she noted that infants may have looked toward the voice less in the mismatched conditions because they were looking toward the person to whom the voice really belonged.

A final study focused more directly on the exploratory activity of looking to a person whose voice is heard. Spelke and Owsley (1979) investigated infants' knowledge of the faces and voices of the mother and father. Infants from 3½ to 7½ months were presented with the two parents sitting side by side. While the parents remained motionless, their tape-recorded voices were heard in turn through a central speaker. Despite the absence of any face-voice synchrony, infants at all ages tended to look to the person whose voice was played. It was concluded that the infants had already come to appreciate which parent's voice belonged with each face. Spelke and Owsley repeated this procedure with 4-month-old infants in two conditions, the first with the mother and father and the second with the mother and an unfamiliar, adult woman. The results of the mother–father condition were as before, but the results of the condition with the two women were not. Contrary to expectations, infants did not look for the woman whose voice was played. In fact, they showed the opposite tendency, looking to the mother when the stranger's voice was played, and vice versa.

Sifting through these studies, it appears that babies do possess some knowledge about the relationship between a parent's face and voice. This knowledge seems to be best reflected in infants' patterns of looking. Except in one condition of Spelke and Owsley's study, infants have tended to look at the mother longer if they heard her voice than if they heard the voice of another person. They also appear to look longer at the father or at an unfamiliar woman if they hear that person's voice. These patterns may reflect an

interesting exploratory activity. Just as babies look in the direction of a sound, and just as they look at an object that is synchronized with a sound, they may look for an object that they know to be the source of a sound.

There is some dispute concerning the age at which infants acquire this knowledge. Carpenter found responses to the mismatched faces and voices of the mother and stranger at 2 weeks, but Cohen observed no such responses until 8 months. The findings of Lewis and Hurowitz and of Spelke and Owsley are more consistent with those of Carpenter, because they yield evidence for knowledge at the youngest ages tested (1 and 4 months). Despite these differences, it is clear that infants are learning about the audible and visible characteristics of familiar people quite early in life. Further studies might profitably focus on the nature of the learning capacities that underlie this accomplishment.

## CONCLUDING REMARKS

Human infants can perceive the communicative actions of others by looking and listening in a coordinated manner. We discussed three aspects of this coordination. First, young infants are usually able to look at a person when they hear the person's voice. They do this because of a very general tendency to look attentively at the time and in the direction in which a sound is heard. Infants may also tend to look specifically for a face when hearing a voice, but this possibility has received no clear experimental support. Because infants will tend to explore a person by looking and listening at the same time under normal viewing conditions, they are in a position to detect specific temporal relationships between the person's speech and facial movements. They are also in a position to learn about the relationship between the face and voice of a specific person. The rest of the chapter focused on these abilities.

We turned next to the ability to perceive a relationship between a face and its accompanying voice. Four-month-old infants are able to determine whether a visibly speaking person is the source of a voice by detecting some temporal relationship between the voice and face. Infants appear to be sensitive to the synchrony of certain characteristics of speech with certain visible movements of the speaker. They also appear to be receptive to the common emotional tenor of a visual and vocal communication. We do not yet know which of the many specific relationships between speech and expressive movements are detectable by infants.

Finally, we considered the infant's ability to learn about particular face-voice relationships. Such learning begins early in life. By 2 to 4 months of age, infants appear to expect the mother's face—and not the face of another person—to accompany the mother's voice. They may reveal this expectation through their exploration: When they hear the mother's voice they look



toward the mother's face and away from the face of some other person, even if the mother's voice is played from some inappropriate direction.

In summary, infants explore a person by looking and listening; they can detect the synchrony of the person's face and voice; and they can learn that the person's face and voice are specifically related. These findings are the fruits of rather recent research. Thus it is clear that psychologists have learned, and are learning, about the infant's perception of people, their communications, and their actions. But it is evident that there is a great deal more to learn. Psychologists have only begun to understand how infants coordinate auditory and visual information when they perceive animate events.

The understanding that has been achieved comes primarily from a heterogeneous collection of investigations, rather than a program of systematic research. We believe that further understanding of the infant's perception of others can come only from a serious and close collaboration between those who study infants as perceivers and those who study infants as social partners. We end by presenting some questions to challenge future investigations and by suggesting how those questions could be approached through such a collaboration.

The first question concerns the nature of the temporal relationships between speech and facial movements that infants are able to detect. What is the "synchrony" to which they are sensitive? To address this question, we need a workable description of the auditory and visual information available to infants in a communicative setting and a description of the manifold correspondences between these arrays of information. Then we may attempt to manipulate systematically the information available to infants in order to reveal what relationships they detect and, ultimately, how they detect them. Both tasks may be undertaken by a collaboration between the student of social communication—who seeks to describe and analyze patterns of activity in a social exchange—and the student of perception—who could attempt, through systematic experiments, to analyze the infant's sensitivity to those patterns.

The second question concerns the development of perception of auditory-visual relationships. What kinds of correspondences between auditory and visual information do infants detect, and which are detected earliest in development? Are infants first sensitive to the common spatial location of a face and voice, to the temporal synchrony of speech and movement, to certain specific correspondences between faces and voices, or simply to the repeated cooccurrence of a voice and a particular visible face? This question has theoretical importance. According to traditional theories of learning by association, the cooccurrence or "temporal contiuguity" of faces and voices should be the first relationship to which an infant responds. According to a theory that bases intermodal perception on the capacity to detect a common stimulus relationship in light and sound (Gibson, 1969), infants should first be

sensitive to either spatial or temporal relationships between a voice and a speaking face. An association theory might have difficulty accounting for the ease and speed with which 4-month-old infants come to appreciate intermodal relationships of complex kinds, such as the synchrony of speech and expressive movements. But, it is difficult to decide among these theories until we understand the newborn infant's sensitivity to all the available intermodal relationships. Here, again, is a topic on which students of perception and students of social interaction might work together. Careful and systematic observations of infants' social contacts with another person might provide the best suggestions about their abilities, and inabilities, to coordinate auditory and visual information about animate events. The newborn infant's social encounters need to be described and then manipulated to control the information that is available. Such manipulations may elucidate the auditory-visual relationships available to humans at birth, and our inborn capacities to detect these relationships.

The third question concerns the development of perception of people and animate events. This chapter is sprinkled with suggestions about the infant's ability to perceive the attributes, actions, and mental states of a person. But what do infants first perceive in other people, and how does the perception of people develop? On no other question, we believe, does a wedding of studies of social interaction with studies of infant perception hold greater promise. And on no other question has work in the two traditions been more at odds. Within an experimental tradition, most studies of the infant's perception of other people present young babies with photographs or schematic drawings of faces. An implicit or explicit assumption of this research is that infants first become sensitive to static features of faces, then to configurational properties, and later to actions, expressions, communications, and the mental states that give rise to them (see Caron et al., 1973; Fagen, 1972). The simplest displays for a psychologist to describe, it is apparently assumed, are likely to be the simplest displays for a young child to perceive. In contrast, students of social interaction observe infants in natural communicative interactions with active, expressive people. The infant's ability to respond to a person's actions, and even a person's intentions, is sometimes taken as a foundation for other basic perceptual, cognitive, and linguistic developments (see Bruner, 1977; Stern et al., 1977; Trevarthen, 1977). These traditions differ not only in their assumptions and their methods, but also in their findings. Studies of perception of face photographs or drawings suggest a slow development of sensitivity to other people (cf. Carey & Diamond, 1977). Studies of social interaction suggest a rather remarkable competence to perceive others and to participate in social exchanges early in life.

Which view may we trust? Are students of face perception missing the infant's greatest capacities by focusing on impoverished stimulus displays? Or are students of social interaction inflating their estimates of the infant's



abilities by focusing on situations in which the infant has an enormous prop: An adult partner who can regulate the social exchange? Does perception of people develop from detecting features to perceiving configurations, acts, and expressions, or is the developmental progression more nearly the reverse? These questions can only be answered by an approach that incorporates the best aspects of the observational and experimental traditions. Studies of the origins of perception of people should endeavor to examine infants in richly informative, social situations, but these studies must attempt to describe each situation and specify the potential sensory information available in it. Only thus can we analyze infants' sensitivity to that information and their ability to make sense of the social encounter.

These questions carry us far beyond the investigations that we have discussed. They point, however, to what we believe is an exciting direction for the study of infant social cognition. Studies of the infant as a perceiver and as a social agent might both benefit from investigations of the infant's perception of the actions, intentions, feelings, and social encounters of human beings. For the student of perceptual development, such research may yield glimpses of the infant's most advanced perceptual capacities, including capacities for coordinating sensory information from different modalities. For the student of social development, such research could shed light on some of the foundations of human social understanding.

## REFERENCES


Aronson, E., & Rosenbloom, S. Space perception in early infancy: Perception within a common auditory-visual space. *Science*, 1971, *172*, 1161-1163.

Bahrick, L. *Infants' perception of properties of objects as specified by amodal information in auditory-visual events*. Unpublished doctoral dissertation, Cornell University, 1979.

Barrera, M. E. The perception of facial expressions by the three-month-old infants. Paper presented at the meeting of the Society for Research in Child Development, San Francisco, March 1979.

Bloom, K. Operant baseline procedures suppress infant social behavior. *Journal of Experimental Child Psychology*, 1977, *23*, 128-132.

Bower, T. G. R. The visual world of infants. *Scientific American*, 1966, *215*, 80-92.

Bower, T. G. R. *Human development*. San Francisco: Freeman, 1979.

Brazelton, T. B., Koslowski, B., & Main, M. The origins of reciprocity: The early mother-infant interaction. In M. Lewis & L. Rosenblum (Eds.), *The effect of the infant on its caregiver*. New York: Wiley, 1975.

Bruner, J. S. Early social interaction and language acquisition. In H. R. Schaffer (Ed.), *Studies in mother-infant interaction*. New York: Academic Press, 1977.

Carey, S., & Diamond, R. From piecemeal to configurational representation in faces. *Science*, 1977, *195*, 312-315.

Caron, A. T., Caron, R. F., Caldwell, R. C., & Weiss, S. T. Infant perception of the structural properties of the face. *Developmental Psychology*, 1973, *9*, 385-399.

Churferential study of cognition in early childhood. In J. Hellmuth (Ew York: Brunner Mazel, Inc., 1970.
Cohedifferences in infants' attentional responses to face-voice incranger. *Child Development*, 1974, *45*, 1155-1158.
Cond& Neisser, U. Infant sensitivity to audio-visual discrepancy: A fait*the Psychonomic Society*, 1977,*9*, 431-432.
Cornemination of photographs of faces following redundant pro*erimental Child Psychology*, 1974, *18*, 98-106.
Doddhe perception of speech. *Perception*, 1977, *6*, 31-40.
Dodd.: Attention to speech presented in- and out-of-synchrony. *Cd 11*, 478-484.
Eilers,on, W. R. Linguistic experience and phonemic perception in infidy. *Child Development*, 1979, *50*, 14-18.
Eimas early infancy. In L. B. Cohen & P. Salapatek (Eds.), *Infant pe*ademic Press, 1975.
Eimas, P., & Vigorito, J. Speech perception in infants. *Science*, 1971, *17*.
Eisenb*etence in early life*. Baltimore: University Park Press, 1976.
Fagangnition memory for faces. *Journal of Experimental Child Ps*6.
Fantzborn infants. *Science*, 1963, *140*, 296-297.
Field,l. P., & Muir, D. Auditory-visual coordination in 2 ½-month-ol*and Development*, 1979, *2*, 113-122.
Fitzgeary reflex activity during early infancy and its relation to social an *Journal of Experimental Child Psychology*, 1968, *6*, 470-482.
Fogel*leave taking in one- to three-month-old infants*. Paper presented at/ for Research in Child Development, San Francisco, March 19
Fogelghurst, B. H., & Demos, V. *Alteration of infant behavior as a reson of maternal behavior*. Paper presented at the meeting of the Sd Development, San Francisco, March 1979.
Frey,deotaped sequences of their own and unfamiliar mothers. Uark University, Worcester, Mass., 1977. Cited in: Fabouti-Mis, I. C. The mother-infant communciation system. In I. C. U*on and communication during infancy*. San Francisco: Jossey-B.
Friedspeaker identity, voice inflection, vocabulary, and message retion of vocal reinforcement. *Journal of Experimental Child P*).
Gibs*otual learning and development*. New York: Appleton-Century-C
Gore*nate form preferences, and visually mediated head-turning in h*esented at the meeting of the Society for Research in Child D 1975.
Gott Bridger, W. H. Cross-modal transfer in human infants. *Child L*124.
Hainnges in visual scanning of face and non-face patterns by infants. *Jild Psychology*, 1978, *25*, 90-115.
Hait infants. In L. J. Stone, H. T. Smith, & J. B. Murphy (Eds.), *The c*: Basic Books, 1973.
HaitMoore, M. J. Eye contact and face scanning in early infancy. *S*
Hamm, M., Russell, M., & Koepke, J. *Neonatal imitation?* Paper presented at the meeting of the Society for Research in Child Development, San Francisco, March 1979.
Hayes, L. A., & Watson, J. S. *Neonatal imitation: Fact or artifact?* Paper presented at the meeting of the Society for Research in Child Development, San Francisco, March 1979.
Horowitz, F. D. Visual attention, auditory stimulation, and language discrimination in young infants. *Monographs of the Society for Research in Child Development*, 1974, *39*, (5-6, Serial No. 158).
Hutt, S. J., Lenard, H. G., & Prechtl, H. F. R. Psychophysiological studies in newborn infants. In L. P. Lipsitt & H. W. Reese (Eds.), *Advances in child development and behavior IV*. New York: Academic Press, 1969.
Jacobson, S. W., & Kagan, J. *Released responses in early infancy: Evidence contradicting selective imitation*. Paper presented at the meeting of the International Conference on Infant Studies, Providence, R.I., March 1978.
Jusczyk, P. W., Rosner, B. S., Cutting, J. E., Foard, C. F., & Smith, L. B. Categorical perception of non-speech sounds by two-month-old infants. *Perception and Psychophysics*, 1977, *21*, 50-54.
Juscyzk, P. W., Walley, A., & Pisoni, D. B. *Infants' discrimination of tone onset time differences: Some implications for voicing perception*. Paper presented at the meeting of the Society for Research in Child Development, San Francisco, March 1979.
Kessen, W., Levine, J., & Wendrich, A. The imitation of pitch in infants. *Infant Behavior and Devleopment*, 1979, *2*, 93-100.
Koch, J. Conditioned orienting reactions in two-month-old infants. *British Journal of Psychology*, 1967, *58*, 105-110.
Kreutzer, M. A., & Charlesworth, W. R. *Infants' reactions to different expressions of emotions*. Paper presented at the meeting of the Society for Research in Child Development, Philadelphia, March 1973.
laBarbera, J. D., Izard, C. E., Vietze, P., & Parisi, S. A. Four and six month old infants' visual responses to joy, anger, and neutral expressions. *Child Development*, 1976, *47*, 535-538.
Lewis, M., & Hurowitz, L. *Intermodal person schema in infancy: Perception within a common auditory-visual space*. Paper presented at the meeting of the Eastern Psychological Assocation, Boston, April 1977.
Lieberman, P. *Intonation, perception and language*. Cambridge, Mass.: MIT Press, 1967.
MacDonald, J., & McGurk, H. Visual influences on speech perception processes. *Perception and Psychophysics*, 1978, *24*, 253-257.
Mann, V. A., Diamond, R., & Carey, S. Development of voice recognition: Parallels with face recognition. *Journal of Experimental Child Psychology*, 1979, *27*, 153-165.
Maurer, D., & Salapatek, P. Developmental changes in the scanning of faces by young infants. *Child Development*, 1976, *47*, 523-527.
McGurk, H., & Lewis, M. Space perception in early infancy: Perception within a common auditory-visual space? *Science*, 1974, *186*, 649-650.
Mehler, J., Bertoncini, J., Barriere, M., & Jassik-Gerschenfeld, D. Infant recognition of the mother's voice. *Perception*, 1978, *7*, 491-497.
Meltzoff, A., & Borton, R. *Cross-modal transfer of shape in one-month-olds*. Paper presented at the meeting of the Society for Research in Child Development, San Francisco, March 1979.
Meltzoff, A. N., & Moore, M. K. Imitation of facial and manual gestures by human neonates. *Science*, 1977, *198*, 75-78.
Mendelson, M. J., & Haith, M. M. The relation between audition and vision in the human newborn. *Monographs of the Society for Research in Child Development*, 1976, 41 (4, Serial No. 167).
Muir, D., & Field, J. Newborn infants orient to sounds. *Child Development*, 1979, *50*, 431-436.