David J. Lewkowicz
New York State Institute for Basic Research in Developmental Disabilities
Staten Island, NY 10314

Infants' Response to the Audible and Visible Properties of the Human Face:

II. Discrimination of Differences between Singing and Adult-Directed Speech

Received 6 May 1997; accepted 19 August 1997

ABSTRACT: *Human infants' responsiveness to the audible and visible features of human faces was studied by habituating them to a person speaking a prepared script in an adult-directed manner and then administering a series of separate test trials where a person could be seen, heard, or seen and heard singing. When habituated to a female person speaking in an adult-directed manner and tested with a singing female 4, 6, and 8-month-old infants responded to the audible, visible, and bimodal changes, whereas 3-month-old infants only responded to the visual and bimodal changes. In contrast, when habituated to a male person speaking in an adult-directed manner and tested with a singing female, all age groups discriminated all three types of changes. These findings demonstrate that infants are responsive to differences between low- and high-prosody content inherent in both the facial and vocal characteristics of the human face and that, whereas responsiveness to the visible and bimodal features associated with differences between adult-directed speech and singing is present as early as 3 months of age, responsiveness to the audible features emerges between 3 and 4 months of age depending on whether gender differences are present as well.* *© 1998 John Wiley & Sons, Inc.* Dev Psychobiol 32: 261–274, 1998

Keywords: *human infants; intersensory perception; face and voice perception; auditory prosody*

Infants are highly responsive to a stylized type of speech that adults normally use when communicating with them. This type of speech is known as infant-directed (ID) speech and differs from adult-directed (AD) speech in that it is higher in pitch, has a wider pitch range, and has smooth and highly modulated pitch contours (Fernald & Simon, 1984; Papousek, Papousek, & Bornstein, 1985). Moreover, the overall tempo of ID speech is slower, the utterances are shorter, the pauses between utterances are longer, and there is greater prosodic repetition (Fernald, 1985; Papousek, Papousek, & Haekel, 1987; Snow, 1977;

Correspondence to: D. J. Lewkowicz

CCC 0012-1630/98/040261-14

Stern, Spieker, & MacKain, 1982). Studies have shown that infants prefer to listen to ID speech when given a choice between ID and AD speech (Cooper & Aslin, 1990; Fernald, 1985; Fernald & Kuhl, 1987; Papousek & Papousek, 1981; Werker & McLeod, 1989).

With a few notable exceptions (Pegg, Werker, & McLeod, 1992; Werker & McLeod, 1989; Werker, Pegg, & McLeod, 1994), the evidence indicating that infants prefer ID over AD speech is based either on the presentation of only the audible features of ID speech or on the presentation of the audible features of ID speech together with a static photograph of a human face (e.g., Fernald, 1993) or a checkerboard. Under normal circumstances, however, speech and many other social communicative events are dynamic, multimodal events. For example, when an infant interacts with a caretaker, the infant usually sees and hears the caretaker at the same time. Consequently, it is reasonable to ask whether the visible attributes of the type of exaggerated prosody inherent in ID speech contribute in some way to infant response and whether the interaction of the visible and audible features of exaggerated prosody play a role in responsiveness. This, in turn, raises a more general question: Does responsiveness to the multimodal representation of an event differ from responsiveness to the unimodal representation of the same event?

A good deal of evidence from studies of infants' responsiveness to multimodal events suggests that responsiveness to multimodal information differs from responsiveness to unimodal information. For example, adult subjects detect and identify speech that is simultaneously specified by its audible and visible features considerably better than speech that is specified just by its audible features (Massaro & Cohen, 1990; Summerfield, 1979). Also, studies have shown that adults' perception of auditory information can be dramatically affected by concurrent visual information in that the visual information actually can change what the subject hears. The two best known examples of these effects are the "McGurk effect" (McGurk & MacDonald, 1976), where subjects' perception of an audible syllable is changed by a conflicting visible syllable, and the "ventriloquism effect" (Jack & Thurlow, 1973) where subjects localize the site of a sound to be at the site of a visible event despite the fact that the two are spatially discrepant. Research with infants also has shown that they respond differently to multimodal than to unimodal information. For example, Lewkowicz (1988a, 1988b, 1992) has shown that whereas infants always can discriminate a simultaneous change in the audible and visible components of a compound auditory–visual stimulus, they often do not respond to a change in the separate, constituent components of the same compound stimulus. Moreover, as shown by Lewkowicz (1988a), infants do not respond to input in one modality when simultaneous input in a second modality is present but do respond to that same input when input in the second modality is absent. Finally, in addition to the modification, enhancement, or suppression effects that input in one modality can have on responsiveness to input in a second modality, simultaneous inputs in different sensory modalities sometimes can have multiplicative effects which cannot be predicted from the effects of the unimodal components comprising the compound stimulus (Stein, Meredith, & Wallace, 1994).

Although the differential effects of unimodal versus multimodal stimulation on responsiveness must be taken into consideration, two other factors must be taken into account when considering the effects of multimodal stimulation in a developing system. First, the different sensory modalities develop at different rates (Gottlieb, 1971). As a result, the auditory modality develops before the visual modality. This, in turn, means that the auditory modality is developmentally more advanced during early development and thus might be better in responding to the auditory component of an auditory–visual compound stimulus. The second factor that must be taken into account is that the different sensory modalities are specialized for the processing of different kinds of stimulus attributes. For example, the auditory modality is specialized for responding to the temporal attributes of stimulation whereas the visual modality is specialized for responsiveness to motion and the spatial attributes of stimulation (Kubovy, 1988). This also can lead to differential responsiveness to compound auditory–visual stimulation.

The functional consequences of the differential rates of sensory system development and of sensory specialization are illustrated by findings from a series of studies of infants' responsiveness to auditory–visual compounds by Lewkowicz (1988a, 1988b, 1992). In these studies, Lewkowicz found that the auditory aspects (i.e., temporal changes) of a compound auditory–visual stimulus blocked or dominated responsiveness to the concurrent visual attributes of the compound stimulus when the visual component was spatially static. In contrast, when the visual component was spatially dynamic (e.g., moving) the auditory component did not dominate responsiveness. In fact, when the visual component was spatially dynamic, infants responded more to it than to a concurrent auditory component (Lewkowicz, 1992). These types of effects make predictions regarding the effects of intermodal interactions a complex problem (Bushnell, 1994) and suggest that generalizations made on the basis of responsiveness to unimodal inputs may be misleading. Moreover, generalizations and characterizations of infants' responsiveness to unimodally presented features of exaggerated prosody may not extend to the multimodal case because (a) most communicative events are multimodal, (b) developmental and specialization factors play an important role in intersensory responsiveness, and (c) the interactive effects of simultaneous heteromodal inputs cannot be predicted from the effects of unimodal inputs. Findings from studies of infant attentional response to multimodally represented ID speech support these conclusions in that infants respond differently to multimodal representations of ID speech at different ages. For example, 7-week-old infants can discriminate between ID and AD speech on the basis of its audible characteristics alone, but they do not exhibit a preference for ID speech when it is simultaneously specified by audible and visible features (Pegg et al., 1992). In contrast, infants older than 4 months of age exhibit a visual preference for ID speech only when it is simultaneously specified by audible and visible attributes (Werker & McLeod, 1989). Findings such as these, as well as findings from other studies (Cooper & Aslin, 1990; Fernald, 1985), indicate that very young infants can discriminate and prefer ID speech when it is presented either alone or together with a non-face-like visual stimulus whereas older infants exhibit an attentional preference for ID speech best when it is specified concurrently in the auditory and visual modalities.

This developmental pattern parallels the pattern observed in animal studies of responsiveness to multimodal compounds. For example, when given a choice between a stuffed hen model of a conspecific adult female and a model of a female from another species, newly hatched bobwhite quail chicks will choose to go to the conspecific model. The basis for this behavior changes during the first few days after hatching. Lickliter and Virkar (1989) showed that the preference is first based solely on the auditory cues of the stuffed hen at 24 and 48 hr following hatching but that by 72 hr after hatching the chicks' preference is based on both visual and auditory maternal cues. What makes this finding particularly interesting is that this developmental pattern is dependent on early experience. If chicks are deprived of the normal prenatal stimulation provided by their broodmates, they continue to respond solely on the basis of auditory maternal cues later into development (Lickliter & Lewkowicz, 1995). This finding suggests that in the case of human infant preference for the combined audible and visible attributes of ID speech later in development, the preference may emerge as a result of experience with multimodally unified exemplars of such an event.

In an attempt to systematically investigate the separate role that the visible and audible attributes of socially meaningful, multimodal compound stimuli play in perceptual development, Lewkowicz (1996a) conducted a series of studies testing infant response to various attributes of face/voice compounds. One of the attributes studied was ID speech. Infants were first habituated to a videotape of a person speaking in the AD or ID manner and then a series of test trials was given where either the audible, visible, or both attributes of the face/voice compound stimulus were changed. In contrast to the finding that infants as young as 1 month of age exhibit responsiveness to audible ID speech, Lewkowicz (1996a) found that the earliest age at which infants responded to the audible difference between AD and ID speech when it was presented together with a dynamic face was 6 months of age. The finding that infants younger than 6 months of age did not respond to the audible difference between AD and ID speech is markedly different from the previously described finding that infants as young as 1 month of age prefer to listen to ID speech when it is specified solely in the auditory modality (Cooper & Aslin, 1990). This suggests that responsiveness to the kinds of exaggerated pitch and prosody cues characteristic of ID speech can differ considerably depending on whether these cues are unimodal or multimodal.

The finding that 4-month-old infants did not respond to the difference between AD and ID speech in Lewkowicz's (1996a) studies is similar to Werker and McLeod's (1989) finding that 4.5-month-old infants did not exhibit a visual preference for ID speech when it was not bimodally specified. Specifically, in Lewkowicz's (1996a) studies, the 4-month-old infants did not respond to the change from AD to ID speech when it occurred against the background of an AD face. Likewise, in Werker and McLeod's (1989) studies, infants did not prefer ID speech when it was presented against the background of a neutral face. Both findings suggest that discrimination of manner of speech when it is specified in the auditory modality without concurrent specification in the visual modality is possible for 6- and 8-month-old, but not for 4-month-old, infants.

There are some interesting parallels between the results from studies of infants' response to multimodal compounds composed of simple attributes (Lewkowicz, 1988a, 1988b, 1992) and the results from studies of infants' response to unimodal and multimodal exemplars of ID speech. When infants are presented with multimodally specified ID speech, they first begin to exhibit evidence of responsiveness to this type of speech and to its difference from AD speech at 4 months of age if the difference is specified bimodally (Lewkowicz, 1996a; Werker & McLeod, 1989). By 6 months of age, however, in addition to being able to respond to the bimodally specified difference between ID and AD speech, infants begin to exhibit a capacity to respond to the separable components by responding to changes in the audible and visible components separately (Lewkowicz, 1996a). Thus, what appears to emerge later in development is responsiveness to the unimodal components of the multimodal compound. What might be the explanation for this kind of developmental pattern? One may be that multimodally specified information, which in most circumstances is unified, provides maximal information and thus provides infants with multiple discriminative cues. A second may be that responsiveness to the integral nature of the face/voice emerges first in development and, as a result, the infant is less disposed to attend to the separable nature of the components. Of course, the first explanation does not require that infants respond to the compound stimulus as a unified whole because responsiveness may simply be enhanced by greater overall amount of information rather than by the integral nature of the compound stimulus. In contrast, the second explanation does presuppose that infants respond to the integral nature of the compound stimulus and, in fact, studies of infants' responsiveness to auditory–visual equivalence indicate that this ability first emerges around the 4th month of life (for examples see Lewkowicz & Lickliter, 1994a). Thus, it could be that the developmentally later emergence of responsiveness to the unimodal (i.e., audible) attributes of multimodally specified ID speech occurs because the ability to differentiate the separable, unimodal properties of multimodal compounds can only develop after the ability to integrate them develops. In other words, the developmental pattern might be initial detection of differences within single modalities (as evidenced by infant ability to discriminate the difference between ID and AD speech solely on the basis of audible differences), followed by the emergence of intermodal integration, followed by differentiation of the unimodal components that are now part of compound objects and events.

As noted earlier, the differential rate of sensory system development and the differential specialization of the different sensory modalities make predictions about responsiveness to specific types of multimodal events difficult and make the postulation of a general developmental sequence such as the one proposed here a rather complex undertaking. The most effective way to provide a more-reliable basis for making predictions and for specifying more-precise developmental principles regarding infants' response to multimodal objects and events is to gather data on infant response to a variety of events. Thus, one of the questions that Lewkowicz's (1996a) findings raise is whether the 4-month-old infants' failure to respond to the audible attributes of ID speech when they compete with the visual attributes of AD speech reflects a general inability of infants at this stage of development to respond to the exaggerated prosody and pitch characteristics of the audible features of multimodal ID speech. If so, then a similar response pattern should be observed for other types of communicative acts characterized by exaggerated prosody and pitch characteristics.

A ubiquitous type of communicative activity that caretakers engage in with their infants is singing. Singing is highly similar to ID speech in that it is also characterized by exaggerated rhythms, exaggerated intonation contours, and exaggerated stress patterns (Trehub, Trainor, & Unyk, 1993). Moreover, infants are responsive to most features of songs. For example, infants respond to the contours (Trehub, Bull, & Thorpe, 1984), tempos (Trehub & Thorpe, 1989), and timbres (Trehub, Endman, & Thorpe, 1990) of melodies, and to the grouping of the elements of musical patterns on the basis of their similarity (Demany, 1982), and their proximity (Trehub & Thorpe, 1989). Indeed, the many parallels in infant response to ID speech and singing have led Trehub et al. (1993) to conclude that common processing strategies might underlie the processing of speech and music in prelinguistic and premusical listeners. Given the universality of singing and the many parallels between ID speech and singing, we felt that it would be useful to explore infant response to the multimodal attributes specifying the difference between speech and singing. Because singing is characterized by even greater prosody and pitch variations than ID speech, it was expected that responding to the audible attributes of a person seen and heard singing would be discriminable earlier in development than the audible attributes of a person seen and heard speaking in the ID manner.

EXPERIMENT 1

To investigate infants' responsiveness to the various heteromodal attributes specifying the difference between multimodally represented AD speech and singing, infants were habituated to a female speaking in the AD manner and then tested with attributes of a different female engaged in the act of singing. Lewkowicz (1996a) found that the youngest age at which infants responded to the audible difference between AD and ID speech presented in a multimodal context was 6 months of age. Given that singing is generally highly prosodic in nature, it was expected that infants would respond to the difference between multimodally represented speech and singing at least as early in development as they do to differences in the manner of speech and, most likely, even earlier. To make direct comparisons possible, the methods employed in the current experiment were the same as those employed by Lewkowicz (1996a).

Method

Subjects. Separate groups of 3-, 4-, 6-, and 8-month-old infants were tested. There were 32 infants in each group, with an equal number of boys and girls. The mean age of the 3-month group was 14.5 weeks (*SD* = .7 weeks). Two additional 3-month-old infants were tested but not used because 1 fussed and 1 was inattentive. The mean age of the 4-month group was 19.7 weeks (*SD* = .7 weeks). Nine additional 4-month-old infants were tested but not used because 5 fussed and 4 were inattentive. The mean age of the 6-month group was 28.4 weeks (*SD* = .8 weeks). Eight additional 6-month-old infants were tested but not used because 4 were fussy and 4 were inattentive. The mean age of the 8-month group was 37.3 weeks (*SD* = .8 weeks). Five additional 8-month-old infants were tested but not used because 1 was fussy and 4 were inattentive. The subjects in this experiment, as well as in Experiment 2, were full-term at the time of birth, had 1- and 5-min Apgar scores of 7 or higher, and were healthy at the time of testing.

Apparatus and Stimuli. Each infant sat in an infant seat in front of a 13-in. color video monitor (Panasonic CT-1331Y) that was located 50 cm from the infant. A color video camera (Panasonic AG-190) was placed on top of the monitor and was used to videotape the infant's visual behavior throughout the testing session. Infants who were fussy when first placed in the infant seat were then placed on the parent's lap. One 3-month-old and two 6-month-old infants were tested on the lap because they refused to sit in the infant seat. The parent holding the baby was asked to sit as still as possible, refrain from interacting with the baby, and was given no information regarding the experimental hypothesis under examination.

A curtain that extended from the monitor past the point where the infant was seated was placed on each side to occlude the infant's peripheral view. A continuous videotape contained a total of 18 trials, each 20 s in duration. The first trial on the videotape was a pretest trial during which a segment of a "Winnie-the-Pooh" cartoon was presented. The last trial on the videotape was a posttest trial during which the same cartoon was presented. The novelty of each of these two trials at the beginning and end of the test session permitted an assessment of fatigue effects and made it possible to determine whether fatigue might account for any failures to exhibit response recovery during the test trials. The second trial on the videotape commenced the habituation phase. During the habituation phase, the infants saw and heard a female person, whose face filled most of the monitor screen, recite a prepared script in the AD manner of speech.[1] While reciting the script, the person held her head still during the entire 20-s segment and recited the script with a minimum amount of intonation and at a pace that was characteristic of a person reading a passage aloud in the presence of other adults. The average sound pressure level of the person's voice while reading the script ranged between 62 and 67 dB (A scale, re. .0002 dynes/cm^2).

Once the habituation phase ended, a testing phase followed with the presentation of a series of four test trials that included a familiar (F), auditory (A) change, visual (V) change, and an auditory/visual (AV) change test trial. During the F trial, the infants saw and heard the same face and voice seen and heard during the habituation phase. During the A test trial, the infants heard a different female person singing a segment of a song[2] but saw the familiar female face mouthing the AD script. The sound pressure level of the singing voice ranged between 59 and 67 dB. During the V test trial, the infants saw the new person engaged in the act of singing but heard the familiar voice reciting the original AD script. It should be noted that because the lips were seen producing something different than the voice was heard producing, the voice and the lips were out of synchrony with respect to one another during the A and V test trials. Finally, during the AV test trial, the infants saw and heard a new person singing the song (The lips and voice were in synchrony.)

Procedure and Design. Once the infant was seated in front of the monitor, the videotape containing all 18

---

[1] The script was as follows: "A paradox within science is that artificial, or controlled situations help scientists understand the natural world. In natural settings many aspects of any two events differ; the scientist sets up artificial situations to make events similar to each other in many ways."

[2] The song was an excerpt from the Broadway show "Oklahoma" and consisted of the following selection: "Oh what a beautiful morning, oh, what a beautiful day. I've got a beautiful feeling, everything is going our way. There's a bright, golden haze on the meadow, there's a bright golden haze on the meadow."

trials was started and allowed to run to completion. Each trial was separated from the next trial by a 3-s blank and silent screen. For each infant, the test session consisted of the pretest trial, 12 habituation trials, four test trials, and the posttest trial in that order. The four test trials were presented according to one of four orders across the infants in each age group, respectively, such that each of the four test trials was presented equally often in each ordinal position.

The same dependent measure used by Lewkowicz (1996a) was employed in the current experiment. It consisted of recording the total amount of time that the infant spent fixating the video monitor during each trial. This was done by having an observer watch the videotape record of each infant's testing session offline and having that observer score the amount of time that the infant looked at the video monitor during each trial. To make sure that the person scoring the infant's behavior was unaware of the stimulus conditions, the sound track portion of the videotape was turned off and the observer could only see the start and end of each trial (The screen showing the image of the infant's face became brighter when the visual stimulus came on and dimmer when it went off but provided no clues regarding the identity of the visual display.) One may object to the fact that whereas the recording of the amount of visual attention devoted to a visual stimulus is of obvious value, the value of recording visual attention directed toward an auditory stimulus is a less-obvious and useful measure of auditory responsiveness. Prior studies (Horowitz, 1974), however, have shown that visual attention to changing auditory stimulation can provide a reliable and useful measure of auditory processing. This is the case for studies of infants' responsiveness to the auditory component of an auditory–visual compound stimulus (Lewkowicz, 1988a, 1988b, 1992, 1996a), as well as for studies of infants' responsiveness to an auditory stimulus alone (Jusczyk & Kemler Nelson, 1996).

Results and Discussion

As a first step, the data from the habituation trials were analyzed by way of a $4 \times 2 \times 12$ (Age × Gender × Trials) repeated measures MANOVA. This analysis yielded a significant multivariate trials effect, Rao $R(11, 110) = 4.09$, $p < .001$, indicating that the infants as a group decreased their looking over the 12 habituation trials. An orthogonal trend analysis indicated that the linear orthogonal component of this effect was statistically significant, $F(1, 120) = 36.35$, $p < .001$. The MANOVA also yielded a significant univariate Age × Trials interaction, $F(33, 1320) = 10.92$, $p < .05$, although the multivariate interaction did not reach statistical significance. Finally, the analysis yielded a significant age effect, $F(3, 120) = 62.83$, $p < .001$, indicating that there was a difference in the overall magnitude of looking at the different ages. Figure 1 shows the data from the habituation trials separately for each age group. To determine the source of the age effect, a Sheffé post-hoc test was used to test the difference in overall amount of looking between the different ages. The results of this test revealed that the 3-month-old infants looked longer than the 4- ($p < .01$), 6-, and 8-month-old infants (ps $<$ .001), and that the 4-month-old infants looked longer

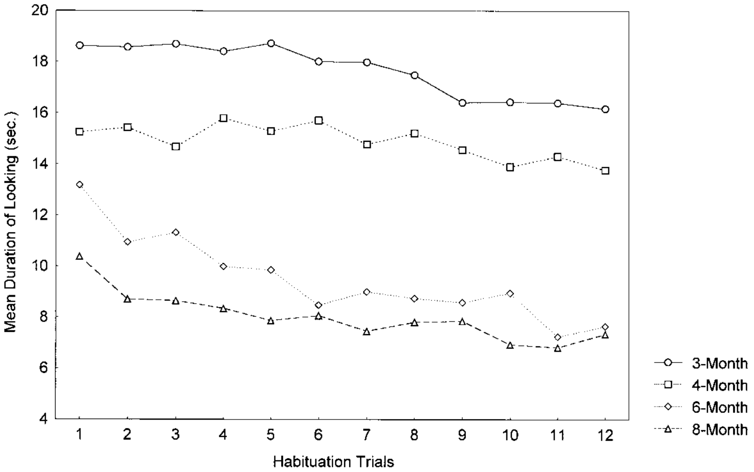

FIGURE 1 Mean duration of looking over the 12 habituation trials in Experiment 1, plotted separately for each age group.

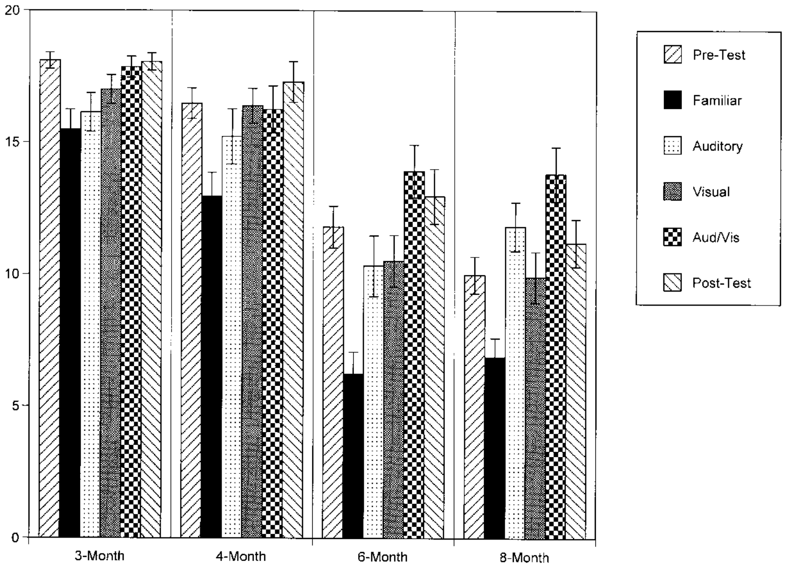

FIGURE 2 Mean duration of looking during the pretest, test, and posttest trials, plotted separately for each age group in Experiment 1. Error bars signify standard error of the mean *(SEM)*.

than the 6- ($p < .001$), and the 8-month-old ($p < .001$) infants.

The results from the pretest, test, and posttest trials are depicted in Figure 2. The first question addressed by the analyses was whether the infants became fatigued during the test session. To answer this question, two different analyses were performed. The first compared the results from the pretest and posttest trials by a $2 \times 4 \times 2$ (Trials × Age × Gender) MANOVA which showed that there was a marginally significant increase in looking between the beginning (pretest trial: 14.1 s) and end of the test session (posttest trial: 14.9 s), $F(1, 120) = 2.98$, $p < .10$. The second analysis compared the amount of looking in the posttest trial versus the amount of looking in the F trial. If the infants had not become fatigued, and if they recognized the novelty of the stimulus presented during the posttest trial, they should have exhibited a significant response recovery. Consistent with this expectation, the average amount of looking during the posttest trial (14.9 s) was significantly greater, $F(1, 120) = 108.63$, $p < .001$, than the average amount of looking in the F trial (10.4 s). This finding indicates that any failure to exhibit significant response recovery during any of the test trials could not have been due to fatigue effects, as all of the test trials preceded the posttest trial. Taken together, the results of the preceding two analyses indicate that the infants not only did not become fatigued during the test session but that they actually became more interested as the session progressed.

The question of greatest interest was whether the infants responded to the various changes in the face/voice compound stimulus. Because Lewkowicz (1996a) previously showed that responsiveness to similar types of changes differs across age, a preliminary analysis of the test trial data was conducted first to determine whether the age of the infants affected responsiveness in the current experiment as well. To do so, the data from the F, A, V, and AV trials were submitted to a $4 \times 2 \times 4$ (Age × Gender × Trials) MANOVA. Results of this analysis yielded a highly significant univariate, $F(3, 360) = 45.93$, $p < .001$, and multivariate, Rao $R(3, 118) = 50.12$, $p < .001$, effect of trials, a significant age effect, $F(3, 120) = 22.88$, $p < .001$, and a significant multivariate Age × Trials interaction, Rao $R(9, 287) = 3.61$, $p < .001$.

Given that the Age × Trials interaction was significant, the test trial data from each age group were analyzed separately by conducting separate sets of planned contrast analyses at each age. These planned contrasts compared the duration of looking in the A, V, and AV test trial, respectively, with the amount of looking in the F trial. Significantly greater looking in any of these trials relative to the amount of looking in the F trial indicated the infants discriminated the change in that test trial. Table 1 shows the results of these contrasts. As can be seen, the 4-, 6-, and 8-month-old infants discriminated all three types of changes (see Figure 2). In contrast, although the 3-month-old infants responded to the visible and

Table 1. Results of Planned Comparison Tests on the Data from the Test Trials in Experiment 1

Test Trial	3-Month	4-Month	6-Month	8-Month
Auditory	$F(1, 120)=.50$, $p=.47$	$F(1,120)=6.36$, $p<.025$	$F(1, 120)=20.79$, $p<.001$	$F(1, 120)=30.42$, $p<.001$
Visual	$F(1, 120)=4.12$, $p<.05$	$F(1, 120)=21.21$, $p<.001$	$F(1, 120)=33.38$, $p<.001$	$F(1, 120)=16.72$, $p<.001$
Aud–Vis	$F(1, 120)=7.13$, $p<.01$	$F(1, 120)=13.88$, $p<.001$	$F(1, 120)=75.98$, $p<.001$	$F(1, 120)=61.83$, $p<.001$

audible/visible change, they did not respond to the audible change.

The results from this experiment indicate that 4-, 6-, and 8-month-old infants discriminated the audible, visible, and combined attributes of a multimodally represented person. In the present case, the contrasts that the infants were capable of discriminating involved the difference between a person engaged in the act of speaking in an adult-directed manner and another person engaged in the act of singing. Even though 3-month-old infants also responded to the visible and bimodal changes, in contrast to the older infants, the youngest groups of infants did not respond to the audible change from AD speech to singing. Taken together, these findings suggest that sensitivity to the kinds of audible differences between AD speech and singing presented here emerges between 3 and 4 months of age.

A comparison of the current findings with those from Lewkowicz's (1996a) studies reveals one interesting difference. Whereas the 4-month-old infants in the current experiment responded to the change in the person's audible features, the infants in Lewkowicz's (1996a) studies did not. The principal difference between the two studies was that the audible change in the current experiment involved a change from a female speaking in the AD manner to a female singing, whereas the change in Lewkowicz's study was from AD to ID speech. The fact that responsiveness to the difference between AD speech and singing is present 2 months earlier supports the prediction that it should be easier for infants to detect the latter type of difference presumably because singing is characterized by greater overall prosody relative to ID speech.

EXPERIMENT 2

The findings from Experiment 1 showed that 4-, 6-, and 8-month-old infants responded to all three types of changes in the face/voice compound stimulus but that 3-month-olds did not respond to the changes in its audible component. When these findings are considered together with Lewkowicz's (1996a) earlier findings, the following pattern emerges. Infants do not respond to an audible difference in the face/voice compound when the difference is specified by AD speech even as late as 8 months of age. In contrast, when the audible difference is between AD and ID speech, 6- and 8-month-old infants respond to it, and when the difference is between AD speech and singing, infants as young as 4 months of age respond to it. This pattern suggests that the closer that the audible differences approximate the infant's real-world experience, the earlier in development do infants show responsiveness to the audible features of the face/voice compound. If that is the case, what other features of the world might be salient to young infants?

One of the most ubiquitous perceptual distinctions in the infant's social world is gender. Infants are constantly exposed to gender differences as they interact with the people around them and empirical evidence indicates that they are responsive to the audible, visible, and combined differences in gender. For example, Miller (1983) found that 6-month-old infants respond to the categorical differences in the syllable "hi" spoken by males and females and Jusczyk, Pisoni, and Mullenix (1992) found that even 2-month-olds respond to the categorical differences in syllables spoken by people of different gender. Cornell (1974) and Fagan and Singer (1979) found that 5- and 6-month-olds can recognize static faces on the basis of gender, and Leinbach and Fagot (1993) showed that 9- and 12-month-old infants can perform categorical gender discrimination of pictures of different faces. Finally, Walker-Andrews, Bahrick, Raglioni, and Diaz (1991) found that 6-month-old infants could match the gender of a voice with the gender of a face in a study of cross-modal matching of dynamically presented faces and voices.

Given that infants respond to auditory gender differences as early as 2 months of age, would infants younger than 4 months of age also respond to audible differences inherent in a face/voice compound if gender was one of the differences? That was the question addressed in this experiment, and to do so, infants first were habituated to a male speaking in the AD manner and then were tested with a female who was engaged in the act of singing.

Method

Subjects. Separate groups of 3-, 4-, 6-, and 8-month-old infants were tested. There were 32 infants in each group, with an equal number of boys and girls. The mean age of the 3-month group was 14.8 weeks ($SD = .7$). Three additional 3-month-old infants were tested but not used because 2 were fussy and 1 was inattentive. The mean age of the 4-month group was 19.7 weeks ($SD = .8$ weeks). Four additional 4-month-old infants were tested but not used because 3 fussed and 1 was inattentive. The mean age of the 6-month group was 28.3 weeks ($SD = .8$ weeks). Seven additional 6-month-old infants were tested but not used because 5 were fussy and 2 were inattentive. The mean age of the 8-month group was 37 weeks ($SD = .7$ weeks). Seven additional 8-month-old infants were tested but not used because 5 were fussy and 2 were inattentive.

Apparatus and Stimuli. The apparatus and stimuli were the same as in Experiment 1, with one exception. In this experiment, a male person was seen and heard during the habituation phase reciting the same AD script that the female in Experiment 1 recited. The average sound pressure level of the male's voice while reciting the script ranged between 62 and 67 dB (A scale, re. .0002 dynes/cm^2). Following the habituation phase, the same novel female presented in the test phase of Experiment 1 was seen and heard singing.

Procedure. The procedure was the same as the procedure used in Experiment 1. Three 3-month-old, two 6-month-old, and two 8-month-old infants were tested on the parent's lap.

Results and Discussion

Figure 3 shows the habituation data separately for each age group. Analysis of these data by way of an Age × Gender × Trials (4 × 2 × 12) repeated measures MANOVA indicated that, as a group, the infants decreased their looking over the 12 habituation trials. This was manifest in a significant multivariate trials effect, Rao $R(11, 110) = 4.03$, $p < .001$. An orthogonal trend analysis indicated that the linear orthogonal component of this effect was statistically significant, $F(1, 120) = 34.78$, $p < .001$. In addition, the age effect was significant, $F(3, 120) = 29.24$, $p < .001$. The Sheffé post-hoc test indicated that the age effect was due to the fact that the 3- and 4-month-old infants, respectively, looked longer than the 6-month-old, $p < .001$, and the 8-month-old, $p < .001$, infants, and that the 6-month-old infants looked longer than the 8-month-old infants, $p < .05$.

Figure 4 shows the data from the pretest, test, and posttest trials, separately for each age group. Comparison of the pretest and posttest results by way of a Trials × Age × Gender (2 × 4 × 2) MANOVA indicated that, like in Experiment 1, there was a marginally significant increase in looking from the beginning (15 s) to the end (15.7 s) of the test session, $F(1, 120) = 3.57$, $p < .07$. This shows that rather than becoming fatigued during the test session, the infants actually became more interested as the session progressed. Additional evidence that the infants did not become fatigued comes from the finding that the infants exhibited significant response recovery when viewing the cartoon segment in the posttest trial, $F(1, 120) = 72.67$, $p < .001$.

The preliminary analysis of the test-trial data by

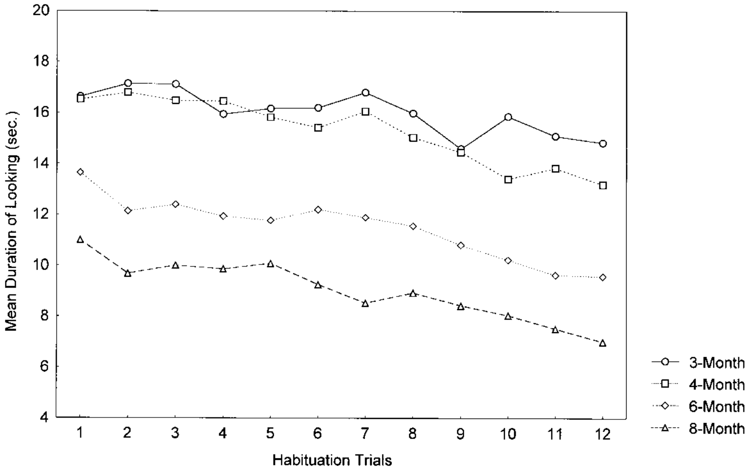

FIGURE 3 Mean duration of looking over the 12 habituation trials in Experiment 2, plotted separately for each age group.

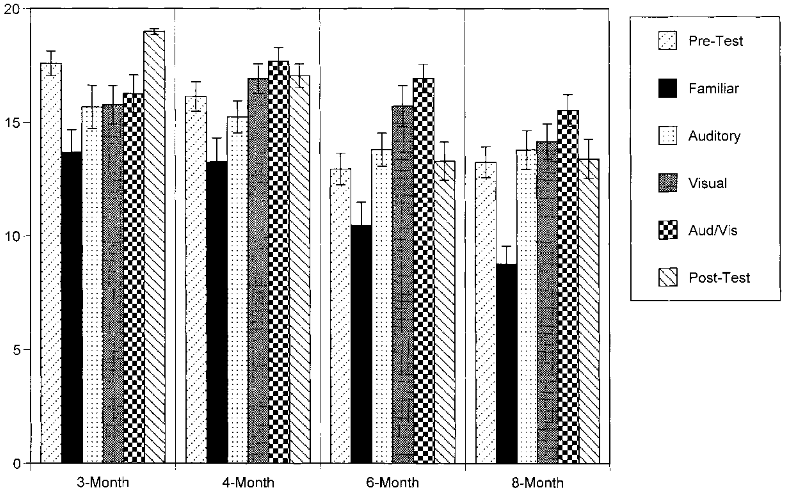

FIGURE 4 Mean duration of looking during the pretest, test, and posttest trials, plotted separately for each age group in Experiment 2. Error bars signify standard error of the mean *(SEM).*

way of an Age × Gender × Trials (4 × 2 × 4) MANOVA indicated that there was a significant univariate, $F(3, 360) = 53.27$, $p < .001$, and multivariate effect of trials, Rao $R(3, 118) = 39.7$, $p < .001$, a significant age effect, $F(3, 120) = 3.61$, $p < .025$, and a significant multivariate Age × Trials interaction, Rao $R(9, 287) = 1.89$, $p = .05$. Given the significant Age × Trials effect, the test-trial data from each age were analyzed separately by way of planned contrast analyses as was done in Experiment 1. The results of these comparisons can be seen in Table 2 and show that infants of all ages discriminated all three types of changes. Thus, in contrast to the results from Experiment 1, the current results showed that in addition to discriminating the changes in the V and the AV test trials, the 3-month-old infants also discriminated the audible change.

A comparison of the results from the current experiment and those from Experiment 1 indicates that whereas the 3-month-old infants in Experiment 1 did not respond to the audible difference between AD speech and singing, the infants in the current experiment did. The one critical difference between the two experiments was that the person speaking in AD and the person singing were of different gender in the current experiment. Thus, it appears that the addition of gender cue differences to prosody cue differences made the audible difference salient enough to the 3-month-old infants.

It is interesting to note that the gender difference in the current experiment made discrimination of changes in the audible characteristics of the face/voice compound possible. This stands in contrast to Lewkowicz's (1996a) finding indicating that a gender difference did not aid in the discrimination of the audible attributes of a person. The difference in outcome is most likely due to the fact that in the current experiment gender differed along with a difference between AD speech and singing, whereas the gender difference accompanied a difference between two segments of AD speech in Lewkowicz's (1996a) studies. This suggests that the combination of a gender and prosody difference enables young infants (i.e., 3-month-olds) to detect changes in the audible attributes of a face/voice compound. At the same time, however, it should be noted that older infants do not require concurrent gender differences to discriminate the difference between AD speech and singing. This suggests that as

Table 2. Results of Planned Comparison Tests on the Data from the Test Trials in Experiment 2

Test Trial	3-Month	4-Month	6-Month	8-Month
Auditory	$F(1, 120)=4.20$, $p<.05$	$F(1, 120)=4.16$, $p<.05$	$F(1, 120)=11.96$, $p<.001$	$F(1, 120)=26.94$, $p<.001$
Visual	$F(1, 120)=5.57$, $p<.05$	$F(1, 120)=16.47$, $p<.001$	$F(1, 120)=35.17$, $p<.001$	$F(1, 120)=37.00$, $p<.001$
Aud–Vis	$F(1, 120)=7.46$, $p<.01$	$F(1, 120)=21.57$, $p<.001$	$F(1, 120)=47.47$, $p<.001$	$F(1, 120)=50.71$, $p<.001$

development progresses, infants become increasingly more responsive to various audible attributes of face/voice compounds and that, as a result, there is a shift from an initial reliance on multiple discriminative cues to a later ability to respond on the basis of single cues.

GENERAL DISCUSSION

The present experiments extend previous studies of infants' response to the audible, visible, and bimodal attributes of human faces. The previous studies (Lewkowicz, 1996a) showed that 4-, 6-, and 8-month-old infants did not respond to audible differences between different samples of continuous AD speech, even though they did respond to the visible and bimodal differences between such samples of AD speech. In addition, the previous studies showed that when infants were asked to discriminate between a sample of continuous AD speech and a sample of continuous ID speech, 6- and 8-month-old infants responded not only to the visible and bimodal differences but to the audible difference as well, whereas 4-month-old infants did not respond to the audible difference. These findings raised the question of whether the 4-month-old infants were unresponsive to audibly specified prosody differences in general, or whether they were unresponsive to the difference between AD and ID speech in particular. Consequently, the objective of the current experiments was to investigate this question by examining responsiveness to other types of prosodic variations normally present in an infant's ecological setting.

Utilizing methods similar to those used in the prior studies (Lewkowicz, 1996a), infants were familiarized with a person speaking in the AD manner and then were given a series of test trials during which another person either could be heard, seen, or heard and seen singing. The results from Experiment 1 indicated that the 4-month-old infants' failure to respond to the audible differences in Lewkowicz's (1996a) studies did not reflect a general insensitivity to audible differences in prosody in the multimodal context. Specifically, the results from Experiment 1 indicated that infants as young as 4 months of age responded not only to the visible and bimodal differences between AD speech and singing, but that they also responded to the audible differences. This finding suggests that the kinds of exaggerated prosody and spectral cues characteristic of singing are more salient to 4-month-old infants than are those characteristic of ID speech. The findings from the 6- and 8-month-old infants confirmed and extended the prior findings in showing that 6- and 8-month-old infants can discriminate the audible, visible, and bimodal attributes of faces not only when these attributes are specified by the difference between a person speaking in the AD and ID manner but also when they are specified by the difference between a person speaking in the AD manner and a person singing.

The new and interesting finding in the current studies was that 3-month-old infants did not respond to the difference between what might be described as a low-prosody-variation event (AD speech) and a high-prosody-variation event (singing) in Experiment 1. This raised the same question that was raised with regard to the 4-month-old infants' response to AD–ID speech differences, namely, is this difference not salient enough and/or are the infant's perceptual capacities not differentiated enough to detect this difference? To test this question, the infants in Experiment 2 were provided with the additional discriminative cue of gender that is known to be discriminable to infants. As expected, when provided with a gender difference together with a difference in the degree of prosody variation, the 3-month-old infants, like the older groups of infants, discriminated the audible difference. It is unlikely that the successful discrimination of the audible difference by the 3-month-old infants was simply a response to the gender difference because Lewkowicz (1996a) showed that gender differences themselves are not sufficient for infants to respond to the audible differences in face/voice compounds differing in their AD characteristics. Thus, the most conservative interpretation of the finding is that the 3-month-old infants relied on the combination of gender and prosody differences to make the successful discrimination.

It could be argued that the successful discriminative response in the A and V test trials does not mean that infants successfully detected the audible and visible differences between AD speech and singing. Rather, because a change in the lexical/semantic content of the utterance in the A test trial and a disruption of audio–visual synchrony in both the A and V test trials occurred during these two trials, it might be argued that these two differences were responsible for the successful discrimination. This possibility can be ruled out for three reasons. First, infants of all ages responded to the bimodal change despite the fact that audio–visual synchrony was not disrupted in that trial. In fact, and in direct contrast to the argument that temporal asynchrony might have contributed to the successful discrimination, is the finding that the response to the bimodal change was consistently the most robust even though asynchrony was not involved in that trial at all. Second, the 3-month-old infants in Experiment 1 and the 6- and 8-month-old infants in Lewkowicz's (1996a) studies responded to the change in the V test trial but did not respond to the change in the A test trial. For of the 3-month-old infants, this difference was between AD speech and singing and in the case of the two older groups it was the difference between different segments of AD speech. If a disruption of synchrony were a salient aspect of the test trials, then the infants should have responded to the change in the A test trial. Evidence indicates that infants are capable of responding to auditory–visual asynchrony (Lewkowicz, 1992, 1996b). Finally, it should be noted that in studies utilizing the same testing procedures, Lewkowicz (1996a) found that 4-month-old infants did not respond to the change in either the A or the V test trials even though audio–visual synchrony was disrupted in each type of test trial, and even though the lexical/semantic content of the utterance changed in the A test trial. In sum, the weight of the evidence strongly suggests that neither synchrony disruption nor lexical/syntactic content differences played a significant role in the differential responsiveness observed in the current experiments.

When the current findings are compared with previous findings from studies of infants' responsiveness to the multimodal properties of the human face, they suggest that the exaggerated prosody cues and exaggerated spectral cues associated with various forms of multimodally represented, infant-directed communicative acts may depend to some extent on the specific properties of the cues. Thus, the findings from the current experiments show that infants as young as 4 months of age can discriminate between the audible features of multimodally specified AD speech and singing in same-gender individuals, and that infants as young as 3 months of age can make such discriminations in different-gender individuals. In contrast, Lewkowicz's (1996a) findings showed that 4-month-old infants do not respond to the audible difference between multimodally represented AD and ID speech. The two findings of 4-month-old infants' failure to respond to the audible component of ID speech in conjunction with the current finding of successful discrimination of the difference between AD speech and singing suggests that singing carries additional prosody and spectral cues which are presumably more salient. In addition, the current findings suggest that the prosody and spectral cues associated with singing are sufficiently salient to infants as young as 4 months of age in a same-gender context and to infants as young as 3 months of age in a different-gender context that they are able to ignore unchanging and discrepant visual information and detect the audible change.

Although the present findings do not provide any direct information on what specific features of the difference between AD speech and singing might have mediated responsiveness, investigations of infants' responsiveness to music provide some suggestions regarding the cues that might have served as the basis for the discriminations. One of the most consistent findings from such investigations is that infants' responsiveness to musical patterns is dominated by pitch-contour and pitch-range processing (Trehub et al., 1993). This suggests that the earlier emergence of responsiveness to the difference between AD speech and singing relative to the emergence of responsiveness to the difference between AD and ID speech is likely due to greater variations in pitch contours in the singing presented in the current experiments compared to the ID speech presented in Lewkowicz's (1996a) studies. It is also likely that exaggerated prosody cues contributed to the successful response, although the relative contribution of exaggerated prosody cues cannot be determined.

The current findings also provide some interesting insights into the power of the visible attributes of faces in determining responsiveness to the concurrent audible attributes of the face. The findings from Experiment 1 indicated that the 3-month-old infants did not respond to the audible difference between AD speech and singing even though they responded to the visible and bimodal differences. Lewkowicz (1996a) also found instances where infants did not respond to the audible differences between different segments of AD speech spoken by different people or where 4-month-old infants did not respond to the difference between AD and ID speech even though in all of these cases the infants responded to the visible and bimodal differences. Could it be that the visible properties of the face can sometimes override its audible properties and suppress responding to the audible attributes? Findings on infants' responsiveness to sentential prosody suggest that this might be the case. Mandel, Jusczyk, and Kemler Nelson (1994) showed that infants as young as 2 months of age respond to sentential prosody; after being habituated to a sentence, the infants exhibited response recovery when two of the words in the sentence were changed. If infants as young as 2 months of age are responsive to sentential prosody presented in the auditory modality alone, then the findings of infants' failure to respond to differences between different segments of AD speech, to differences between AD and ID speech, or to differences between AD speech and singing in the presence of the face, all of which not only differed in sentential prosody but in a number of other ways, suggest that the face probably played some role in the infants' responsiveness to the audible information. This effect could have been either the result of the suppression of the responsiveness to the audible information by the visible alone, or could have been the result of the combination of visible suppression and perceptual insensitivity to the audible information. Future studies need to disentangle these possibilities.

In conclusion, Lewkowicz's (1996a) and the current findings suggest that the timing underlying the emergence of responsiveness to various audible and visible attributes of multimodally specified faces is likely to be the result of a complex interaction between the sensory/perceptual status of the infant at a given point in development and his or her experience up to that point in development (Gottlieb, 1991; Lewkowicz & Lickliter, 1994b; Thelen & Smith, 1994). The response propensities resulting from this kind of dynamic interaction between experience and growth act on the information provided by the infant's social partner to determine responsiveness at each point in development. At any one time in development, the perceptual system possesses a range of response capacities and the most effective way to uncover this range of response capacities is by studying the infant's response to a range of stimulation (Lewkowicz & Lickliter, 1994b). Thus, infants can appear to be either more or less sophisticated depending on the type of information that is presented. For example, infants appear to be relatively unresponsive to audible differences between different segments of AD speech even as late as 8 months of age. In contrast, when audible differences between AD and ID speech are presented, infants exhibit responsiveness by 6 months of age. When differences between AD speech and singing are involved, infants exhibit even more impressive abilities in that they respond to such differences as early as the 4th or 3rd month of life depending on concomitant gender attributes. There is little doubt that these kinds of developmental differences reflect the interaction between differential experience and inherent changes in sensory/perceptual capacities. In addition, infants can appear to be more or less sophisticated depending on the response system tested. For example, when affective responsiveness is measured, infants exhibit sensitivity to AD and ID speech contrasts even though they do not evidence such sensitivity when visual preferences are measured (Fernald, 1993) and they exhibit such sensitivity earlier in development than they do when visual preferences are measured (Pegg et al., 1992). This suggests that differential maturation of different response systems must be considered along with differential experience, inherent changes in sensory/perceptual capacities, and properties of the stimulus events when attempting to understand the development of the kinds of perceptual capacities exemplified by responsiveness to multimodally represented social events (Lewkowicz & Lickliter, 1994b).

NOTES

This work was supported in part by funds from the NYS Office of Mental Retardation and Developmental Disabilities. I thank Marci Dabbene and Daisy Edmondson for their assistance and the staff at St. Vincent's Medical Center of Richmond County, Staten Island, NY for their assistance in subject recruitment. I also thank Robert Freedland for comments on an earlier version of this article.

REFERENCES

Bushnell, E. W. (1994). A dual-processing approach to cross-modal matching: Implications for development. In D. J. Lewkowicz & R. Lickliter (Eds.), *Development of intersensory perception: Comparative perspectives* (pp. 19–38). Norwood, NJ: Erlbaum.

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61*, 1584–1595.

Cornell, E. H. (1974). Infants' discrimination of photographs of faces following redundant presentations. *Journal of Experimental Child Psychology, 18*, 98–106.

Demany, L. (1982). Auditory stream segregation in infancy. *Infant Behavior and Development, 5*, 261–276.

Fagan, J. F., & Singer, L. T. (1979). The role of simple feature differences in infant recognition of faces. *Infant Behavior and Development, 2*, 39–46.

Fernald, A. (1985). Five-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8*, 181–195.

Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development, 64*, 657–674.

Fernald, A., & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese. *Infant Behavior and Development, 10*, 279–293.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20*, 104–113.

Gottlieb, G. (1971). Ontogenesis of sensory function in birds and mammals. In E. Tobach, L. R. Aronson, & E. Shaw (Eds.), *The biopsychology of development* (pp. 67–128). New York: Academic Press.

Gottlieb, G. (1991). Experiential canalization of behavioral development: Theory. *Developmental Psychology, 27*, 35–39.

Horowitz, F. D. (1974). Visual attention, auditory stimulation, and language discrimination in young infants. *Monographs of the Society for Research in Child Development, 39* (5–6) (Serial No. 158).

Jack, C. E., & Thurlow, W. R. (1973). Effects of degree of visual association and angle of displacement on the "ventriloquism" effect. *Perceptual and Motor Skills, 37*, 967–979.

Jusczyk, P. W., & Kemler Nelson, D. G. (1996). Syntactic units, prosody, and psychological reality during infancy. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax:*

Bootstrapping from speech to grammar in early acquisition. (pp. 389–408). Mahwah, NJ: Erlbaum.

Jusczyk, P. M., Pisoni, D. B., & Mullenix, J. (1992). Some consequences of stimulus variability on speech processing by two-month-old infants. *Cognition, 43*, 253–291.

Kubovy, M. (1988). Should we resist the seductiveness of the space:time::vision:audition analogy? *Journal of Experimental Psychology: Human Perception and Performance, 14*, 318–320.

Leinbach, M. D., & Fagot, B. I. (1993). Categorical habituation to male and female faces: Gender schematic processing in infancy. *Infant Behavior and Development, 16*, 317–332.

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

Lewkowicz, D. J. (1988b). Sensory dominance in infants: II. Ten-month-old infants' response to auditory–visual compounds. *Developmental Psychology, 24*, 172–182.

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

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

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

Lewkowicz, D. J., & Lickliter, R. (1994a). *The development of intersensory perception: Comparative perspectives*. Hillsdale, NJ: Erlbaum.

Lewkowicz, D. J., & Lickliter, R. (1994b). Insights into mechanisms of intersensory development: The value of a comparative, convergent-operations approach. In D. J. Lewkowicz & R. Lickliter (Eds.), *The development of intersensory perception: Comparative perspectives* (pp. 403–413). Hillsdale, NJ: Erlbaum.

Lickliter, R., & Lewkowicz, D. J. (1995). Intersensory experience and early perceptual development: Attenuated prenatal sensory stimulation affects postnatal auditory and visual responsiveness in bobwhite quail chicks (*Colinus virginianus*). *Developmental Psychology, 31*, 609–618.

Lickliter, R., & Virkar, P. (1989). Intersensory functioning in bobwhite quail chicks: Early sensory dominance. *Developmental Psychobiology, 22*, 651–667.

Mandel, D. R., Jusczyk, P. W., & Kemler Nelson, D. G. (1994). Does sentential prosody help infants organize and remember speech information? *Cognition, 53*, 155–180.

Massaro, D. W., & Cohen, M. M. (1990). Perception of synthesized audible and visible speech. *Psychological Science, 1*, 55–63.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 229–239.

Miller, C. L. (1983). Developmental changes in male/female voice classification by infants. *Infant Behavior and Development, 6*, 313–330.

Papousek, M., & Papousek, H. (1981). Musical elements in the infant's vocalization: Their significance for communication, cognition, and creativity. *Advances in Infancy Research, 1*, 163–224.

Papousek, M., Papousek, H., & Bornstein, M. H. (1985). The naturalistic environment of young infants: On the significance of homogeneity and variability in parental speech. In T. M. Field & N. A. Fox (Eds.), *Social perception in infants* (pp. 269–297). Norwood, NJ: Ablex.

Papousek, M., Papousek, H., & Haekel, M. (1987). Didactic adjustments in fathers' and mothers' speech to their three-month-old infants. *Journal of Psycholinguistic Research, 16*, 306–319.

Pegg, J., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior and Development, 15*, 325–345.

Snow, C. E. (1977). The development of conversation between mothers and babies. *Journal of Child Language, 4*, 1–22.

Stein, B. E., Meredith, M. A., & Wallace, M. T. (1994). Development and neural basis of multisensory integration. In D. J. Lewkowicz & R. Lickliter (Eds), *The development of intersensory perception: Comparative perspectives* (pp. 81–105). Hillsdale, NJ: Erlbaum.

Stern, D. N., Spieker, S., & MacKain, K. (1982). Intonation contours as signals in maternal speech to prelinguistic infants. *Developmental Psychology, 18*, 727–735.

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

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

Trehub, S. E., Bull, D., & Thorpe, L. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development, 55*, 821–830.

Trehub, S. E., Endman, M. W., & Thorpe, L. A. (1990). Infants' perception of timbre: Classification of complex tones by spectral structure. *Journal of Experimental Child Psychology, 49*, 300–313.

Trehub, S. E., & Thorpe, L. A. (1989). Infants' perception of rhythm: Categorization of auditory sequences by temporal structure. *Canadian Journal of Psychology, 43*, 217–229.

Trehub, S. E., Trainor, L. J., & Unyk, A. M. (1993). Music and speech processing in the first year of life. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 24, pp. 1–35). San Diego: Academic Press.

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

Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology, 43*, 230–246.

Werker, J. F., Pegg, J. E., & McLeod, P. J. (1994). A cross-language investigation of infant preference for infant-directed communication. *Infant Behavior and Development, 17*, 323–333.