Research on infant perceptual development has often focused on how infants process information in their environment, and how that information guides interaction with their surround. Given that the world is replete with complex relationships, the information that guides perceptual processes such as attention is also apt to be of a complex nature. That is, a multiplicity of factors likely determines infant attention to any event. It has been suggested that both stimulus factors (e.g., color, size, amplitude, rate of stimulation) as well as factors specific to the infant (e.g., state of arousal, state of sensory receptor adaptation, prior response history with the stimulus event) interact to determine attention to events (McGuire & Turkewitz, 1979; Schneirla, 1959, 1965).

Investigations of what factors affect infant attention to infant-directed (ID) speech has recently turned from focusing on pitch characteristics, to focusing on rate of speaking. Generally, the results of these studies have shown that both 1- and 4-month-old infants prefer slower to faster rates of ID speech when pitch information is held constant (J. Cooper, R. Cooper, Ostroff, & Aslin, 2001). Although, these results are the first to illustrate how speaking rate affects infant speech preferences, they are best considered a first step toward more fully understanding how multimodal sources of information (e.g., auditory, visual, tactile) influence infant attention. The present study was designed to further investigate the role of multimodal factors in attention by exploring how variations in visual stimulation can affect infant speech preferences, especially for rate of speaking.

Factors Affecting Infant Attention to Visual Events

Most investigations of speech perception have focused on singular features of speech. Historically, research on infant attention to visual events has involved both featural analyses and investigations of how multimodal factors determine infant attention. In regard to the former, a substantial amount of research has investigated attention to differing rates of visual stimulation (e.g., frequency of flashing checkerboard displays). Overall, studies find that infants prefer visual displays flashing at higher relative frequencies (see Lewkowicz, 1989, for a review). For instance, Lewkowicz (1985) found that 6-month-old infants looked more to displays flashing at 8 Hz. over displays flashed at 2 and 4 Hz. Balaban and Dannemiller (1992) found similar results with 12-week-olds.
Research has also found that multiple factors can concurrently affect attention to visual events. For example, McCarvill and Karmel (1976) found that 9- and 13- week old infants looked more at visual displays of moderate contour density and moderate luminance levels, with decreased attention to displays containing higher levels of either factor. Hence, it seems that multiple factors within a singular modality can affect infant attention to events.

An interesting facet of the McCarvill and Karmel (1976) study is that infants looked more to displays of lower contour density in the context of higher levels of luminance, and looked more to higher contour displays in the context of lower luminance levels. A possible explanation for this behavior is that the infants sought to maintain a constant level of stimulation. That is, as the stimulus energy in the luminance domain decreased, infants sought out circumstances where the amount of available energy in the contour domain was higher. Hence, the infant may have sought to maintain a stable level of stimulus energy intake across variations in visual display characteristics (Maurer, 1993).

The idea that infants seek to maintain a relatively constant level of stimulation across avenues of input is commonly referred to as arousal modulated attention. Here, the term arousal refers to any type of stimulation the infant may experience. This stimulation may be exogenous (e.g., stimulus characteristics) or endogenous (e.g., physiological arousal). In the former, this would refer to events like sensory receptor excitation; in the later, this would Indeed, much research has focused on understanding how arousal modulated attention affects infant attention visual events.

In addition to looking at how multiple factors within the visual modality concurrently affect visual attention, it has also been found that factors across sensory modalities may affect infant attention to visual events. For example, Lawson and Turkewitz (1980) presented newborn infants with visual displays comprised of low, intermediate, or high numbers of black-and-white cubes. Half of the infants received a burst of white noise (80 dB.) concomitant with the presentation of the visual displays, while the other infants received no sound when presented the displays. Compared to the no-sound group, infants who received sound trials looked significantly more at visual displays containing lower numbers of cubes and less at displays containing higher numbers of cubes. Mendelson and Haith (1976) observed a similar effect when they presented newborn infants with a visual display consisting of a single bar and three stripes (located on
opposite sides of the display) either in the presence or absence of a voice or modulated tone. In the absence of sound, infants’ visual fixations were closer to the higher intensity (i.e., multiple stripe) region of the display, and farther from the lower-intensity single bar. The opposite pattern was observed in the presence of sound.

Additionally, it has been suggested that factors specific to the organism may determine infant attention to stimulus events. These factors include state of arousal, state of sensory receptor adaptation, and prior response history with the stimulus event (McGuire & Turkewitz, 1979; Schneirla, 1959, 1965). Research exists to support this idea. For example, Lewkowicz and Turkewitz (1981) investigated visual preferences of 1-month-old infants who either had or had not been exposed to sound before the start of a visual preference task. Here, all infants were presented with visual stimuli at low, intermediate, and high luminance levels. Infants not exposed to sound showed maximum attention to displays of intermediate intensity, with decreased attention to the low and high intensity displays (i.e., inverted U-shaped relationship). However, the looking behavior of infants exposed to sound took the form of a linear relationship, with maximum attention to low intensity displays. Lewkowicz and Turkewitz interpreted that the arousal level of the sound-exposed group was higher than that of the group not exposed to sound. Therefore, in the context of elevated internal arousal, infants seemed to have sought out external events containing less stimulus energy.

Gardner and Karmel (1995) undertook a similar investigation. Newborn and 1-month old infants were presented with unpatterned light panels flashing at frequencies between 1 and 8 Hz. All infants were tested on different occasions both before and after having been fed. Infants were classified as having higher levels of arousal when tested prior to feeding, and having lower levels of arousal when tested subsequent to feeding. In the high arousal condition, infants looked more to displays flashing at slower frequencies, while attending more to displays of faster frequency in the low arousal condition.

It is apparent from the research reviewed that several factors may influence infant attention to visual events and that young infants seem to maintain a relatively constant level of stimulation across sensory modalities (Maurer, 1993). As the amount of energy from a source changes, infants may adjust their looking behaviors in an effort to sustain a consistent amount of stimulation. Some studies have systematically manipulating the amount of stimulus energy or
organismic state to affect levels of internal arousal (i.e., arousal modulated attention) and influence infant attention to visual events (e.g., Gardner & Karmel, 1995; Lewkowicz & Turkewitz, 1981). It bears mentioning, however, that no direct measures of infant physiological arousal were taken in these investigations. Instead, inferences were made from an existing body of work showing that feeding (Gardner & Turkewitz, 1982; Harper, Hoppenbrouwers, Bannett, Hodgman, Sterman, & McGinty, 1977) and white-noise stimulation (Bartoshuk, 1964; Steinschneider, Lipton, & Richmond, 1966; Turkewitz, Moreau, Birch, & Davis, 1970) produce autonomic and behavioral changes in infant state.

In addition to studies on how stimulus energy affects infant preferences for visual events, researchers have also examined how the structure and meaningfulness of visual stimuli affect infant attention. A large portion of this research has looked specifically at differences in infant attention to face-like and non-face-like stimulus displays. Overall, it seems that infants prefer face-like over non-face-like visual stimuli from birth (Morton & Johnson, 1991). For instance, Fantz (1961) found that neonates preferentially fixated to schematic face-like patterns over abstract patterns (e.g., a bulls-eye) across variations in brightness, number of pattern elements, and pattern symmetry. Why neonates prefer face-like displays has been of considerable interest given that infants younger than 2-months do not process internal features of faces (Bushnell, 1979; Bushnell, Gerry, & Burt, 1983; Maurer & Salapatek, 1976).

One position is that neonates prefer faces because they contain large (i.e., low spatial frequency), high contrast features that are arranged symmetrically and are enticing for infants given their poor visual processing capabilities. Furthermore, face-like stimuli may have an implicit level of stimulus energy that seems to effectively match the perceptual capacities of young infants (e.g., Banks, 1985; Karmel & Maisel, 1975). As infants' perceptual capabilities become more refined during the postnatal period, the reason infants prefer faces may change. That is, as infants gain more experience with faces, preference for face-like stimuli may become less a function of stimulus energy and more a function of social relevance (Dannemiller & Stephens, 1988; Kleiner & Banks, 1987).

Indeed, studies have found evidence that a change occurs between 2- and 4-months in what motivates infant preferences for faces. For instance, Kleiner (1987) found that neonates did not show a preference for a face-like over non-face-like visual display if both contained the same
amount of stimulus energy (as measured by the amplitude spectrum of a Fourier transfer).
Kleiner and Banks (1987) repeated this experiment with 2-month-olds and found those infants
preferred the face-like displays. Similarly, Dannemiller and Stephens (1988) found that 12-week-
old infants prefer schematic faces that are not contrast reversed over ones that are, whereas 6-
week-old infants do not exhibit such a preference. In addition, research has also shown that by 2-
to 4-months of age, infants show a preference for a normal schematic face over the same
schematic face with the location of the features rearranged (e.g., Fantz, 1961; Maurer & Barrerra,

Overall, it seems that a multiplicity of factors determine infant attention to visual events.
While some analyses of single features have been performed (e.g., rate; Balaban & Dannemiller,
1992; Lewkowicz, 1985), other investigations have demonstrated that infant visual attention can
be influenced by multiple stimulus factors (e.g., luminance, structure of the stimulus display,
presence of concurrent cross modal stimulation; Kleiner, 1987; Lawson & Turkewitz, 1980;
Mendelson & Haith, 1976), as well as organismic factors (e.g., level of endogenous arousal;
Gardner & Karmel, 1995; Lewkowicz & Turkewitz, 1981). These types of investigations have
provided us with an enhanced understanding of infant attention to visual events. However, we do
not comparably understand the intricacies of infant attention to auditory events.

Factors Affecting Attention to Infant-Directed Speech

Similar to research on visual perception, research in the field of auditory perception has
focused on discovering what factors guide infant attention. One of the most heavily researched
auditory events is infant-directed (ID) speech. Compared to adult directed (AD) speech, ID
speech is higher in vocal pitch, contains more pitch variability, consists of shorter utterances and
longer pauses between utterances, has a slower tempo/rate, increased phrase repetition, and
increased amplitude (Fernald & Mazzie, 1991; Fernald & Simon, 1984; Jacobson, Boersma,
Fields, & Olson, 1983; Masataka, 1992; 1996; Papousek, Papousek, & Symmes, 1991). These types of vocal changes have been observed in males and females as well as parents and non-parents (Fernald, Taeshcher, Dunn, Papousek, de Boysson-Bardies, & Fukui, 1989; Jacobson et al., 1983). Furthermore, similar modifications have been observed in other languages/dialects such as British-English, French, Italian, German, Mandarin Chinese, Japanese, and sign language (Fernald & Simon, 1984; Fernald et al., 1989; Masataka, 1992; Papousek et al., 1991; Shute & Wheldall, 1989).

Additionally, infants ranging in age from newborn to 18-months prefer ID speech to AD speech (R. Cooper, Abraham, Berman, & Staska, 1997; R. Cooper & Aslin, 1990; 1994; Fernald, 1985; Glenn & Cunningham, 1983; Kaplan, Goldstein, Huckeby, Owren, & R. Cooper, 1995; Pegg, Werker, & McLeod, 1992; Werker & McLeod, 1989). Due to the importance of ID speech in directing infant attention, much research has investigated what factor(s) are primarily responsible for its perceptual draw. Recently, research has focused on how rate of speaking affects infant attention to ID speech.

In one series of experiments (J. Cooper et al., 2001), we presented 1-month old infants with ID-normal speech (i.e., ID speech at the rate infants would normally hear spoken) and ID-fast speech (i.e., ID speech at a rate twice that of the normal speech, equivalent to the rate of AD speech). Importantly, these two sets of recordings were alike in pitch characteristics and differed only in rate of speaking. The infants preferred the ID-normal speech recordings. We next sought to test the idea that a window of effective rates of speaking exists to which infants show maximal attention. That is, if infants heard ID speech that was slower than the normal rate, would they show decreased attention as they did to the ID-fast speech? To test this we created auditory recordings of ID speech at a rate half that of the ID-normal recordings used previously (i.e., twice as slow). As before, all recordings were equivalent in pitch information and differed only in the rate of speaking. Here, we found that infants did not show a significant preference for the ID-normal speech but attended more to the ID-slow speech samples.

We thought perhaps this result was a function of infants not having had sufficient experience with normal ID speech. To investigate this, we repeated the same experimental procedure using 4-month old infants. These infants were tested in the same manner as the 1-month olds and were presented the same ID-normal and ID-slow speech samples. Four-month
olds were selected because by this time infants have had significantly more exposure to maternal speech than at 1-month. In addition, ID speech to 4-month-olds has been found to be more exaggerated than at any other age (Fernald, 1992). Interestingly, we found that the 4-month old infants also preferred the ID-slow speech. This finding is intriguing given infants would likely have never heard ID speech of this nature (i.e., high in pitch but extremely slow). It would be difficult for an individual to speak in an infant-directed manner at such a slow rate and maintain the same pitch level and degree of modulation. Hence, these results do not seem readily attributable to experience or learning. One possibility is that factors inherent to the testing situation affected infant state such that infants preferred slower rates of speaking when they might be more likely to prefer normal rates of speaking under different circumstances.

The procedure we used to test for infant speech preferences involved pairing the speech with a visual display and measuring infant looking-time depending on the auditory event presented. Conceivably the visual display used across the experiments (i.e., a colored bulls-eye) may have stimulated the infant’s perceptual system to a degree where normal rates of infant directed speaking became overstimulating and the slower rate was preferred because it allowed infants to maintain their desired level of excitation. This is based on the assumption that faster speaking rates provide higher levels of stimulation. Perhaps if factors such as complexity and structure of the visual display were different, speech preferences may have been affected. The present investigation was an attempt to explore this possibility.

Intent of Present Investigation

The present study was an attempt to better understand infants’ preferences for ID-slow speech. The first experiment addressed whether infant attention to rate of speaking in the previous experiments could have been affected by the complexity of the visual display used in the experiments. In Experiment 1, ID-normal and ID-slow speech was paired with displays of lower and higher complexity. Given the ideas about how stimulation in one sensory modality (e.g., visual) can affect attention to events in other modalities (e.g., auditory), an interaction was expected between the complexity of the visual display and speech type. It was hypothesized that in the presence of the higher complexity display infants would prefer the slower ID speech, but would prefer the normal rate of ID speech in the presence of the lower complexity display. This hypothesis was based on the idea that young infants seek to maintain a relatively constant level of
stimulation across sensory modalities, and would adjust their attention to events depending on the overall amount of stimulation available.

Whereas the first experiment was designed to investigate how complexity of the visual display affected infant attention to rate of speaking, the second experiment examined the effect of structure on infant attention. In this experiment, ID-normal and ID-slow speech samples were paired with a face-like and non-face-like/scrambled display. An interaction was anticipated between structure of the visual display and speech type, such that infants would prefer slower ID speech in the presence of the scrambled (i.e., non-face-like) display but prefer normal ID speech in the presence of the face-like display.

The present investigation involved infants 3- to 4-months of age for several reasons. For one, Gardner and Karmel (1995) found that manipulations of infant state shown to affect neonates' attention to flashing visual displays (i.e., testing prior to being fed or pre-trial stimulation with 8 Hz. tone) did not have the same effect on 4- to 5-month-olds. That is, the older infants attended more to displays flashed at faster relative frequencies regardless of whether they had been fed prior to testing or received pre-trial stimulation. Hence, the nature of arousal-modulated attention seems to change between birth and 5-months. In order to observe differences in auditory preferences as a function of differing visual stimulation (i.e., complexity of visual display; Experiment 1), it was therefore necessary to study infants younger than 5-months.

In addition, research has shown that infants do not seem to process the structural features of face-like stimuli until around 3-months of age (e.g., Balaban & Stephens, 1988; Maurer & Barrerra, 1981). Therefore, infants older than 3-months were studied in order to address whether the structure of the visual display (face-like or scrambled; Experiment 2) would influence infant speech preferences. Hence, studying infants between 3- and 4-months of age seemed optimal for examining the effect visual display complexity and structure have on infant preferences for rates of ID speaking.
Experiment 1: Complexity of Visual Display

Methods

Participants

The final sample was composed of 24 3- to 4-month-old infants (M age = 108.58 days, SD = 4.68; 12 females and 12 males). An additional 9 infants failed to complete sessions due to excessive fussiness (8) or inattention to the displays (1). All subjects were recruited from the community through birth announcements published in local newspapers. Parents were initially contacted via mail with a letter explaining the nature of the study and subsequently contacted by phone. At time of testing all subjects were reported by their parents to be healthy and full-term with no prenatal or postnatal complications. The mean age of mothers was 29.5 years (SD = 4.15); mean education of the mothers and fathers (in years) was 15.58 (SD = 1.91) and 16.08 (SD = 2.70) respectively; 91.7% of the sample were Caucasian; 91.7% were married; 45.8% had no other children; and the mean gestational age of the infants at birth was 39.54 weeks (SD = 1.76).

Visual Displays

Two visual displays were used for the present experiment, one of higher complexity and one of lower complexity (see Figure 1). Again, complexity was defined as number of contrast elements (e.g., ovals, circles, curved lines) with luminosity being consistent across all displays.

Figure 1

Visual displays used in Experiment 1

![Visual displays](image)
Table 1
Luminosity measurements of visual displays

<table>
<thead>
<tr>
<th>Display</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher Complexity/Face</td>
<td>172.37</td>
<td>118.40</td>
</tr>
<tr>
<td>Higher Complexity/Scrambled</td>
<td>172.37</td>
<td>118.40</td>
</tr>
<tr>
<td>Lower Complexity</td>
<td>172.22</td>
<td>118.25</td>
</tr>
</tbody>
</table>

(see Table 1). The higher complexity display had six elements in a scrambled arrangement. Three of the six internal elements from this higher complexity display were used to create a lower complexity display. These displays and arrangements were modeled after those used previously to test infant preferences for face-like and non-face-like visual displays (Kleiner, 1987; Kleiner & Banks, 1987; Johnson, Dziurawiec, Ellis, & Morton, 1991). Arrangement of elements in all conditions produced vertically symmetrical displays.

Audio Recordings

The same auditory recordings used by Jamie Cooper et al. (2001) were utilized in the present experiment. These auditory recordings were created using Digital Performer, a waveform-editor program for the Power MacIntosh. Using a sophisticated algorithm, this program can affect rate of speaking by manipulating steady-state portions of digitized utterances. In most speech, but particularly in ID speech (Kuhl, Andruski, Chistovich, Chistovich, Lozhevnikova, Ryskina, Stolyarova, Sundberg, & Lacerda, 1997), vowel space tends to be relatively long and stable in regard to pitch and amplitude. By mathematically examining such periods, an algorithm can essentially slow speaking rate by duplicating portions of the vowel space. This produces longer utterances (compared to originals) without affecting pitch characteristics. The two recordings used were ID-normal (unaltered ID speech) and ID-slow (ID speech slowed to 50% of the rate of normal ID speech, see Table 2 and Figure 2).
Table 2
Mean pitch, pitch variability and duration of utterances

<table>
<thead>
<tr>
<th></th>
<th>ID-Normal</th>
<th></th>
<th></th>
<th>ID-Slow</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (Mhz.)</td>
<td>SD (Mhz.)</td>
<td>Duration (sec.)</td>
<td>M (Mhz.)</td>
<td>SD (Mhz.)</td>
</tr>
<tr>
<td>“Good Morning!”</td>
<td>374.27</td>
<td>153.59</td>
<td>1.28</td>
<td>367.44</td>
<td>155.75</td>
</tr>
<tr>
<td>“How are you today?”</td>
<td>317.60</td>
<td>105.30</td>
<td>1.23</td>
<td>323.87</td>
<td>120.16</td>
</tr>
<tr>
<td>“What are you doing?”</td>
<td>373.41</td>
<td>151.31</td>
<td>1.25</td>
<td>362.02</td>
<td>145.53</td>
</tr>
<tr>
<td>“Let’s go for a walk!”</td>
<td>339.72</td>
<td>129.96</td>
<td>1.46</td>
<td>351.51</td>
<td>128.17</td>
</tr>
</tbody>
</table>

Figure 2
Waveforms of “Good morning!” and “Let’s go for a walk!” in ID-normal and ID-slow

Good morning!

Let’s go for a walk!

Apparatus
Each infant was placed on his/her parent’s lap facing a flat black plywood stimulus display (80 cm. x 60 cm.), with floor-length black fabric curtains extending 24 inches from right
and left of the display to create a three-sided enclosure. A 15 inch computer monitor was positioned behind an opening in the plywood panel, offset 7.6 cm to the right of midline and approximately 30 cm away from the infant. A small speaker (Jamo compact 60) was located below the monitor. A video camcorder (Panasonic model AG-180) was also located behind the panel with its lens positioned in a hole (3.5 cm radius) 2 cm left of midline. The camcorder recorded the infant’s responses and provided a view of the infant to an observer watching a 16.7 cm by 14 cm black-and-white television monitor (Magnavox, Model RX4030-WA02).

The observer controlled the presentation of the visual and auditory stimuli and recorded looking times by pressing the spacebar on a Power MacIntosh computer. This computer was running an infant-controlled preference procedure executed by the *MacroMedia Director* software program, which coordinated the presentation of the stimulus events and recorded trial length. The audio output from the computer was sent through an amplifier (Harmon-Kardon Model PM635) and presented through the panel speaker at 60-65 db SPL (B-scale; measured at the infant’s head) as determined by a sound-level monitor (Radio Shack, Cat. No. 33-2050).

**Procedure**

All testing for the present study occurred at the Infant Speech Perception Laboratory at Virginia Tech. Upon arrival, the parent(s) were asked to fill out an informed consent form. After the infant was assessed by the parent(s) and experimenter to be in a quiet, alert state the infant and parent were seated facing the front panel of the 3-sided enclosure and the overhead lights were dimmed.

The parent holding the infant and the observer coding the infant’s looking behavior were deafened to the speech presentations by loud vocal music delivered via headphones. The observer watched the infant’s eye and head movements during the session to determine when the infant was looking at the visual display. Each infant was oriented toward the front of the apparatus prior to each trial by a small flashing red dot on a white background. When the infant was judged to be looking at the screen, the observer initiated presentation of the visual and auditory stimuli by pressing the space bar on the computer keyboard. The visual and auditory stimuli remained present as long as the infant looked at the screen. When the infant was judged to look away from the display (at least 1 s.), the observer pressed the space bar again and the trial
Table 3

Experiment 1: Analysis for Interaction between Complexity of Visual Display and Speech Type

<table>
<thead>
<tr>
<th>Display Type</th>
<th>Lower Complexity</th>
<th>Higher Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech Type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ID-normal</td>
<td>(1)</td>
<td>(3)</td>
</tr>
<tr>
<td>ID-slow</td>
<td>(2)</td>
<td>(4)</td>
</tr>
</tbody>
</table>

*24 Orders = 24 Subjects

was terminated (i.e., presentation of the target visual and auditory stimuli ceased). The red dot was again presented and the same sequence was repeated.

There were four separate display type-speech type pairings: Lower/ID-normal (1), Lower/ID-slow (2), Higher/ID-normal (3), Higher/ID-slow (4) (see Table 3). Presentation of the four conditions alternated across trials and was repeated four times for a total of 16 trials per testing session. The presentation order of conditions was completely randomized; therefore no two infants received the same order of conditions. The primary dependent measure was looking time in each condition.

Results and Discussion: Experiment 1

The display type-speech type pairings were grouped together in blocks of four trials, with each pairing presented once per block. To be included in the final analysis, each infant had to complete at least 12 trials out of a possible 16. The minimum 12 trials of data had to consist of three complete blocks, and therefore four trials of each pairing. In order to determine if there was a preference for certain visual-auditory pairings, mean looking times were calculated for each of the four combinations (i.e., the total looking time to each combination was divided by its respective number of trials). A 4 x 2 x 2 mixed analysis of variance (ANOVA) was computed on the individual subject means, with condition first (higher complexity/ID-normal, higher complexity/ID-slow, lower complexity/ID-normal, lower complexity/ID-slow) as the between-subjects variable and display type (lower complexity, higher complexity) and speech type (ID-normal, ID-slow) as the within-subjects variables. This ANOVA revealed a significant main effect for speech type, $F(1, 20) = 8.82, p < .05$, with the mean looking time to ID-slow (M =
Table 4

Experiment 1: Overall Means and Standard Deviations

<table>
<thead>
<tr>
<th>Variable</th>
<th>Condition</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech Type</td>
<td>ID-normal</td>
<td>13.65</td>
<td>7.36</td>
</tr>
<tr>
<td></td>
<td>ID-slow</td>
<td>19.03*</td>
<td>13.03</td>
</tr>
<tr>
<td>Visual Type</td>
<td>Lower complexity</td>
<td>17.06</td>
<td>11.47</td>
</tr>
<tr>
<td></td>
<td>Higher complexity</td>
<td>15.63</td>
<td>8.92</td>
</tr>
</tbody>
</table>

* p < .05

19.03, SD = 13.03) being longer than that for ID-normal speech segments (M = 13.65, SD = 7.36). No other main effects or interactions were significant.

Because order of condition was not a significant factor, a second 2 x 2 completely within-subjects ANOVA was computed on individual subject means with display type (lower complexity, higher complexity) and speech type (ID-normal, ID-slow) as the independent variables. Excluding the condition variable increased the power of the analysis to detect the hypothesized interaction. However, there was no main effect for display type (F(1, 23) = .63, p > .05), nor was there an interaction between speech type and display type (F(1,23) = .02, p > .05). As above, this ANOVA showed a significant main effect for speech type, F(1, 23) = 7.131, p < .05, with the mean looking time to ID-slow being longer than that for ID-normal speech segments (see Tables 4 and 5). The effect size associated with this difference was .53 (moderate according to Cohen, 1992). In total, 18 of 24 infants had longer looking times to ID-slow speech (p < .01, two-tailed binomial).

The results of the first experiment did not show the hypothesized interaction between speech type and display type. That is, 3- to 4-month old infants did not attend more to the ID-slow speech when it was paired with the higher complexity display versus the lower complexity display. Also, these infants did not attend more to the ID-normal speech when it was paired with the lower complexity display compared to the higher complexity display. However, the results replicated our finding that infants between 1- and 4-months-of-age preferred slower speaking
Table 5
Experiment 1: Means and Standard Deviations for Individual Conditions

<table>
<thead>
<tr>
<th>Speech Condition</th>
<th>Visual Condition</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID-normal</td>
<td>Lower complexity</td>
<td>14.47</td>
<td>8.53</td>
</tr>
<tr>
<td>ID-normal</td>
<td>Higher complexity</td>
<td>12.83</td>
<td>6.19</td>
</tr>
<tr>
<td>ID-slow</td>
<td>Lower complexity</td>
<td>19.64*</td>
<td>14.41</td>
</tr>
<tr>
<td>ID-slow</td>
<td>Higher complexity</td>
<td>18.43*</td>
<td>11.64</td>
</tr>
</tbody>
</table>

* p < .05

rates over faster speaking rates. Therefore, it does not appear from the present results that infant attention to rate of speaking was affected by complexity of the visual display.

Given that complexity is but one of the factors known to affect infant attention to visual displays, it was important to look at the influence of other factors, such as the structure of visual displays (i.e., face-like vs. non-face-like). Around 3- and 4-months of age infants’ preferences for face-like and non-face-like displays seems to shift from being guided by quantitative factors (i.e., complexity) to being guided more by qualitative factors (i.e., structure/meaningfulness). Therefore, the second experiment was undertaken to assess whether structure of the visual displays (i.e., face-like vs. scrambled/non-face-like) would affect infant attention to rate of speaking.
Experiment 2: Structure of Visual Display

In this second experiment, two visual displays were used that contained the same components and were equal in complexity but differed in structure or ‘meaningfulness’ (i.e., face vs. scrambled). As in Experiment 1, the two visual displays were paired with both ID-slow and ID-normal speech, resulting in four separate structure-speech type pairings. It was predicted that the infants in Experiment 2 would prefer ID-slow speech in the presence of the high complexity, geometric display but would prefer ID-normal speech when the same display was reconfigured to form a schematic of a human face.

Methods

Participants

The final sample for Experiment 2 was composed of 24 3- to 4-month-old infants (M age = 106.92 days, SD = 3.44; 10 females and 14 males). An additional 2 infants failed to complete sessions due to excessive fussiness (1) or inattention to the displays (1). All subjects were recruited from the community through birth announcements published in local newspapers. Parents were initially contacted via mail with a letter explaining the nature of the study and subsequently contacted by phone. At time of testing all subjects were reported by their parents to be healthy and full-term with no prenatal or postnatal complications. The mean age of the mother was 29.08 years (SD = 4.33); mean education of the mothers and fathers (in years) was 15.71 (SD = 2.44) and 16.57(SD = 2.97) respectively; 95.8% of the sample was Caucasian; 91.7% were married; 37.5% had no other children; and the mean gestational age of the infants at birth was 39.20 (SD = 1.89).

Visual Displays

Two visual displays differing in structure but not complexity were used in the present experiment. Again, complexity was defined as number of contrast elements (e.g., ovals, circles, curved lines) with luminosity being consistent across all displays (see Table 1). In one condition, the arrangement of the elements produced a schematic face (Face), whereas the other condition contained these components in a scrambled arrangement (Scrambled; see Figure 3). The Higher complexity display used in Experiment 1 was used as the Scrambled display in Experiment 2. Arrangement of elements in all conditions produced vertically symmetrical displays.
**Audio Recordings**

The auditory recordings were the same as in Experiment 1.

**Apparatus**

The apparatus was the same as in Experiment 1.

**Procedure**

The same procedure as Experiment 1 was used in Experiment 2. The only difference was that the parent holding the infant wore opaque, black goggles. So in addition to not being able to hear the speech, they also could not see the display being presented to their infant. There were four separate display type-speech type pairings: Face/ID-normal (1), Face/ID-slow (2), Scrambled/ID-normal (3), Scrambled/ID-slow (4) (see Table 6). As in Experiment 1, presentation of these four conditions alternated across trials and were repeated four times for a total of 16 trials per testing session. Complete randomization of presentation order was again performed such that no two infants received the same order of conditions (see Table 6). The primary dependent measure was still looking time in each condition.

**Results and Discussion: Experiment 2**

To be included in the final analysis, each infant had to complete at least 12 trials (3 trials...
Table 6
Experiment 2: Analysis for Interaction Between Structure of Visual Display and Speech Type

<table>
<thead>
<tr>
<th>Display Type</th>
<th>Face</th>
<th>Scrambled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech Type</td>
<td>ID-normal (1)</td>
<td>ID-normal (3)</td>
</tr>
<tr>
<td></td>
<td>ID-slow (2)</td>
<td>ID-slow (4)</td>
</tr>
</tbody>
</table>

*24 Orders = 24 Subjects

of each display type-speech type pairing) out of a possible 16. In order to determine if there was a preference for certain pairings, mean looking times were calculated for each of the four combinations (i.e., the total looking time to each combination was divided by its respective number of trials). A 4 x 2 x 2 mixed ANOVA was computed on the individual subject means to assess whether there was an interaction between the display type-speech type pairing the infant received first and their preference for the speech or display types. Condition first (face-like/ID-Normal, face-like/ID-slow, scrambled/ID-Normal, scrambled/ID-slow) was the between-subjects variable and display type (face-like, scrambled) and speech type (ID-normal, ID-slow) were the within-subjects variables. The results of this ANOVA showed a significant main effect for display type, $F(1, 20) = 19.85, p < .01$, with the mean looking time to the face-like display ($M = 19.07, SD = 9.20$) being longer than that for the scrambled display ($M = 14.72, SD = 6.51$). No other main effects or interactions were significant.

As in the first experiment, a 2 x 2 completely within-subjects ANOVA was computed on individual subject means with display type (face-like, scrambled) and speech type (ID-normal, ID-slow) as the within-subjects variables. However, there was no main effect for speech type ($F(1, 23) = 1.47, p > .05$), nor was there an interaction between speech type and display type as was hypothesized ($F(1, 23) = .14, p > .05$). The results of this second ANOVA did show a significant main effect for visual display, $F(1, 23) = 16.43, p < .05$, with the mean looking time to the face-like display being longer than that for the scrambled display (see Tables 7 and 8). The effect size associated with this difference was .55 (moderate according to Cohen, 1992). In total, 18 of 24
Table 7

Experiment 2: Overall Means and Standard Deviations

<table>
<thead>
<tr>
<th>Variable</th>
<th>Condition</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech Type</td>
<td>ID-normal</td>
<td>17.79</td>
<td>8.31</td>
</tr>
<tr>
<td></td>
<td>ID-slow</td>
<td>16.01</td>
<td>7.41</td>
</tr>
<tr>
<td>Visual Type</td>
<td>Face</td>
<td>19.07*</td>
<td>9.20</td>
</tr>
<tr>
<td></td>
<td>Scrambled</td>
<td>14.72</td>
<td>6.51</td>
</tr>
</tbody>
</table>

* p < .05

infants had longer looking times to the face-like display (p < .01, two-tailed binomial).

The results of Experiment 2 showed that the hypothesized interaction between speech type and display type was not observed. That is, infants did not attend more to the ID-normal speech when it was paired with the face-like display versus the non-face-like display; nor did they attend more to the ID-slow speech when it was paired with the non-face-like/scrambled display compared to the face-like display. However, unlike the results of the first experiment, there was a main effect for visual type instead of speech type. Therefore, in this experiment,

Table 8

Experiment 2: Means and Standard Deviations for Individual Conditions

<table>
<thead>
<tr>
<th>Speech Condition</th>
<th>Visual Condition</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID-normal</td>
<td>Face</td>
<td>20.22*</td>
<td>10.09</td>
</tr>
<tr>
<td>ID-normal</td>
<td>Scrambled</td>
<td>15.35</td>
<td>6.53</td>
</tr>
<tr>
<td>ID-slow</td>
<td>Face</td>
<td>17.92*</td>
<td>8.32</td>
</tr>
<tr>
<td>ID-slow</td>
<td>Scrambled</td>
<td>14.09</td>
<td>6.49</td>
</tr>
</tbody>
</table>

* p < .05
infants attended more to the visual components of the testing situation and less to the auditory components. The finding that infants preferred the face-like display over the scrambled display replicates an earlier study by Kleiner and Banks (1987) who found that infants older than 2-months preferred a face-like over a non-facelike display, even though both displays contained the same amount of stimulus energy.

General Discussion

The purpose of the present set of experiments was to explore whether 3- to 4-month-old infants’ preferences for ID speech that varied in rate could be affected by pairing the speech with visual displays differing in either complexity (Experiment 1) or structure (Experiment 2). In Experiment 1 it was predicted that infants would look more to a lower complexity display in the presence of the ID-normal speech, but look more to a higher complexity display in the presence of the ID-slow speech. In the second experiment, it was hypothesized that infants would attend more to a face-like display when it was paired with ID-normal speech, and attend more to a scrambled display when it was paired with ID-slow speech (with these two displays being equal in complexity). Neither of these predicted patterns were supported by the data.

Experiment 1 was primarily conducted to gain a better understanding of infant preferences for ID-slow speech observed in previous experiments (J. Cooper et al., 2001). It was thought that if the visual display used in previous work (i.e., a colored bulls-eye) was high in visual stimulation, the infants may have sought out events in other sensory modalities (e.g., auditory) that were relatively less stimulating. The results of Experiment 1 indicated that this may not have been the case. Contrary to what was predicted, infants did not look at the lower complexity display more when it was paired with the ID-normal speech, nor did they look to the higher complexity display more when it was paired with the ID-slow speech. Instead, infants looked more to both types of displays when they were paired with ID-slow speech. There are several interpretations of this preference for ID-slow speech.

It is possible that the lack of interaction between speech type and display type was because the visual displays used in Experiment 1 did not differ enough in terms of complexity to produce the hypothesized preferences. Although the displays were modeled after those previously employed to investigate infant preferences for face-like and non-face-like displays (e.g., Johnson et al., 1991; Kleiner, 1987; Kleiner & Banks, 1987), they have not been used in the same type of
investigation as Experiment 1. However, these displays were created using the same parameters as in past research to produce arousal modulated attention (e.g., McCarvill & Karmel, 1976; Lawson & Turkewitz, 1980). That is, they were equivalent in the ratio of black to white, with one display containing substantially fewer and larger components than the other. Nonetheless, the hypothesized interaction may have been observed if the visual displays differed more in terms of complexity, or if other types of manipulations were invoked. For instance, McCarvill and Karmel (1976) showed that increasing or decreasing level of luminance in the display also affected arousal modulated attention. Therefore, if the lower complexity display had been presented at a lower level of luminance (i.e., decrease level of stimulation), and the higher complexity display presented at a higher level of luminance (i.e., increase level of stimulation), the hypothesized interaction may have been observed. Regardless, it will be necessary to follow up Experiment 1 with an additional assessment of whether infants at this age can discriminate the lower from higher complexity displays.

Assuming that 3- to 4-month-olds can discriminate the displays used in Experiment 1, it is also possible that the infants’ interest in ID speech (especially when it was slow) overshadowed their attention to concurrent visual information. That is, the infants may have simply not attended to any differences across the two visual displays when listening to ID slow speech. However, the results of Experiment 2 argue against this interpretation because infants’ attention to speech was affected by the structural organization of visual information. That is, in the second study, infants attended differentially to the visual information and not to the auditory information. The lack of a preference for the ID-slow speech here was unexpected, especially given that infants preferred ID-slow over ID-normal speech when it was paired with the scrambled/higher complexity display in Experiment 1. Thus, it is not the case that one kind of information (visual vs. auditory) overshadowed the other; rather, characteristics of the visual and auditory events seemed to differentially impact infant attention across the two experiments.

Another possibility is that infants prefer ID-slow speech due to its novelty. As already iterated, it is unlikely that infants would have encountered speech at this rate that was also high overall pitch and pitch modulation. However, previous research does not seem to support a novelty preference. J. Cooper et al. (2001) presented one-month-olds with ID-normal speech and ID-fast speech, and found that infants preferred the ID-normal speech. Even though the ID-fast
speech was also a novel event (i.e., it is also unlikely that infants have encountered speech at this faster rate with the same overall pitch and pitch modulation as normal ID speech), in this case infants preferred the familiar auditory event. Importantly, J. Cooper et al. found that infants of this same age preferred ID-slow over ID-normal. Therefore, it would seem unlikely that a preference for novelty would be responsible for infant attention to ID-slow speech, but not ID-fast.

The results of Experiment 2 also lend support to the idea that infant preference for ID-slow speech is not based on novelty. In that experiment, infants did not prefer ID-slow speech but preferred the more familiar event of a face. Again, it seems unlikely that a novelty preference would guide infant attention in Experiment 1, while a preference for familiarity would guide infant attention in Experiment 2.

It is also possible that even though the complexity of the visual display did not seem to affect infant attention to speaking rate at 3- to 4-months of age, it may have an effect at younger ages. Much of the research demonstrating arousal modulated attention has been conducted with infants three-months of age or younger (e.g., Lawson & Turkewitz, 1980; Lewkowicz & Turkewitz, 1981; McCarvill & Karmel, 1976; Mendelson & Haith, 1976). The only study that has researched infants older than this age found visual attention was not affected by the presence of an auditory stimulus or increased internal physiological arousal at 4- to 5-months of age (Gardner & Karmel, 1995). Therefore, it is possible that the infants in the present study were too old to be affected by manipulations in the complexity of the visual display. If Experiment 1 were repeated with 1-month-olds, it is possible that the hypothesized interaction between display type and speech type would have been observed.

Although the results of Experiment 1 showed infants’ preferences for ID-slow speech were not affected by manipulating complexity of the visual display, the results of Experiment 2 indicated that infants did not differentially attend to the rate of ID speech in the presence of a schematic face. It was hypothesized that infants would attend more to the face-like display in the presence of ID-normal speech, and attend more to the scrambled display in the presence of the ID-slow speech. However, no interaction was found between visual display and speech type, but there was an overall preference for the face-like display independent of speech type.
One possible explanation for the results of Experiment 2 is that the face-like display created a ceiling effect, such that the infants’ attention to the face-like display exhausted their attentional resources. In other words, the face-like display may have maximized infant attention so that the addition of the ID-slow speech was not effective beyond this point. However, the results of Experiment 1 cast some doubt on this explanation. That is, the mean looking time to the face-like display in Experiment 2 was virtually the same as the mean looking time to the lower and higher complexity displays when they were paired with ID-slow speech in Experiment 1. If the face-like display had elevated infant attention to its limit, and the addition of ID-slow speech could not increase looking times, then we would expect to see a mean looking time to the face-like display significantly different than that to the conditions in Experiment 1 where ID-slow speech was paired with less appealing visual displays. However, this was not the case. In fact, the average looking time to the high complexity display during ID slow speech was much lower in Experiment 2 compared to Experiment 1, suggesting that the presence of the face acted to minimize attention to non-face displays. Nonetheless, it may be informative to repeat Experiment 2 using a protocol that may allow infants more flexibility in terms of attentional distribution. For example, infants could be presented with two visual displays of the schematic face for a fixed amount of time. One display would be paired with ID-normal speech, and the other with ID-slow. Infants could then determine by their looking behaviors which type of speech they are presented. This would afford a dynamic assessment of infant attention to rate of speaking in that both speech types would be concurrently available, hence perhaps providing a more sensitive evaluation infant preferences.

Even though no other studies have investigated how the presence of a face affects infant attention to speaking rate, other studies have found that infants behave differently when they are presented with a face-like image in combination with speech versus if they are presented speech and a non-face-like display (e.g., geometric form). For example, Cooper and Aslin (1990) found that newborn and 1-month old infants discriminated and preferred ID speech to AD speech when the speech was paired with a checkerboard display. However, Lewkowicz (1996) found that when 4-month-olds were habituated to a dynamic human face speaking ID speech the infants did not dishabituate when they were presented AD speech. Lewkowicz suggested that prior to 6-months of age, infants do not process the auditory component of face-and-voice events, which
seems relevant to the findings of Experiment 2 in the present study. In contrast, however, McCartney (1999) habituated 4-month-old infants to a dynamic face reciting a story, and found that they dishabituated to a change in speaker gender. Additionally, Werker and McLeod (1989) found that 4- to 5.5-month olds preferred ID speech over AD speech when it was paired with a dynamic face. Therefore, it seems that young infants can process auditory information in the presence of a face-like display, although the likelihood that they will do so may depend on task specifics. One way to test this idea with the present set of stimuli would be to habituate 4-month olds to either ID-slow or ID-normal speech paired with the face-like display, and then switch to the other type of speech. If infants are in fact attending to the auditory information, then we should see dishabituation. However, if the presence of the face is negating infant attention to the auditory component, then we would not expect to see discrimination.

Although the results of the present set of experiments were unexpected and leave several questions unanswered, this study indicates that characteristics of the visual display can indeed affect infant attention to rate of speaking. Therefore, assertions about the power of ID-slow speech to direct infant attention should not fail to address other factors in the testing situation, namely the characteristics of the visual display used in testing. Past research on arousal modulated attention has already established the need for researchers to more fully consider the factors that can affect infant behavior in any testing situation. However, the implications of this investigation for arousal modulated attention remains unclear. Nonetheless, the present study may serve as an impetus for re-evaluating the existing literature on ID speech. That is, in order for researchers to more fully understand the nature of infant perceptual development, it is important to widen the lens of inquiry to include multiple elements of the system of influences on infant attention, including those specific to the testing situation (e.g., visual, auditory) as well as those specific to the infant (e.g., endogenous arousal).
References


PERSONAL INFORMATION

Business Address:  
Division of Human Development  
Department of Psychology  
St. Mary’s College of Maryland  
18952 Fisher Rd.  
St. Mary’s City, MD 20686

Business Phone:  (301) 862-0288
Business Fax:   (301) 862-0436

Home Address:  20904 Jo Marie Way  
Callaway, MD 20620

Home Phone:   (301) 994-3440

E-mail Address:   jscooper@osprey.smcm.edu

EDUCATION

Ph.D.  Virginia Polytechnic Institute & State University, 2001
   Major field: Psychological Science (Developmental)
   Dissertation: Infant attention to rate of speaking in the context of varying levels of visual stimulation

M.S.   Virginia Polytechnic Institute & State University, 1998
   Major field: Psychological Science (Developmental)
   Thesis: The ability of speaking rate to influence infants’ preference for infant-directed speech

M.A.   Tennessee Technological University, 1996
   Major field: Educational Psychology
   Thesis: Psychosocial development in college students: Gender, age, class standing, and other differences.

B.S.   Tennessee Technological University, 1995
   Major field: Psychology
   Summa Cum Laude
**POSITIONS HELD**

- **August 2000 – Present**
  Visiting Instructor; St. Mary's College of Maryland: Dual Appointment in Departments of Psychology & Educational Studies

- **May 1999 – July 2000**
  Course Instructor; Virginia Tech: Department of Psychology

- **August 1997 – July 2000**
  Student Services Advisor, Undergraduate Information Center; Virginia Tech: Department of Psychology

- **June 1997 - August 1997**
  Data Specialist; Virginia Tech Center for Research on Health Behavior

- **August 1996 - May 1997**
  Graduate Teaching Assistant; Virginia Tech: Department of Psychology

- **May 1996 - July 1996**
  Data Specialist; Tennessee Technological University: Counseling Center

- **January 1996 - May 1996**
  Graduate Assistant; Tennessee Technological University: Counseling Center

- **February 1995 - January 1996**
  Teacher/Counselor; FAMILIES Therapeutic Foster Care Agency

- **June 1994 - February 1995**
  Intern; FAMILIES Therapeutic Foster Care Agency

**PROFESSIONAL AFFILIATIONS**

- International Society for Infant Study
- International Society for Developmental Psychobiology
- American Psychological Association
- APA Division 2 [Society for the Teaching of Psychology]
- APA Division 6 [Behavioral Neuroscience and Comparative Psychology]
- APA Division 7 [Developmental Psychology]
COURSES TAUGHT

Introduction to Psychology (Spring 2001 - St. Mary’s College)

Introduction to Women’s Studies (Lecture Participant; Spring 2001 - St. Mary’s College)

Principles of Psychological Research (Fall 1999, Summer 2000 - Virginia Tech)

Developmental Psychology
- Infant & Child Development (Fall 2000, Spring 2001 - St. Mary’s College)
- Infancy through Adolescence (Spring 2000 - Virginia Tech)
- Lifespan Development (Summer 1999 - Virginia Tech; Fall 2000 - St. Mary’s College)

Human Sexuality (Fall 2000 - St. Mary’s College)

Exceptionality: Introduction to Special Education (Spring 2001 - St. Mary’s College)

PUBLICATIONS & ARTICLES IN PREPARATION

Cooper, J. S., Cooper, R. P., Ostroff, W. L. & Aslin, R. (in preparation). Young infants differentially attend to infant-directed speech depending on rate of presentation.


PAPER PRESENTATIONS


POSTER PRESENTATIONS


CONFERENCES & SYMPOSIA ATTENDED


**AWARDS**

Travel Award from the Graduate Student Assembly of Virginia Tech, April 1999.

Small Grant Award from the Graduate Student Assembly of Virginia Tech: endowed to the Developmental Science Society of Virginia Tech, J. Cooper, President; April 1999.

Award, Humanities and Social Sciences Division: 15th Annual Virginia Tech Research Symposium, March 1999.

Travel Award from Carnegie-Mellon University to attend the 29th Carnegie Symposium on Cognition, October 1998.

Small Grant Award from the Graduate Student Assembly of Virginia Tech: endowed to the Developmental Science Society of Virginia Tech, J. Cooper, President; May 1998.

Award, Humanities and Social Sciences Division; 14th Annual Virginia Tech Research Symposium, April 1998.

Travel Award from the Graduate Student Assembly of Virginia Tech, April 1998.