Visual Recognition Memory in 5- and 8-Month-Old Infants and its Relation to Vocabulary Development

Swapna Musunuru, Derek M. Houston and Sarah Pope

Speech Research Laboratory
Department of Psychological and Brain Sciences
Indiana University
Bloomington, Indiana 47405

1 This work was supported by a research grant awarded to Derek Houston (R01DC006235) and a training grant to David Pisoni (T32DC00012). We would like to acknowledge the staff of the Infant Language Lab, especially coordinator Kabreea York, for their assistance in recruiting and testing infants. We are also very grateful to the infants and their parents who participated in the research.

2 Department of Otolaryngology – Head & Neck Surgery; Indiana University School of Medicine, Indianapolis, IN
Visual Recognition Memory in 5- and 8-Month-Old Infants and its Relation to Vocabulary Development

Abstract. Early infancy measures that may predict language and cognitive outcomes are important diagnostic tools when considering early intervention strategies. These infancy measures are especially of value to hearing impaired populations whose language development is often significantly delayed. These populations can benefit tremendously from early intervention. Visual recognition memory is one such infancy measure that relies on measuring short-term memory capacity to predict later cognitive outcomes. In this study, we used a variant of a visual recognition memory task designed by Rose et al. (2001) to test normal hearing and hearing impaired infants’ visual recognition memory. Results suggested that 8.5-month-olds recognized the third object in the largest span length, though 5-month-olds did not demonstrate any notable preferences. A correlation between looking scores and the MacArthur Bates CDI also indicated that infant ability to recognize objects is related to their ability to their early gesturing skills.

Introduction

Finding early predictors for language and cognitive outcomes is important for identifying infants at risk for later deficits, especially since early identification is the first step towards implementing an intervention. Outcome predictors are especially valuable to obtain for hearing impaired populations as several studies have demonstrated the benefits of early intervention for children with hearing impairment. Typically, these populations experience significant delays in language development and academic achievement (Yoshinaga-Itano et al., 1998). Early identification of hearing loss in children prior to six months of age immediately followed by an appropriate intervention has been shown to improve language, speech, and social-emotional development in comparison to children identified later in life (Yoshinaga-Itano, 1999; Downs & Yoshinaga-Itano, 1999; Yoshinaga-Itano et al., 1998). In particular, those with hearing losses identified before six months of age had higher receptive and expressive language quotients, larger productive vocabulary lexicons, increased speech intelligibility, better academic achievement, and less developmental delay compared to those identified at a later age (Yoshinaga-Itano, 1999; Downs & Yoshinaga-Itano, 1999; Yoshinaga-Itano et al., 1998).

Since language abilities are frequently impaired in populations with hearing loss, determining which abilities predict language development is especially important for children with hearing impairments (Miyamoto et al., 2003). In normal hearing infants, later language development may be predicted through expressive and receptive language performances, early phonetic speech perception, speech segmentation ability, as well as various speech processing tasks (Newman et al., 2006; Tsao, Liu, & Kuhl 2004; Hohm et al., 2007). However, these auditory-based measures cannot be used to predict language outcomes in deaf infants; more appropriate measures may be a range of non-auditory cognitive tasks.

Since attention is an important component of learning in infancy, understanding the connection between attention and language outcomes could provide insight regarding which early abilities are predictive of later cognition. Tasks that examine language and cognitive predictors but do not depend on auditory experience have been conducted with normal hearing children, though only a few have examined predictors in hearing impaired populations. Among studies that have examined pre-cochlear
implant predictors, Bergeson and colleagues (2001) demonstrated that pre-implantation Pediatric Sentence Intelligibility scores are strongly correlated with vocabulary, receptive and expressive language, and speech intelligibility scores obtained two years post-implantation. These findings imply that the pre-implantation lip-reading and audiovisual speech perception scores measured by Pediatric Sentence Intelligibility can be used to predict speech and language skills after several years of implant use. Audiovisual speech perception measures may also provide reliable behavioral markers that can be used to predict and identify which children will benefit the most from their cochlear implants (Bergeson et al., 2001). Other studies conducted by Horn and colleagues (2004) have demonstrated the link between motor development and perceptual-motor functions with post implant language measures. Individual differences in visuomotor integration skills in prelingually deaf children are predictive of post-implant performance on open-set speech perception, auditory sentence comprehension, and speech intelligibility skills when measured over three years of cochlear implant experience (Horn et al., 2004). Motor development measures taken from behavioral functioning assessments (specifically the Vineland Adaptive Behavioral Scales) have also been predictive of performance on spoken-word recognition, receptive and expressive language, and vocabulary knowledge tests measured three years post-implantation (Horn et al., 2005). It is important to note that the previously mentioned measures are all indicated for older CI children, since early identification and intervention often does not take place in infancy. As of yet there are not many studies examining early abilities in infancy for children with hearing impairment.

In contrast, a number of studies have looked at non-auditory cognitive predictors for language and cognitive outcomes in children with normal hearing. For instance, multiple components of information processing such as memory, processing speed, attention, representational competence and cross modal transfer have all been shown to predict later differences in IQ and the Mental Development Index (Rose et al., 2005; Rose and Feldman, 1995; Rose & Feldman, 1997). Information processing in infancy is even linked to adolescent intelligence scores (Sigman & Beckwith, 1997) and young adult IQ and academic achievement (Fagan, Holland, & Wheeler, 2007). Studies on infant habituation and novelty preference have demonstrated a link between attention and cognitive outcomes, such that shorter looking times were indicative of better outcomes in childhood (McCall & Carriger, 1993; Colombo, Shaddy, Richman, Maikranz, & Blaga, 2004). In addition, language impairments, such as specific language impairment (SLI), are associated with impaired cognitive processing (Benasich & Tallal, 2002) and impaired auditory processing (Choudhury, Leppanen, Leeyers & Benasich, 2007). In typically developing populations, visual novelty preferences in infancy predict not only IQ, but also language abilities independent of IQ level (Thompson, Fagan & Fulker, 1991).

One particular component of information processing, visual recognition memory (VRM), is highly related to outcomes (Colombo, Shaddy, Richman, Maikranz & Blaga, 2004; Fagan & Krahe McGrath, 1981). Specifically, better information processing is correlated with better visual recognition memory and consequently better cognitive outcomes (Rose et al., 2001; Rose & Feldman, 1997). Visual recognition memory is considered a particularly strong predictive measure for later specific cognitive outcomes and demonstrates that early cognitive processes are important for later language acquisition (Rose et al., 1992; Rose et al., 1991). VRM requires learning to attend to some features and to ignore others and form perceptual categories. Formation of these categories is important for early language acquisition. Children’s abstraction of perceptual features forms the basis for concepts of objects and these concepts need to be in place before language may be acquired (Rose et al., 1991). Specific language outcomes that have been measured through VRM include receptive and expressive language at 2.5, 3, 4, and 6 years, vocabulary scores at 4, 7, and 11 years, IQ at 3, 4, 5, 6, and 11 years and language proficiency at 3 years (Rose et al., 2004; Rose et al., 1991).
The goal of the present study was to extend studies conducted on normal hearing infants to hearing impaired infants and toddlers in hopes of finding pre-implant predictors in infancy and early childhood. Employing a similar span-task paired-comparison paradigm as used in Rose et al. (2001), we tested both normal hearing and hearing impaired infants on their VRM. A series of images was shown for familiarization, followed by the same images paired with new images for the test phase. The target images were in either a series of two or three (spans) and the percentage of time looking at the new image was calculated. This preference for looking at the new image indicates short term memory, or their visual recognition memory. We will then observe how the individual differences in infant VRM scores predict later language outcomes, as measured by the MacArthur Bates Cognitive Development Index and other language measures. It is hoped that by assessing VRM for these infants, we can predict the nature and severity of later language deficit they may incur and consequently implement appropriate early intervention measures.

**Method**

**Participants**

Twenty 5-month-old infants (mean age: 5.0, range: 4.3-5.6) and fifteen 8.5-month-olds (mean age: 8.6, range: 8.1-9.0) with normal hearing participated in this study. All participants were recruited from the greater Indianapolis metropolitan area and passed their newborn hearing screenings.

![Diagram](image.png)

**Figure 1.** During the span tasks, the caregiver sits in a neutral position looking forward and holds the infant in his or her lap. The visual stimuli appear at the left and right stimulus locations. The “attention getter” appears at the center stimulus location.
Apparatus

Testing was conducted in a custom-made double-walled IAC sound booth (see Figure 1). Infants sat on their caregivers’ laps approximately five feet in front of a 55 inch wide-aspect television monitor. Visual stimuli were displayed on the monitor at approximately eye level to the infants. The experimenter observed the infant from a separate room via a hidden, closed-circuit digital camera and controlled the experiment using the Habit software package (Cohen et al. 2004) running on a Macintosh G5 desktop computer.

Stimuli

The stimuli consisted of 24 images of colorful objects. Images were found using an image search on the Internet and were selected if it was unlikely that the infant would already be familiar with the image (e.g. an image of a spoon would not have been selected since it is likely that the infant has seen a spoon before, but an image of a unique candleholder would have been selected). Images were then organized into 12 pairs: 2 pairs for familiarizing the infant to the testing procedure and 10 pairs for the experiment. Image pairs were designed to be easily discriminable from each other yet equal in attractiveness – no single image in a paired set was significantly more intricate or colorful than its corresponding paired image (See Figure 2). In order to create the paired image slide, Photoshop was used to create an initial 12x8 inch blue background template slide with two equally sized white boxes placed side by side on top of the blue background. Next, the individual images from each pair were made an equal size (0.75x0.75 inches) and pasted within the white boxes. This process was repeated for each paired set of images to yield a collection of slides (12 total) that were identical to one another except for the unique image in each white box. Two additional familiarization slides were then made for each corresponding paired image slide. These familiarization slides consisted of the same 12x8 inch blue background slide with a single centrally located white box which was equivalent in size to the paired image slide boxes. Each familiarization slide corresponding to a given paired image slide consisted of one of the two images in the pair. For example, if a paired image slide consisted of images A and A’, two familiarization slides were created - one with image A in the centrally located box, and the other with image A’ in the centrally located box. This was done for counterbalancing purposes so that any bias for one image over the other would be canceled out since half of the infants would be familiarized with image A while the other half was familiarized with A’. All of these images were again resized to 0.75x0.75 inches before they were pasted into the centrally located box. This process was repeated for every paired image slide of the experiment phase and for only one of the two images in each of the two familiarization phase slides, thus yielding a total of 22 familiarization slides (20 for the experiment phase and 2 for the familiarization phase).

Once pairs were created, pairs were further organized into two sets of spans of two and three. Pairs within each span were designed to be highly discriminable from other pairs in the span. Extra care was also taken to make sure that images that appeared in the left boxes were all extremely different from one another, and images that appeared in the right boxes were extremely different from one another.
Figure 2. Images were placed in pairs of equal attractiveness, and then further grouped into two groups of span 2 problems and span 3 problems.

Aside from the 24 stimuli, an “attention-getter” video clip was also created. This clip consisted of a black screen with an animation of a baby laughing located at the center of the screen. This was used to redirect the infants’ attention back to the screen in between trials.

Procedure

The basic design of the experiment consisted of 2 sets of 5 problems each, in spans of two or three items. The procedure for each of these spans followed a paired-comparison paradigm. For each span, the infant was familiarized to two or three images in succession depending on span length, and then given a series of test trials with each successive familiar image now paired with a new image. For example, for a span of two, the infant would first be familiarized with images “A” and “B” in succession, and then tested with A vs. A’ and B vs. B’ (the prime mark denotes the new image in the pair). All spans were presented in ascending order (from span 2 to span 3). Previous studies have demonstrated that ascending versus descending order of span length does not affect outcome of results (Rose et al., 2001).

Two initial pre-test problems were given first to familiarize the infant with the testing procedure. Each problem consisted of a single familiarization slide followed by its corresponding paired image slide, with the novel image being on the right side for pre-test test trial 1 and on the left side for pre-test test trial 2. Each slide was shown for a total of 10 seconds and in between slides the brief “attention-getter” clip was shown to redirect the infant’s attention to the screen.

For the test phase of the experiment, infants were presented with two sets of the five problems mentioned earlier. There was no break between the two sets. Each of the familiarization slides were
shown for a total of 10 seconds, followed immediately by the paired image slides, also each shown for 10 seconds. Familiarization slides were selected at random to ensure that there would be an equal number of paired image slides with the novel stimuli on both the right and left sides to control for side preference. In addition, each infant had a different randomized set of familiarization images, which also helped control error due to side preference and image preference. In between each slide, the infant’s attention was always redirected to the screen by the “attention-getter” clip.

The infant’s looks during testing were recorded by a digital video camera operated by a research assistant who adjusted the camera throughout testing to maintain focus on the infant’s eyes. The digital video recordings were then coded for eye-movements offline, via a frame-by-frame analysis. During coding, the researcher was aware of when the infant was looking at a paired versus a familiarization slide, but was not aware of which side the novel stimuli was located on in the paired slides. During familiarization trials, looks were coded as either “center” or “away.” During test trials, looks were coded as either “left,” “right,” or “away.” The beginning of each trial was determined as the moment that the infant’s face became bright as consequence of the TV monitor shining light from the light-colored study slides into the dark testing room.

Once raw frame-by-frame data was obtained from coding the videos, it was entered into a spreadsheet which then computed the total looking time and longest looks in each direction and to target versus non-target via use of a macro designed to compute these factors. These looking times were then entered into SPSS for further analysis.

**Results**

**Span Three**

Because the number of conditions differed between the span 2 and span 3 strings, the data from each span were analyzed separately. For Span 3, we performed a three-way, repeated measures analysis of variance (ANOVA) with age (5 months, 8.5 months) as a between-subjects factor, and stimulus condition (novel, old), serial position (one, two, three) and phase (one, two) as within-subject factors. The analysis demonstrated a main effect of phase, $F(1,33) = 6.38$, $p = .02$, and an interaction between target, serial position and age group, $F(2,33) = 5.88$, $p = .007$. Due to this interaction we decided to perform two-way repeated measures ANOVAs combined across phases by age group and serial position. Among 5-month-olds in span 3 there were no main effects or interactions. However, in the 8.5-month-old population, there was a main effect of phase in serial position 3, $F(1,14) = 12.9$, $p = .003$, as well as an effect of target that neared significance, $F(1, 14) = 4.30$, $p = .057$. Since there was no significant interaction between target and phase, we grouped across phase and used t-tests to assess looking time differences at each serial position. We found no significant difference in any of the serial positions for 5-month-olds, but there was a significant difference in looking times in serial position 3 for the 8.5-month-olds, $t(14) = 3.59$, $p = .003$, suggesting 8.5-month-olds recognized the final objects in the span 3 problems.

**Span Two**

We decided to analyze the data from span 2 in the same way that we analyzed the data in span 3, in order to compare the phases to each other. The only difference between the analyses was that the serial position factor had three levels (one, two, three) rather than two. In the 5-month-old age group, there was a significant interaction between phase and target in Span 2, serial position 2, $F(1,19) = 4.69$, $p = .043$. Due to this, we analyzed differences in looking time by both phase and serial position for the 5-month-
olds in span 2. However, there were no significant differences in any of the serial positions, regardless of phase. There were also no significant differences in the 8.5-month-olds.

**MacArthur Bates CDI**

Out of the twenty 5-month-olds, 13 had completed the MacArthur Bates CDI (MCDI) in time for analysis. Out of fifteen 8.5-month-olds, 12 subjects completed the MCDI. Correlations between performance on each phase of the recognition memory task and infants’ vocabulary as measured by the MCDI were assessed. Infants’ overall performance in phase 2 correlated significantly with their early gestures ($r = .59, p = .002$), suggesting that infants’ ability to recognize objects is related to their ability to recognize early gestures.

**Discussion**

For the span 2 task we found no statistically reliable evidence of object recognition for either age. These findings are inconsistent with the Rose et al. (2001) study, where a novelty response was demonstrated in the span 2 task. One possible reason for our failure to replicate the findings is that we did not have enough statistical power. Consistent with that possibility, infants’ mean looking times were in the direction of a novelty preference, though they were not statistically significant. Further testing with additional subjects may lead to the trend being statistically reliable.

Differences between these results and Rose et al. may also be due to differences in the nature of the stimuli. Two-dimensional images on a TV screen were used in this study while 3-D objects placed in a tray were used in the Rose study. While 3-D objects are more complex than 2-D images, the 2-D images may have been less stimulating or they may have been more complex to process since they appeared on a TV screen, a presentation mode that the infant may not be used to. This may have contributed to increased task difficulty, and the subsequent familiarity response.

The span 3 task, in contrast to the span 2, demonstrated interesting serial position effects. The 8.5-month age group demonstrated a recency effect, an indication that the images were stored in their short-term memory (as opposed to moving to their long-term memory). Just as demonstrated by Rose et al. (2001), they had a significant novelty response to novel over familiar in serial position 3 of span 3. The 5-month old age group, on the other hand, did not show object recognition at any serial position.

Aside from measuring performances on the span task, another goal for this study was to correlate results from the span task to language development via use of the MacArthur Cognitive Development Index. We found that performance on the span task did correlate with early gestures. This suggests that while our version of the span task may be more difficult for infants that the Rose et al. task, it is possible that our version may serve as a tool for predicting language acquisition in deaf children before implantation. We are currently testing that possibility.

**References**


