An Operantly Conditioned Looking Task for Assessing Infant Auditory Processing Ability

Jason Nawyn, Cynthia Roesler, Teresa Realpe-Bonilla, Naseem Choudhury, April A. Benasich
Center for Molecular & Behavioral Neuroscience
Rutgers, The State University of New Jersey
197 University Ave.
Newark, NJ 07102
+1 (973) 353-1080 ext. 3216
{nawyn, croesler, trealpe, nchoudhu, benasich}@andromeda.rutgers.edu

ABSTRACT
In this paper, we describe the design and evaluation of a gaze-driven interface for the assessment of rapid auditory processing abilities in infants aged 4 to 6 months. A cross-modal operant conditioning procedure is used to reinforce anticipatory eye movements in response to changes in a continuous auditory stream. Using this procedure, we hope to develop a clinical tool that will enable early identification of individuals at risk for language-based learning impairments. Some of the unique opportunities and challenges inherent to designing for infant-computer interaction are discussed.

Categories and Subject Descriptors
J.4 [**Computer Applications**]: Social and Behavioral Sciences---Psychology

General Terms
Design, Experimentation, Human Factors.

Keywords
Infant-computer interaction, eye tracking, operant conditioning, language impairment, early assessment.

1. INTRODUCTION
During early infancy, the critical foundations for phonemic perception and later language are laid down. One of the fundamental skills for the acquisition of language is the ability to process brief, rapidly changing auditory signals occurring within tens of milliseconds. Deficits in rapid auditory processing are characteristic of individuals with language learning impairment (LLI) [1]. An assessment that can accurately identify infants at highest risk for LLI will allow for implementation of early interventions during critical periods of language development. Current research findings indicate that the potential for altering later outcome may be maximally effective during the first year of life [2].

We are currently developing an early assessment tool to help researchers identify individuals at risk for LLI. We propose using an operantly conditioned looking paradigm to evaluate an infant's ability to process rapid successive auditory signals. In this paper we describe the design and implementation of such an application, and present preliminary results from an evaluation study. If the proposed technology is shown to provide a reliable measure of infant processing, we hope to extend its development into a clinically appropriate tool to enable early identification for children at highest risk for LLI.

2. BACKGROUND
Children with LLI have difficulty acquiring language and later reading skills while other cognitive abilities appear relatively intact [3]. It is estimated that roughly 20% of preschool- and school-aged children suffer from deficits in language [4]. More than 50% of children exhibiting early impairments specific to language continue on to develop dyslexia, a disorder associated with lifelong difficulties in the literacy domain [5, 6].

Longitudinal studies with infants offer insights into the etiology of LLI. From birth, infants possess remarkably sophisticated acoustic capabilities allowing the perception of speech as well as non-speech sounds [7]. This phenomenon enables the study of linguistic precursors well before spoken language emerges. Levels of performance on tasks tapping these early precursors (such as rapid auditory processing) have been shown to be predictive of language skills at 16, 24, and 36 months of age [8].

The mechanism by which lower-level processing skills influence later language outcomes likely occurs early on in development, when acoustic and phonological maps are being constructed [9]. Over time, the cumulative effects of poorly encoded representations may result in delay or impairment of language skills. However, if early intervention were successful in increasing efficiency of early auditory processing, it is possible that later language difficulties could be reduced or eliminated.

ASSETS'07, October 15–17, 2007, Tempe, Arizona, USA.

3. RELATED WORK

The design of the current project draws upon insights from a variety of disciplines, including developmental science, eye-gaze tracking, and human-computer interaction.

3.1 Traditional Infant Processing Tasks

Developmental scientists have devised a number of techniques for studying cognition in infancy. For the most part, these techniques enable inferences about certain cognitive skills through the observation of subtle changes in the infant's behavior. Two processes commonly tapped for such studies are ones that emerge early in development: habituation and operant conditioning.

3.1.1 Habituation

Habituation is based on the propensity of infants to differentially attend to novel as compared to familiar stimuli. It is defined as the progressive decrement in attending to one or more stimuli when they are presented repeatedly or are available continuously. As the child becomes more familiar with a stimulus, their interest in it begins to wane. The introduction of a new, novel stimulus will generally attract more of the infant's attention than the reintroduction of one that is familiar. This tendency has allowed researchers to develop habituation and recognition memory tasks that test an infant's ability to discriminate between two similar stimuli. Although these paradigms have been shown to be reliable measures of infant perceptual processing [8], the nature of the task (essentially requiring the infant to become bored with a stimulus) limits its usefulness for rapidly evaluating many stimulus exemplars.

3.1.2 Operant Conditioning

Operant conditioning paradigms rely on the ability of infants to learn the association between a specific behavioral response and subsequent reinforcement. Infancy researchers have exploited this phenomenon to study auditory processing abilities using such techniques as the conditioned head turn task [10]. In the classic head turn procedure, the infant is taught to turn his or her head to look at an animatronic toy that is activated after a novel sound is presented. By manipulating the difference between familiar and novel sounds, researchers can investigate psychophysical performance at various levels of difficulty across an infant population.

The main disadvantage of a head turn procedure for clinical assessment purposes is that it relies on a relatively mature neck and head response, which typically does not emerge until after 6 months of age.

3.2 Eye tracking studies

The developmental primacy of looking behavior makes automated eye tracking technology a valuable tool for the study of cognition and perception in early infancy. Researchers have, for example, used eye tracking to conduct a two alternative forced choice task with a behavioral response that is learned quickly and maintained over many trials [11]. By capitalizing on pre-existing behavioral tendencies such as the visual anticipation of repeated disappearance and reappearance events, such studies have shown eye movements to be a fast, metabolically efficient, and easily detected response mechanism, traits that make eye-gaze interaction appealing as the basis for an early assessment tool.

3.3 Technologies for infant assessment

The vast majority of language-related intervention and assessment technologies for children have focused on school-aged populations. The pool of language-related technologies for infants is quite small.

One such example is the Baby Babble-Blanket (BBB) [12], a system developed to provide infants and developmentally delayed children with a mechanism that enables them to communicate or control their environment. When lying on the blanket, the infants use simple movements, such as head rolling or leg raising to activate digitized sounds. The BBB intervention was used to train cause-effect relationships in 5 month olds with limited physical ability, and was found to be successful at increasing switch activations in response to the sound of the mother's voice.

Another project involving infant communication is the Early Vocalization Analyzer (EVA) [13], a program that automatically analyzes digitized recordings of infant vocalizations. The first goal of EVA is to analyze prespeech utterances in a standardized fashion to eliminate errors in human coding. By comparing the prelinguistic utterances of normal and at-risk infants with a variety of etiologies, EVA can also be used to assess and possibly predict later language difficulties.

EVA is focused primarily on detecting spoken language disorders, whereas the system we propose may be used to identify children at risk for receptive language deficits. By focusing on early auditory processing impairments, we hope to develop a system that can detect individuals at risk for disorders that may not manifest with abnormal vocalizations.

4. SYSTEM OVERVIEW

In this project, we sought to apply the lessons of prior work to the design of an interface that enables assessment of infant auditory processing ability. We exploit the principles of habituation and operant conditioning to set up a response mechanism, but in contrast to prior work, use the relatively more efficient modality of eye gaze.

The final prototype represents the product of an iterative prototyping process in which 15 pilot sessions were conducted with infants aged 3 to 10 months (mean age 5.4 months). General task parameters were based wherever possible on prior infant research [14, 15] and were refined according to observations made on the behavior of the pilot subjects.

4.1 Hardware Configuration

Despite our ultimate ambition to produce a lightweight and cost-effective clinical system, our early prototype is designed to enable maximal control and flexibility for research and development. The system consists of the main components illustrated in Figure 1: a remote eye tracking system with pan-tilt optics and magnetic head tracker from Applied Science Laboratories (Model 504), a designated PC for controlling the eye tracker, and a second PC for controlling the assessment software. The specific interaction sequence for the task is dependent on the infant's behavior, which is determined from point-of-regard coordinates transmitted at 60Hz over a serial data link from the eye tracker control unit.

Visual stimuli used in this task are presented on a 24" wide-aspect LCD monitor located 28" from the subject's head when seated in an infant high chair. The display area distends approximately 40º of visual angle horizontally and 25º vertically. To reduce the likelihood that subjects will be distracted by objects in the testing environment, the ambient lighting is dimmed to a level of approximately 10 lux at the stimulus display screen.

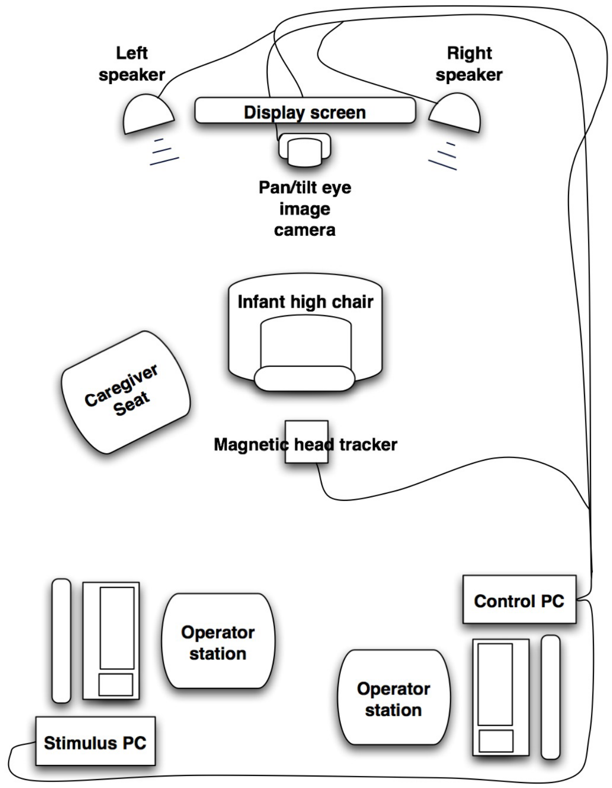

Figure 1. Overall system configuration

During the current system evaluation phase, a minimum of two operators must be present to ensure that the system is functioning properly. A third operator in an adjacent observation room can provide manual blind-coding of infant gaze using a video feed from an infrared camera located at midline below the stimulus display. A parent or caregiver is seated next to (but out of sight from) the infant to attend to any needs that might arise.

4.2 Task Configuration

The overall design of the assessment represents a careful balance among several key elements that must be in place before an operant conditioning procedure will be successful. Specifically, these are (1) a *conditioned stimulus* cue that will be used to signal an upcoming reward, (2) the *reward stimulus* itself, a display that is reinforcing enough to attract the child's attention and motivate sustained engagement with the task, and (3) an *operant response* by which the infant indicates anticipation of the reward.

4.2.1 Conditioned Stimulus (Audio)

Because the goal of this application is to evaluate the infant's ability to process incoming auditory information, a pair of contrastive sounds is used in creating the conditioned stimulus. These two sounds should typically differ on a single acoustic property so that modifying this parameter can result in an array of exemplars that range from very similar to very different. Depending on the specific research questions being explored, such parameters might include frequency, complexity, modulation, amplitude, or various phonological properties of speech.

In the operant conditioning task, one sound is designated as the *familiar* stimulus, and shall remain acoustically invariant. The familiar stimulus is presented repeatedly with a silent inter-trial interval (ITI) of 1.5s between the offset of one instance and the onset of the next. Although the precise ITI may be adjusted based on the nature of the sounds in use, it should be short enough to enable efficient testing, but long enough to ensure the stimuli are processed as discrete units [16]. With an appropriate ITI selected, it is expected that the infant will habituate quickly to the familiar stimulus.

The second sound is designated the *novel* stimulus, and is presented in place of the familiar stimulus at a ratio of approximately 1:5 per recommendations of prior work [17]. It is the novel stimulus that is used initially to alert the child to an upcoming reward as discussed below. Over time the acoustic parameters of the sound can be varied systematically to establish a rough index of the minimal difference the infant is able to detect.

4.2.2 Reward Stimulus (Video)

The reward used in this task is a full-color animated video segment selected from a collection of age-appropriate DVD titles. The video is presented silently to the left of the display screen in a defined area representing 25 by 15 degrees of visual angle. Figure 2(A) shows the blank video display area as it appears during trials with the *familiar* audio stimulus.

During *novel* stimulus trials, the video will appear as shown in Figure 2(B). The reward display lasts 4 seconds before disappearing. The DVD continues to play in the background even when the reward is not visible, so the exact video segment the child will view is varied and non-deterministic. The reward stimulus should be of sufficient complexity and duration to attract the child's attention, but not so stimulating that it interferes with the ability to learn the task.

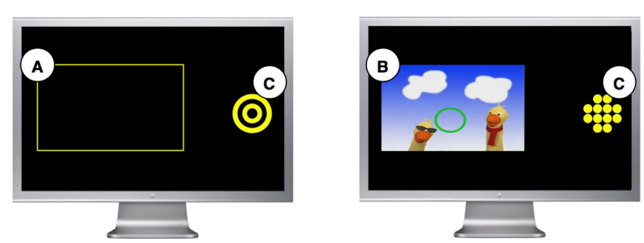

Figure 2. Sample display screens at different points in time showing (A) the empty video display area, (B) the reward video, and (C) the fixation point

4.2.3 Operant Response

A key premise of this task is whether the infant is able to learn the association between the conditioned (novel) stimulus and the onset of the video reward. Successful learning of this contingency is demonstrated when the infant directs his or her gaze into the video display area *after* the novel stimulus is presented, but *before* the reward appears.

To ensure that the subject is not simply staring into the video display area and waiting for the reward to appear, the experimental software will not deliver a novel trial unless the child's gaze is outside the defined area. A fixation point (Figure 2(C)) consisting of geometric patterns that may be animated, pulsing, or flashing is used to distract the child from dwelling in the video area during familiar trials.

4.3 Task Sequence

The timeline in Figure 3 illustrates what a 30 second sample of the task might look like. The familiar (F) audio stimulus recurs 2 to 7 times between each novel (N) trial. If the infant correctly demonstrates the operant response after hearing the novel sound, the video reward is activated immediately. If no response is detected, the system will repeat the novel stimulus up to two times. If no response is received after the second repetition, the system will play the video regardless of infant behavior, thereby reinforcing or reawakening the association between the novel stimulus and the reward. The frequent presentation of the reward video also serves as a short "break" designed to increase the overall amount of time the infant will remain engaged with the task.

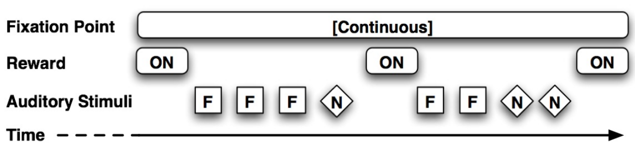

Figure 3. Sample stimulus timeline. 'F' refers to familiar auditory stimuli, 'N' refers to novel stimuli

4.4 Administration Protocol

The operantly conditioned looking task is embedded into an experimental research protocol designed to establish whether the infant can discriminate between the two sounds, and if she is able to learn the operant contingency. A three-phase protocol similar to those used other operantly conditioned tasks [17] is employed in this study.

4.4.1 Training Phase

The first phase of evaluation is used to condition the association between the novel stimulus and the reward. To ensure that the infant can differentiate between the two auditory stimuli, exemplars with the greatest acoustic separation (i.e. the most discriminable) are used. A total of 10 *change trials* are presented. The term "change trials" is simply the name given to those instances wherein the novel stimulus is followed by the reward.

Early in this phase, it is expected that the infant will orient to the video display area only after delivery of the reward, which occurs 1.5 seconds (the length of a standard inter-trial interval) after the second repetition of a novel stimulus. The period of time between the offset of the first novel presentation and the onset of the video is called the scoring interval. As training progresses, infants who learn the stimulus-reward contingency will typically begin looking toward the video area at some point during the scoring interval. This response is scored as a 'hit' and is reinforced by the immediate activation of the video. Change trials in which the infant does not exhibit looking in anticipation of the reward are scored as a 'miss.'

Infants who score 4 or more hits on the last 6 change trials are considered to have trained on the task, and will proceed to the criterion phase. At the discretion of the test administrator, children who did not meet this baseline may repeat the training phase or proceed to criterion.

4.4.2 Criterion Phase

The second phase of evaluation is similar to the training phase, except with the introduction of 10 *no-change* trials interspersed among the change trials. From the infant's perspective, a no-change trial is simply an instance of the familiar stimulus. What differs from the researcher's perspective is the addition of a scoring interval to the familiar stimuli. No-change trials allow the researcher to determine whether the hits generated from the change trials represent correct performance or a looking bias in favor of the video display area. If the infant looks to the reward area during the no-change scoring interval, the trial is scored a 'false alarm.' Otherwise it is scored a 'correct rejection.'

According to traditional infant performance metrics [10], the participant could be considered to have passed criterion when he or she scores at least two hits and two correct rejections on 5 consecutive trials. Upon satisfying this "4/5 rule", the child may move directly into the test phase. In the current protocol, however, all 20 criterion trials are administered and scored regardless of performance.

4.4.3 Testing Phase

The testing phase proceeds similarly to the criterion phase, except that the acoustic separation between the two auditory stimuli is reduced. Depending on the infant's state, up to two 20-trial testing phases can be administered, each with a different level of difficulty. Performance at each difficulty level is reported in terms of the percentage of all trials scored correctly (either hits or correct rejections). Correct percentages above 60% are typically considered suggestive of the fact that the child was able to learn the task and perform the discrimination. Higher values indicate greater confidence that the subject's performance was above chance: at 70%, $t(19) = -2.179$, $p < 0.05$; at 80%, $t(19) = -2.854$, $p = 0.01$.

5. EVALUATION

To a large extent, this project is still a work in progress. Our current objective is to establish whether infants in the target age group are able to learn the task and perform it reliably. We also aim to answer basic questions about the pragmatics of administering the task. How long does it take? How long do infants typically maintain cooperative state? Is this task more appropriate for some ages than others? Are there any patterns in the infant response data that may suggest how to improve the design?

To begin to address these questions, we used word-of-mouth and direct mail recruitment techniques to invite as many infants (and their caregivers) as possible to visit our lab and try out the technology. Visits were scheduled during normal business hours, and were expected to last about an hour. Parents were compensated $25.00 for their time.

After a brief meet-and-greet session during which the caregiver filled out a consent form and short questionnaire, the infant was seated in the high chair and positioned for optimal viewing. During this time, a silent animated video is presented on the stimulus screen.

The video is then repositioned and rescaled in a looming stimulus calibration procedure [18] to verify whether the default eye tracker calibration will be adequate for the assessment task. For most infants it is. If not, up to two saved calibrations are tested before a 2-point "quick" calibration is attempted. If the quick calibration fails, a 9-point "full" calibration is used as a last resort. As a rule, we strive to keep the calibration phase as short as possible to conserve the infant's energy for the experimental procedure.

Once calibrated, the infant was immediately transitioned into the three-phase assessment procedure. Because our intentions at this time were to evaluate the efficacy of the operant looking task rather than study infant processing ability, a relatively simple auditory contrast was used. This strategy helps simplify interpretation of the outcome data by reducing the possible confounding effects that a difficult stimulus might introduce.

Stimuli were constructed as pairs of sinusoidal complex tones with fundamental frequencies of either 800 or 1200 Hz. Each tone included the first 15 harmonics with a 6 dB roll-off per octave. The duration of each tone is 70 ms. The familiar stimulus consisted of two 800 Hz tones separated by a short inter-stimulus interval (ISI). The novel stimulus consisted of one 800 Hz tone and one 1200 Hz tone separated by an equivalent ISI. Similar tone-pair stimuli have been used in work with infants, and have been shown via electrocortical event-related potentials [17] to be discriminable by infants without a family history of language learning disorders.

All sounds were presented at 72 dB SPL from stereo speakers located at each side of the stimulus display screen. The acoustic parameter that was varied between test phases was the duration of the inter-stimulus interval. For training and criterion, an ISI of 200 ms was used, as this slow presentation rate is likely to be discriminable by all infants. Fort the two testing phases, more challenging ISIs of 60 ms and 100 ms were used for Test 1 and Test 2 respectively. To control for likely fatigue effects, the order in which these were presented was counterbalanced.

In order to evaluate how much time infants are willing to spend in the task, we sought to administer as much of the assessment as possible for each child. In practice this meant that most subjects reached a point where their state precluded further testing. Either during the experimental session or afterwards based on video review, the researchers annotated the points in time during which the child's attention was considered alert versus distracted.

After concluding the looking task assessment, the participants were led to another room where a standardized intelligence test was administered.

6. RESULTS

For the research protocol described above, 21 infants under 12 months of age were entered into the evaluation study. Of these, 3 could not be tested satisfactorily due to reasons of temperament, state, technical failure, or methodological irregularities. Of the remaining 18 participants (12 male, 6 female), the mean age was 6m 2d. All were found to be within the normal IQ range on the Bayley Scales of Infant Development (4^{th} Edition) [19]. The population mean for this study was 105 with a standard deviation of 12.4.

Table 1 presents summary data for each of the pilot participants. For criterion and testing phases, the values presented refer to the percentage of correct responses for the total number of trials administered.

Table 1. Data from pilot infants

ID	Age	Sex	Training	Criterion	Test1	Test2
1	2m 25d	M	Pass	90%	[1]65%	55%
2	3m 14d	F	Pass	60%	50%	
3	3m 16d	M	Pass	65%	[1]80%	65%
4	3m 28d	F	Pass	75%		70%
5	4m 15d	M	Pass	60%		
6	4m 18d	F	Pass	80%	50%	
7	5m	M	Pass	80%	75%	
8	5m 14d	F	Pass	75%		70%
9	5m 16d	M	Pass	70%		
10	5m 19d	M	Pass	80%		60%
11	5m 22d	M	Pass	75%	75%	
12	5m 23d	M	Pass	65%		
13	6m 14d	M	Fail	50%	[1]55%	65%
14	6m 26d	M	Pass	80%	60%	
15	8m 10d	M	Pass	75%		55%
16	10m 16d	F	Fail	65%		55%
17	11m 15d	F	Pass	65%		
18	11m 21d	M	Pass	90%	65%	
			Mean	71%	64%	63%

Key: $p < 0.05$ $p \leq 0.01$ [1]First of two tests administered

Overall, 89% of the participants passed the training phase. Under a formal administration protocol, those subjects who failed training would not have proceeded to criterion and testing. These phases were nonetheless administered in the current evaluation to investigate other research questions. The experimental sessions were not terminated until all phases were administered or the infant's state became uncooperative. Fourteen out of 18 infants were calm enough after criterion to receive at least one test administration. Only three of these went on to receive the second test administration.

Group performance on the criterion phase was found to be significantly higher than chance ($t(16) = -8.796$, $p < 0.001$), with 11 of the 16 individuals (69%) who passed training performing significantly above chance. When subjected to the traditional 4/5 rule for criterion, 100% of the trained individuals demonstrated passing performance.

Performance on test phases 1 and 2 was lower than that observed for criterion, but was still significantly higher than chance ($t(14) = -4.750$, $p < 0.001$). No significant differences were observed between ISI test conditions.

On average the complete experimental protocol lasted 11 minutes from the time the infant was placed in the high chair until the session was terminated. This total includes all phases of task administration in addition to time spent on setup, calibration and interaction with the infant. When broken down according to test phases, mean durations were as follows: Training - 2m 15s, Criterion - 3m 14s, Testing - 3m 10s per block.

7. DISCUSSION

As noted earlier, the system evaluation was designed to serve several ends. The first of these is to gather a preliminary dataset from which conclusions might be drawn about its potential to serve its intended function as an auditory processing assessment tool.

7.1 Quantitative Analysis

The data collected in this investigation thus far are suggestive. The proportion of subjects passing the training phase supports the conclusion that an infant can be conditioned to demonstrate anticipatory looking in response to changes in a continuous auditory stream. The high percentage of correct trials observed during the criterion phase suggests further that the operant response is selective to novel auditory stimuli rather than representing a general anticipatory bias.

A within-subjects analysis of the criterion phase also revealed a more nuanced pattern of responses when grouped by age: 2-3 months, 4-5 months, and 6-11 months. Although overall performance was similar for all groups, the youngest infants tended to perform better in the second half of the trials, while the oldest performed better in the first half. Subjects in the middle group performed consistently in both halves.

This observation appears to suggest that younger infants require more time to learn the task, but that they can perform well once they do. In contrast, the oldest infants seem to master the task quickly but then lose interest. Given that optimal performance pattern would be one that is consistent over time, this evidence suggests that the current task may be best suited for individuals aged 4-5 months.

The outcome of the testing phase appears to support a similar interpretation, with the best performance falling in the 3.5-5 month range. However, the overall results of Test 1 and Test 2 are far from conclusive. As designed, the testing phase was intended to provide points of comparison for an infant's performance at multiple levels of difficulty. Because only 3 subjects were cooperative enough to permit administration of both test blocks, no conclusions can be drawn from these data.

When viewed in conjunction with qualitative observations of the child's state, it becomes apparent the decrement in performance from criterion to testing can be largely explained by a progressive loss of state control. After 7 or 8 minutes in the task, the child's attention typically begins to wander, often to their hands or feet or in search of reassurance from the caregiver. Occasionally the infant will re-engage with the display for a short period, but the likelihood decreases rapidly as time passes.

With an average testing session lasting 11 minutes with just one test block, it is clear that performance during the final stages is compromised by infant fatigue. Even under optimal conditions, the administration of one block each of Training, Criterion, and Testing can be expected to last around 9 minutes. Revisions to the task should be implemented to accommodate the likelihood that infants will only engage actively with it for 8 minutes or less. Although it is possible that adjusting certain aspects of the user interface (e.g., the duration of the reward video) might lead to small increases in infant attention, our current findings indicate that restructuring the experimental protocol to include fewer trials in each of the Criterion and Testing phases should allow for delivery of more stimulus exemplars without sacrificing data reliability.

7.2 Design Observations

Although there remains a substantial amount of work to be done before this task can be presumed to have any clinical usefulness, the very act of designing and administering it has proved enlightening. What follows is a discussion of a few key observations gleaned from the experimental sessions with our pilot participants, as well from our work with the infants participating in the iterative design process.

While these observations will be most directly relevant to other researchers interested in early assessment and intervention, it is expected that researchers working on other problems and populations will find some relevance to their own work.

7.2.1 Set realistic expectations for your population.

Non-verbal users such as infants present a unique challenge for interface designers, as the interaction sequence must be one that requires no explicit instruction. Rather, it must be designed to successively elicit appropriate responses by recruiting spontaneous cognitive processes such as visual tracking and auditory habituation.

When designing such an interface, one should as a rule avoid introducing too many variables at once. Presenting multiple new elements simultaneously increases the overall cognitive load of the task, and decreases the amount of attention the user can pay to any single item. For example, our earliest prototypes explored the use of a 2-alternative forced-choice paradigm in which the infant was expected to learn to associate each of two different sounds with a separate visual stimulus. Looking at the correct image when its associated sound is presented would be rewarded.

Although this design is appealing for its psychometric simplicity, we found that infant participants (as well as some adults) were unable to learn and maintain the double association required to perform the task. The change/no-change conditioned looking paradigm ultimately used in our assessment, while less efficient to administer, proved a more viable task for this population.

7.2.2 Find the right amount of stimulation.

Young infants are notorious for their lack of internal state control. In designing an interface for these users, it is essential to achieve an appropriate level of stimulation that both accommodates and responds to changes in the user's state.

In our task, the reward videos were selected to represent the maximum amount of stimulation appropriate for an infant, since its function was to attact the child's attention from elsewhere. Visual overstimulation was not a major concern because of the short display duration, and because no other attentional demands were required during its presentation. However, the potential for understimulation was an issue, particularly for older subjects (those 8 months and over). In time we were able to roughly classify which DVDs were most likely to appeal to infants in various age groups.

The design of an appropriate fixation point was a more delicate exercise. The function of this element is twofold: to distract the infant from dwelling on the empty video display area, and to prevent the infant's gaze from wandering away from the screen during familiar trials. Because successful completion of the task requires the child to process the auditory stimuli while the fixation point is present, it must not be so engaging as to recruit attention away from the auditory signal.

The use of simple animated geometric figures was found to be the optimal compromise based in part on results from an auditory-visual habituation procedure. Images with complex internal features such as faces and even some abstract geometric designs were found to be too engaging, often capturing the child's attention at the expense of the sounds and the video.

7.2.3 Avoid misleading cues.

Setting up a behavioral contingency based only on external stimuli requires careful attention to cues that might detract from the learning process. In our design phase, we discovered that some seemingly innocuous interface features were in fact at cross-purposes with our training goal. The first of these, a chime sound paired with the onset of the video reward was initially intended to orient the infant to the screen in case their gaze had wandered. In practice, however, this chime was much more salient than the conditioned auditory stimulus, and appeared to interfere with the learning of the stimulus-reward contingency. None of the subjects (including adult pilots) would look to the reward until after the chime was presented.

A second misleading association involved the use of auditory spatial localization as a supplementary cue for training the infant's looking patterns. By presenting the familiar stimulus only out of the right-hand speaker, we hoped to strengthen its association with the fixation point on the right-hand side of the screen. We then presented the novel stimulus from the left-hand speaker in hopes that it would be associated with the video display area. It worked. Very well. Our exploratory subjects demonstrated anticipatory looking to the reward area with nearly 100% accuracy during the training phase. However, as soon as the child transitioned into the criterion phase and the localization cue was removed, performance fell to chance.

It appears that the sound localization did not strengthen the audio-visual association as we had hoped, but rather precluded it, transforming the task into a aural tracking exercise instead.

7.2.4 Know your technology and its limitations.

Our stated goal in designing this application is to produce a stand-alone technology that might be deployed in a variety of settings. The use of an automated eye tracking system is a key assumption in this objective, since it theoretically enables a closed-loop system requiring minimal operator intervention. In practice, however, technology for tracking eye gaze in infants may not be ready for widespread adoption.

Although the technology has improved somewhat over the last few years, successfully tracking an infant's gaze still requires an experienced operator and a fair amount of luck regarding the infant's behavioral state. Under optimal conditions, we can capture more than 90% of an infant's looking behavior. Given more typical conditions, that number drops closer to 50%. Because the exact timing of the gaze contingent reinforcment is crucial to learning the association used in our task, we presently only feel comfortable running this task with a human coder providing manual backup scoring in case the eye tracking technology fails.

8. CONCLUSIONS

We believe the results of this research support continued development of the operantly conditioned looking task as a tool for assessing auditory processing ability in young infants. The data collected thus far, though limited, suggest that the basic paradigm is both appropriate for the population and effective enough to produce a workable dataset.

Future work on this project should involve the collection of baseline normative data on a larger and more diverse population of infants. In order to establish its predictive utility as a potential clinical tool, a controlled longitudinal study will be necessary to determine if infant discrimination performance on a variety of stimulus contrasts can predict to later language-based leaning impairments.

At present, we hope that this research, as one of the earliest examples of an interactive application designed for children under 6 months of age, will spur others to consider the potential for new technologies to promote very early assessment and intervention for children at risk of developmental disorders. We hope too that researchers working on other populations with challenging needs might extract useful insights from the methods and observations described herein.

9. ACKNOWLEDGMENTS

This research was made possible by support from the Elizabeth H. Solomon Center for Neurodevelopmental Research and the Don and Linda Carter Foundation. The authors wish to thank Francine Giocondo, Alexandria Reilly, and Nathalie Guthrie for their contributions to this project. Portions of the described apparatus and procedure are protected by patents pending.

10. REFERENCES

[1] Tallal, P. Improving language and literacy is a matter of time. Nature Reviews Neuroscience 2004, 5, 721-728.

[2] Heim, S. and Benasich, A.A. Disorders of Language. In Developmental Psychopathology, 2nd Edition Ed.; D. Chicchetti, D. Cohen, Eds.; John Wiley & Sons: New York, 2006, 268-316.

[3] Tallal, P., Allard, L., Miller, S. and Curtiss, S. Academic Outcomes of Language Impaired Children. In Dyslexia: Biology, Cognition and Intervention; C. Hulne, M. Snowling, Eds.; Whurr Press: London, 1997, 167-181.

[4] Beitchman, J.H., Nair, R., Clegg, M., Ferguson, B. and Patel, P.G. Prevalence of psychiatric disorders in children with speech and language disorders. Journal of the American Academy of Child and Adolescent Psychiatry 1986, 25, 528-535.

[5] Stark, R.E. and Tallal, P. Language, Speech, and Reading Disorders in Children: Neuropsychological Studies; Little, Brown and Co., Inc.: Boston, MA, 1988.

[6] Bishop, D.V.M. and Snowling, M.J. Developmental dyslexia and specific language impairment: same or different? Psychological Bulletin 2005, (130), 858-88.

[7] Aslin, R.N. and Hunt, R.H. Development, plasticity, and learning in the auditory system. In Handbook of Developmental Cognitive Neuroscience; C.A. Nelson, M. Luciana, Eds.; The MIT Press: Cambridge, MA, 2001, 205-220.

[8] Benasich, A.A. and Tallal, P. Infant discrimination of rapid auditory cues predicts later language impairment. Behavioural Brain Research 2002, 136, 31-49.

[9] Benasich, A.A. and Read, H. Representation: Picture or process? Theoretical perspectives in the development of representational (symbolic) thought. In Theoretical perspectives in the development of representational (symbolic) thought.; I.E. Sigel, Ed.; Lawrence Erlbaum Associates, Inc.: Mahwah, NJ, 1998, 33-160.

[10] Kuhl, P.K. Methods in the study of infant speech perception. In Measurement of audition and vision in the first year of postnatal life: A methodological overview; G. Gottleib, N.A. Krasnegor, Eds.; Ablex: Norwood, NJ, 1985, 223-249.

[11] McMurray, B. and Aslin, R.N. Anticipatory eye movements reveal infants' auditory and visual categories. Infancy 2004, 6 (2), 203-229.

[12] Fell, H.J. and Ferrier, L.J. A baby babble-blanket. In INTERACT '93 and CHI '93 conference companion on Human factors in computing systems; ACM Press: Amsterdam, The Netherlands, 1993, 17-18.

[13] Fell, H.J., MacAuslan, J., Chenausky, K. and Ferrier, L.J. Automatic babble recognition for early detection of speech related disorders. In ASSETS 98; ACM Press: Marina del Rey, CA, 1998, 59-66.

[14] Benasich, A.A. and Tallal, P. An operant conditioning paradigm for assessing auditory temporal processing in 6- to 9-month-old infants. In Temporal information processing in the nervous system: Special reference to Dyslexia and Dysphasia.; P. Tallal, A.M. Galaburda, R.R. Llinás, C.v. Euler, Eds.; The New York Academy of Sciences: New York, 1993;Vol. 682.

[15] Benasich, A.A. Auditory temporal processing thresholds, habituation, and recognition memory over the 1st year. Infant Behavior and Development 1996, 19, 339-357.

[16] Fitch, R.H., Read, H. and Benasich, A.A. Neurophysiology of speech perception in normal and impaired systems. In Physiology of the ear, 2nd Ed.; A. Jahn, J. Santos-Sacchi, Eds.; Singular Publishing Group: San Diego, 2001, 651-672.

[17] Benasich, A.A., Thomas, J.J., Choudhury, N. and Leppanen, P.H.T. The importance of rapid auditory processing abilities to early language development: Evidence from converging methodologies. Developmental Psychobiology 2002, 40 (3), 278-292.

[18] McMurray, B. and Aslin, R.N. Automated corneal-reflection eye-tracking in infancy: Methodological developments and applications to cognition. Infancy 2004, 6 (2), 155-163.

[19] Bayley, N. The Bayley Scales of Infant Development: Fourth Edition.; Psychological Corporation: New York, 2006.