RUNNING HEAD: Infants’ Visual Attention to Videos as a Function of Auditory Features of the Program

**Infants’ Visual Attention to Videos as a Function of Auditory Features of the Program**

Author: Marta Perez

Research Mentor: Dr. Sandra Calvert

Senior Psychology Honors Thesis
ABSTRACT

This study examined how the auditory features of video programming affected infants’ visual attention to the screen. Sixty-two children (23 6-month-olds, 19 9-month-olds, and 20 24-month-olds) watched one five minute video in their homes. Their attention to the video was analyzed in relation to the auditory features of the videos. Two types of dependent variables were calculated for each infant: the proportion of time spent looking at the screen while a feature was occurring and the probability of continuing to look at the screen after a feature occurred. Infants were significantly more likely to maintain than to terminate attention after a highly salient event, such as a sound effect or a vocalization. Males attended significantly more during language and music durations and were significantly more likely to maintain attention following sound effects and vocalizations than females. These results support the exploration to search model in which infant’s attention is primarily determined by highly salient program features.
Introduction

In today’s world, many videos are being produced for and marketed to very young children, yet research has yet to uncover how infants attend to features in these videos (Christakis & Zimmerman, 2009). The formal features (production features that structure, mark and represent content) of audiovisual content may influence how children attend to and comprehend the program (Huston & Wright, 1983). Specifically, auditory formal features are unique in that they can influence a child’s attention even when the child is not looking at the screen, while visual features require that the child is looking at the screen in order to observe the features. For example, if a child is not looking at a screen, he or she may here a sound effect or character language that causes him or her to orient their attention back to the program.

Exploration to Search Model

Two models have been proposed to explain how children attend to video programming. The first model is the exploration to search model by Huston and Wright (1983). The exploration to search model concentrates on the perceptual salience of the formal features of the program as the determinant of the attention. This model predicts that children will preferentially attend to and maintain attention to programs that have highly salient formal features, such as rapid pacing, and quick cuts, camera zooms, and visual and auditory special effects. Auditory features that are high in perceptual salience include sound effects and vocalizations. Calvert et al (1982) found that children attended significantly more to salient auditory features than to nonsalient auditory features. Thus, the exploration to search model would predict that infants will attend more to highly
salient auditory features such as sound effects and vocalizations than to nonsalient features, such as language, laughter, and music.

**Comprehensibility Hypothesis**

A second model, formulated by Anderson and colleagues (1981) holds that children selectively attend to program content that they think that they will understand. According to this comprehensibility model, a child will attend more to content that he or she believes is understandable, independent of salient features. Research has found that children older than 18 months attend significantly more to programs in which language is normal compared to programs with distorted, incomprehensible language (Anderson et al., 1981; Pempek et al., 2008). The comprehensibility model as it applies to auditory features focuses primarily on language as the determinant of attention.

**Literature Review**

However, a trend in the body of literature on pre-school aged children and elementary aged children suggests that children move from the exploration to search model to the comprehensibility model as they age. Toddlers attend more to the language of a video (a nonsalient, comprehension driven feature) as age increases from 2.5 years to 5 years (Anderson et al., 1981). Additionally, Pempek and colleagues (2008) found that while distorted language significantly negatively affected pre-school aged children’s attention compared to a program with normal language, this difference did not exist for children younger than 18 months. Additionally, a study by Valkenburg and Vroone (2004) found that salient features stronger predictors of attention for children 6 to 18 months old, while nonsalient features dealing with program content was the primary determinant of attention for children in older age groups. This trend suggests that
language comprehensibility is not the main predictor of attention to media for children who are younger than 18 months.

Prior research has primarily focused on how children two years old and older attend to and process media content. Although Valkenburg and Vroone (2004) included 6-18 month olds in their sample, the study involved exposure to media that was meant for children older than this age group. Research has yet to look at the abundance of media products designed for and marketed to infants and how they attend to the features of this media. Despite the American Academy of Pediatrics recommendation that children under two years be exposed to no media (AAP, 1999), research has found that about half of children in this age group watch DVD’s (Rideout & Hamel, 2006). This study examines how the auditory features of programs designed for infants affect the attention of children who much younger those that have been previously studied, that of six, nine, and twelve-month-olds.

**Hypotheses**

Based on the trends of prior research and the exploration to search model, we predict that children will preferentially attend to auditory features that are high in perceptual salience, particularly sound effects and vocalizations, over features that are low in perceptual salience. We also predict that boys will be more attentive to the videos in general due to prior research finding that pre-school boys pay more attention to media than girls independent of the salience of the programming (Alvarez, 1988).

**METHOD**

*Formal Features*
The formal features of 59 videos marketed to infants were coded in a previous study (Goodrich et al., in press). Of the sample, two of the highest paced videos and two of the lowest paced videos were selected and five minute vignettes were created from each of the four videos. The two low paced videos were Teletubbies and Eebee’s Adventures and the two high paced videos were Baby Signing Time and Baby Road Trip.

Procedure

A previous studied recruited sixty-two children (29 male) aged six (N=23), nine (N=19), and twelve-month-old (N=20) to watch one of the videos in their homes. The child watched either Baby Road Trip, Baby Signing Time, Eebee’s Adventures, or Teletubbies. The infants were recorded while they watched the video. The attention behavior of each child was coded for the amount of time their eyes were on or off the screen. The current study used the attention data of the participants from this prior study. Observer XT, a software program for behavior data, was used to code the auditory characteristics and the attention data of the videos and child’s behavior.

Coding

Auditory features of the four videos were coded so they could be compared to the attention data for each participant. There were two types of auditory features; duration events and point events. Duration events were recorded as the length of time the auditory feature is occurring. There were three duration events: Language (when a character was talking or narrating) Laughter, and Music (background music and foreground music). Point events occurred at a point in time. The two auditory point events were Sound Effects and Vocalizations.

Scoring
The auditory features of each video and each child’s attention to each video were aligned so that the child’s attention during any auditory feature could be determined. Table 1 shows the types of auditory features and their scoring method. For duration events, the proportion of time that a child’s eyes were on the screen while the feature was occurring was compared to the proportion of time a child’s eyes were on the screen while the feature was not occurring. For point events, given that a child was looking at the screen when the point event occurred, the child would either maintain (continue to look) or terminate (look away) in the following three seconds. Analyses used by Alwitt and colleagues were used to calculate the probability that a child would maintain or terminate a look if they were already looking at the screen when a point event occurred (Anderson, Alwitt, Lorch, & Levin, 1979).

RESULTS

Sound Effects

Sound Effects were present in both the High Paced videos, Baby Road Trip and Baby Signing Time. A 2 x 2 x 3 repeated measure ANOVA was used to examine the effect of age and sex on infants’ probability of maintaining or terminating their looks after a sound effect. There was a main effect of sound effect. Infants were significantly more likely to maintain their look to the screen than they were to terminate their look after hearing a sound effect ($F(1,28)=60.78$, $p<.001$). There was also a significant sound effect by sex interaction ($F(1,28)=6.18$, $p<.02$). Post-hoc t-tests revealed that males were significantly more likely to maintain attention than females ($t(17)=11.33$, $p<.001$) and females were significantly more likely to terminate attention then males ($t(15)=3.193$, $p<.01$). These results are depicted in Graph 1.
**Vocalizations**

All four videos contained vocalizations made by characters. Looking behavior for each child in the three seconds following a Vocalization was analyzed using a 2 x 2 x 3 repeated measures ANOVA. There was a main effect of vocalizations. Infants were significantly more likely to maintain their look to the screen than they were to terminate their look after hearing a vocalization ($F(1,61)=52.53$, $p<.001$). There was no effect of age or sex on looking behavior following a Vocalization.

**Language**

Language was present in each video. Language represented both character speaking and narration. There was no significant main effect for language on attention. However, a trend showed that children spent more time looking at the screen while language was occurring than they did when language was not occurring. A repeated measure ANOVA found no effect of age. The ANOVA did find a relationship between attention and sex while language was occurring. There was a significant sex difference between the proportion of time spent looking at the screen while language was occurring, with males spending more time looking at the screen than females ($F(1,56)=4.70$, $p<.04$).

**Laughing**

Laughter was present in three of the videos; Baby Road Trip, Eebee’s Adventures, and Teletubbies. There was no main effect for attention to the screen while laughing was occurring. A repeated measure ANOVA found an effect of video Pace on attention during laughter. There was a significant difference between proportion of eyes on the screen between the High Paced vignette (Baby Road Trip) and the Low Paced vignettes (Eebee’s Adventure’s and Teletubbies). Children spent a longer proportion of time...
looking when laughing was present during the High Paced video than the Low Paced videos ($F(1,40)=1.42, p<.04$).

**Music**

Music was present in three of the videos, with Background Music occurring in Teletubbies and Foreground music occurring in Eebee’s Adventures and Teletubbies. There was no main effect of proportion of time spent looking at the screen while Music was occurring compared to proportion of time spent looking when Music was not occurring. A repeated measure ANOVA found a relationship between Sex and attention to the screen. Males spent significantly more time with eyes on screen when Music was present than females did ($F(1,38)=4.202, p<.05$).

**DISCUSSION**

The purpose of this study was to examine the effects of auditory features on infant’s attention while watching videos. Huston and Wright (1983) proposed the exploration to search model to explain why infants pay more attention to features that are highly salient, such as sound effects and vocalizations. Prior research has found that children spend a significantly more time attending to sound effects and vocalizations, with kindergarten aged children paying more attention to these salient features than other, nonsalient features (Calvert et al, 1982) and with children younger than 18 months attending to highly salient features over content based, nonsalient features (Valkenburg & Vroone, 2004). Our findings are congruent with prior research, but add to the existing literature by focusing specifically on infants and how they attend to highly salient auditory features in programming that is designed for them. These findings support the
exploration to search model, emphasizing that infant’s attention is driven by salient features, such as sound effects and vocalizations.

There was no main effect for language. The proportion of time infants had their eyes on the screen while language was occurring was no different than the proportion of time their eyes were on the screen when language was absent. This effect supports the emerging body of research that young children’s attention to video is not determined by the comprehensibility of linguistic features (Pempek et al, 2008). Prior research found that 2, 3.5, and 5 year old children’s attention was significantly negatively affected by distorting language features in a program, showing that the attention of these pre-school aged children was driven by comprehensibility (Anderson et al, 1981). Research has indicated that toddlers (ages 1-4) attend more to linguistic features as age increases (Anderson and Levin, 1976). Pempek et al (2008) found that children under 18 months did not discriminate between videos with normal and distorted language. Additionally, while language and content was found to predict attention behavior of pre-school children, when children under 18 months watch the same programming, highly salient features better predict their attention than language and content (Valkenburg and Vroone, 2004). Our data shows that children under one year do not attend significantly more when language is present compared to when it is absent. This research and our findings suggest that while pre-school aged children and older may attend to media based on linguistic content, infant’s attention to video is not driven by comprehension, but instead by highly salient features by the exploration to search model.

Laughter, along with language, is a less salient formal feature. There was an effect of the pace of the program during the time that laughter was occurring. High paced
videos are more perceptually salient than low paced videos and highly paced videos elicit more attention than low paced videos (Wright et al, 1984). In congruence with the exploration to search model, young children will attend more to highly salient, high paced videos than less salient, low paced videos. This tendency for children to pay more attention to highly paced videos than lower paced videos, because of their high concentration of highly salient features was found to have an effect when the auditory feature of laughter was occurring. Infants attended significantly more to high paced videos while laughter is occurring than to low paced videos when laughter is occurring. Therefore, the power of the high paced videos, with their high density of highly salient features, seems to affect attention even during a less salient auditory feature of laughter.

A key finding of this study was the influence of sex on attention for both salient and nonsalient features. Males attended significantly more than females during the nonsalient features of language and music. Males were significantly more likely than females to continue looking at the screen and significantly less likely to look away from the screen following the occurrence of a sound effect. Research has found that pre-school aged boys visually attend more than females to video content regardless of the salience of the program (Alvarez, 1988). The characteristics of the program did not make a difference in the Alvarez study, with boys always paying more attention than girls. With the exception of this study, few studies on toddlers have found sex effects for media attention (Anderson, et al., 1981, Pempek, et al., 2008, Valkenburg & Vroone, 2004). Self-diary reports from mothers report that pre-school boys watch more television than girls, but this data does not specify attention to a particular program (Comstock, 1978). Most research on sex differences in media viewing behavior has been centered around
elementary school children. In this age group, boys report that they spend more time watching cartoons than girls (Comstock, 1978). These data on older children and Alvarez’s findings in regards to pre-school children provide evidence that boys are more likely to pay attention to media programming than girls. This may explain why we found effects for sex for nonsalient features and salient features alike. Our findings on sex differences in attention to programming add to research by finding that infant boys attend more to media than infant girls.

There are a few limitations of the present study. Because the four stimulus videos were chosen by their Pacing features, not all of the videos contained all of the auditory features that were analyzed. For example, only Baby Signing Time and Baby Road Trip contained sound effects and laughing and music were each present in three, not all four videos. There was also not a great wealth of diversity in the types of language present. Prior research found significant differences between children’s attention to speaker identity and age, with children attending less to male adult speakers (Anderson and Levin, 1976). Our sample did not contain any adult male language to make such a comparison.

In conclusion, younger children attend preferentially to highly salient features of videos (Calvert et al, 1982). While linguistic comprehension drives attention of preschool and elementary aged children, the same does not appear to be true of infants under two years (Pempek et al, 2008 and Anderson et al, 1981). The current study examined how the auditory features of video programming designed for infants affect infant’s attention to the screen. Consistent with prior research on older children (Calvert et al, 1982), infants attend preferentially to highly salient auditory features of videos, such as sound
effects and vocalizations, than to nonsalient features, such as language, laughter, and music. Regarding language, our results support existing data in finding that visual attention of children under the age of two is not driven by linguistic features (Pempek, et al, 2008). Sex was related to attention for both salient and nonsalient features, with males attending more than females. These findings fit with prior research with older children that found sex differences with males attending more to programming than females (Alvarez, 1988).

In light of the American Academy of Pediatrics recommendation for children under two and the reality of an abundant pool of DVD’s designed for and watched by infants, it is important to understand how infants are watching and attending to media. Our study found that highly salient features were the best predictor of attention for 6, 9, and 12 month olds. Program creators can use this information to create programming that places meaningful or education content in the program after these highly salient features occur, while attention is significantly maintained. Additional research on the effect of formal features on attention could add to this data so that content designed for infants is customized for their viewing behavior.
WORKS CITED


Graph 1. For males and females, the probability of Maintaining a look after a Sound Effect was significantly greater than Terminating attention. Males Maintained their attention significantly more than females. Females were significantly more likely to Terminate their look than males.
<table>
<thead>
<tr>
<th>Feature</th>
<th>Feature Type</th>
<th>Coding Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laughing</td>
<td>Duration Event</td>
<td>Proportion of child is watching the screen while the feature is on</td>
</tr>
<tr>
<td>Language</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Music</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sound Effect</td>
<td>Point Event</td>
<td>Probability that a child will maintain or terminate looking at the screen in</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the 3 seconds following the event</td>
</tr>
<tr>
<td>Vocalization</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Types of auditory features and scoring methods.