Attending to auditory and visual input with flexibility:
Evidence from 4-year-olds

Christopher W. Robinson (robinson.777@osu.edu)
Center for Cognitive Science
The Ohio State University
207D Ohio Stadium East, 1961 Tuttle Park Place
Columbus, OH 43210, USA

Vladimir M. Sloutsky (sloutsky.1@osu.edu)
Center for Cognitive Science
The Ohio State University
208C Ohio Stadium East, 1961 Tuttle Park Place
Columbus, OH 43210, USA

Abstract

Previous research established that infants and young children show a preference for auditory input over visual input. In this research, we hypothesize that young children are flexible attenders, and they may shift their modality preference under different stimuli conditions. The results from the current experiments support the hypothesis that very simple changes in visual stimuli yield attentional shifts in 4-year-olds. Understanding how changes in auditory and visual information influence shifts in attention at various points in development may provide an important tool for understanding lexical and conceptual development.

Introduction

It has been established that very young children attend to non-speech sounds and labels when performing semantic tasks. Using a habituation task, Roberts and Jacob (1991) demonstrated that 15-month-olds were just as likely to form object categories when presented with labels or instrumental music, and Woodward and Hoyne (1999) demonstrated that 13-month-olds were equally likely to associate non-speech sounds or labels with objects in a word learning task, whereas 20-month-olds were less likely to associate non-speech sounds with objects. It appears that very young children attend to a wide range of auditory input, which slowly becomes more refined through learning.

Not only do young children attend to non-speech sounds, but there is also evidence suggesting that auditory input may actually overshadow visual information in infancy. In a series of studies, Lewkowicz (1988a, 1988b) habituated 6- and 10-month-olds to an auditory-visual compound stimulus. At test, infants were presented with the old visual stimulus and a new auditory stimulus or a new visual stimulus and the old auditory stimulus. Although infants in this study were capable of using visual information when presented in isolation, 6-month-olds did not detect a change in the visual component when paired with an auditory stimulus. It appears that any type of auditory information has a privileged status for very young children.

In a recent study, Sloutsky and Napolitano (2003) demonstrated that even children as old as 4-years of age have a preference for information presented to the auditory modality. Here, 4-year-olds were presented with two consecutive auditory-visual compound stimuli and asked to determine whether the two stimuli were the same or different. Across two different sets of visual stimuli, 4-year-olds primarily used auditory information when making same-different responses.

Although infants and young children primarily attend to auditory information, it is also known that children can flexibly shift attention among different stimuli properties (e.g., Smith, Jones, & Landau, 1996). Therefore, it is possible that young children change their modality preference under different stimuli conditions. In particular, it is possible that stimulus simplicity and/or familiarity plays a role in young children's attention to a particular modality. The reported research addresses this issue.

In a series of two experiments, 4-year-olds and adults were presented with two different auditory-visual compound stimuli. During training, these compound stimuli could be used to predict where an animal would appear. After training, two new compound stimuli were created by switching the auditory and visual information, so that auditory information predicted that the animal would appear in one location, and the visual information predicted that the animal would appear in a different location. We also manipulated visual stimuli to determine the stimuli conditions that would lead participants to exhibit auditory or visual preference.
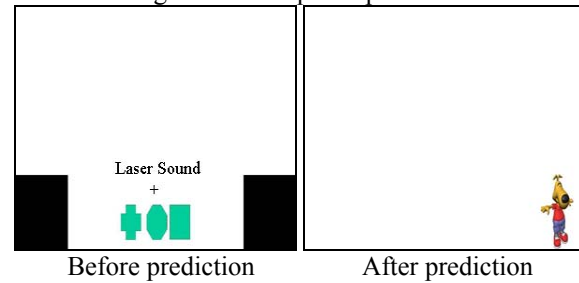
Experiment 1

Method

Participants Twenty-five 4-year-olds (13 Males and 12 Females, $M = 4.59$ years, $SD = 0.34$) and 17 adults (9 Males and 8 Females, $M = 19.39$ years, $SD = 1.77$) participated in this experiment. Young children were recruited through local daycare centers in the Columbus area, and adults participated for course credit.

Stimuli The experiment included two sets of auditory stimuli (A_1 and A_2) and two sets of visual stimuli (V_1 and V_2). The auditory stimuli consisted of a laser sound and a static sound. Each sound lasted 1000 ms in duration and was presented at 68 dB. The visual stimuli consisted of two different three-shape patterns. One of the three-shape patterns consisted of a circle, pentagon, and triangle, and the other three-shape pattern consisted of a cross, octagon, and square. The three geometric shapes were presented in a horizontal line and were identical to stimuli used in Sloutsky and Napolitano (2003). Each geometric shape was of green color and measured 2.54 cm x 2.54 cm in size, and the total three-shape pattern was roughly 10 cm x 5 cm. The stimulus presentation consisted of two auditory-visual training compounds ($V_1 A_1$ and $V_2 A_2$), two auditory-visual test compounds where the auditory and visual information switched ($V_1 A_2$ and $V_2 A_1$), two black panels, and two cartoon-like animals with accompanying melody. The auditory and visual stimuli were perfectly correlated so that the onset and offset of each component occurred at the same time. The two black panels were 5.08 cm x 7.62 cm rectangles and presented at roughly the same height as the compound stimulus. One panel appeared to the left of the compound stimulus and the other panel appeared to the right of the compound stimulus, see Figure 1. The panels were used to mark the location of where animals would appear, see below for a more detailed description of the procedure. After children and adults guessed where the animal would appear, a colorful cartoon-like dog or cartoon-like bird (each roughly 3.81 cm x 7.62 cm) replaced one of the black panels. Both animals, which were animated using Macromedia Flash MX, appeared for 2000 ms. The animation consisted of the animal moving up for 1000 ms and moving down for 1000 ms, resembling a jumping motion. Each animal was accompanied by a short 2000 ms melody.

Figure 1: Example of procedure

| Before prediction | After prediction |

Procedure The procedure consisted of two phases, a training phase and a test phase. During training, participants could use auditory, visual, or both auditory and visual information to predict where an animal would appear. For example, participants may rely on V_1, or on A_1 to predict a dog appearing to the right. At test, participants were presented with a new compound stimulus created by switching the auditory and visual components: they were presented with the $V_1 A_2$ combination and the $V_2 A_1$ combination. If participants primarily use visual information to make inferences about the location of the animal, they should rely on V_1 and V_2, respectively. Alternatively, if they primarily use auditory information, then they should make predictions relying on A_1 and A_2, respectively.

Four-year-olds were presented with a short story at the beginning of the study. Prior to training, 4-year-olds heard: *I have a fun new game where you have to guess where the animal will pop-up. One animal will pop-up here* (pointing to the left panel), *and other will pop-up here* (pointing to the right panel). *I will first give you a clue that will help you to know where the animal will pop-up. Try to use this clue to figure out if the animal will pop-up over here* (pointing to left panel) *or over here* (pointing to right panel). At this point, the experimenter presented $V_1 A_1$ to the child. *When you get this clue an animal will pop-up here* (pointing to animal). *Every time you get this clue, an animal will pop-up here. When you get this clue* (the experimenter presented $V_2 A_2$ at this point), *an animal will pop-up here. Every time you get this clue, an animal will pop-up here. Would you like to play? Here is your first clue.* After the stimulus blinked twice, the experimenter asked: *Where do you think the animal will pop-up?*. If the child did not make a response, the experimenter asked: *Do you think the animal will pop-up over here* (pointing to left) *or over here* (pointing to right)? *If you want to, it's OK to guess.*

Children were tested in a quiet room in local daycare centers using a Dell Inspiron laptop computer. Presentation software was used for stimulus presentation and to record participant's responses. The first two training trials consisted of

the experimenter providing the clues for children, see above for verbal instructions. The experimenter determined the onset of each trial by pressing the space bar. At the beginning of the training trial, one of the auditory-visual compound stimuli appeared for 1000 ms, disappeared for 500 ms, and reappeared for an additional 1000 ms. The two black panels were visible for the entire 2500 ms. If the child did not point to where s/he thought the animal would appear, the experimenter prompted the children by asking, "Where do you think the animal will pop up?". The two black panels remained for an additional 3500 ms or until a response was made. Children's responses were recorded by the experimenter pressing "1" if the child predicted that the animal would appear to the left (i.e., pointing to left panel) or a "0" if the child pointed to the right. After the child guessed where the animal would appear, one of the cartoon-like animals, with accompanying melody, appeared for 2000 ms. Experimenters then provided feedback by saying, "Good job! You got it right. Let's try another one." for correct responses and "Oops, that wasn't the right answer. Let's try another one." for incorrect responses. Children received a total of 16 training trials, with the experimenter explicitly attracting their attention to V_1A_1 and V_2A_2 for the first two trials. The auditory-visual compound stimulus and location of animal were counterbalanced between subjects. Order of stimulus presentation was pseudo-randomized for each subject so that each compound stimulus appeared equally throughout training.

After the 16 training trials, children were presented with 12 pseudo-randomized test trials. At test, the auditory and visual components of the compound stimulus switched so that the auditory input predicted that the animal would appear in one location and the visual input predicted that the animal would appear in a different location. The bird and dog were removed from the test trials, and children did not receive feedback as to whether their responses were correct or incorrect. Prior to test children heard: *Now we are going to play another game that is a little bit different. I am going to give you the same clues that you had in the first part and you will do the same thing as before. This time will be different because you will have to guess where the animal will appear, over here* (pointing left) *or over here* (pointing right), *but you won't see the animal appearing like it did before. This time, you'll have to make all your guesses before the animal appears. If you're not sure how to answer, just guess and when you get through twelve guesses, then the animal will appear on the screen, and that's how you will know you did a good job.* The animal automatically appeared after the last trial regardless of how children did at test. Children received a small prize for their participation.

With several exceptions, the adult procedure was very similar to the procedure used with children. First, adults were instructed to predict where the animal would appear by pressing a "1" if they thought the animal would appear to the left and "0" if they thought the animal would appear to the right, as opposed to pointing to one of the panels. Second, adults were not provided with verbal feedback after each trial, using instead the location of the appeared animal as feedback. Third, adults only received 12 training trials, compared to 16 trials for children. Fourth, inter-trial intervals lasted 1000 ms for adults, compared to experimenter controlled for children. Fifth, adults were not informed that they could use the compound stimulus as a clue to predict where the animal would appear. Finally, adults were not presented with the game scenario and did not receive a prize for completing the study.

Results and Discussion

Participants who correctly predicted where the animal appeared on 4 out of the last 6 training trials or correctly predicted where the animal appeared on the last three training trials were included in the following analyses. Seventeen of the 4-year-olds (68%) and 13 adults (76%) met this criterion. Informal questionnaires revealed that most of the participants who did not meet criterion were trying to detect patterns between trials (e.g., left, left, right, left, left, right, etc.) The proportion of correct predictions during the last six training trials was submitted to a one-way ANOVA with age as a between subjects factor to determine if there were differences in accuracy between the age groups. The proportion of correct responses during training did not differ between the 4-year-olds ($M = .89$, $MSE = 0.03$) and the adults ($M = .95$, $MSE = 0.03$), $F(1,28) = 1.27$, $p > .2$.

Overall, 68% of 4-year-olds' responses and 22% of adults' responses were auditory-based, above chance and below chance, respectively, both one sample $ts > 3.4$, $ps < .01$. A one-way ANOVA with age as a between subjects factor confirmed that the proportion of auditory-based responses at test differed significantly between the 4-year-olds and the adults $F(1,28) = 25.89$, $p < .0001$.

Further analyses focused on individual patterns of responses. Those participants who made at least 8 out 12 auditory responses at test were identified as auditory responders, those who made 4 or less auditory responses were identified as visual responders, and those who made between 4 and 8 auditory responses were identified as mixed responders. Overall, none of the children were categorized as visual responders, seven were mixed responders, and 10 children were categorized as auditory responders. In contrast, nine adults were categorized as visual responders, three were mixed responders, and one adult was categorized as an auditory responder. A chi square analysis revealed

that the numbers of visual, mixed, and auditory responders differed between children and adults, χ^2 (2, N = 30) = 17.75, $p < .001$. The analysis of standardized residuals indicated that children were mostly using auditory information to predict where an animal would appear, and adults were primarily using visual information (all $ps < .05$).

The auditory preference in children and the visual preference in adults could not be explained by an inability for the 4-year-olds to use visual information or an inability for the adults to use auditory information. A control study was conducted using the same stimuli and same methodology as in Experiment 1. However, in the control study, the auditory component of the compound stimulus was removed for the 4-year-olds and the visual component of the compound stimulus was removed for adults. Twenty 4-year-olds (14 Males and 6 Females, M = 4.70 years, SD = 0.46) and 19 adults (7 Males and 12 Females, M = 20.11 years, SD = 2.45) participated in this study. Both the 4-year-olds (M = .76, MSE = 0.04) and the adults (M = .91, MSE = 0.05) had no difficulty using the visual (4-year-olds) or auditory (adults) components when presented in isolation, both $ts > 5$, $ps < .001$.

Using the same visual stimuli as Sloutsky and Napolitano (2003), this study replicated their findings by demonstrating that 4-year-olds are primarily attending to non-speech sounds. However, Experiment 1 extended this by demonstrating that children not only attend to non-speech sounds, but they use this auditory information to make predictions about their world.

Experiment 2

The purpose of Experiment 2 was to determine if small changes in visual information would influence whether 4-year-olds attended to auditory or visual information. In particular, more simple or familiar visual stimuli could shift young children's attention to visual stimuli and away from auditory stimuli. If confirmed, this information would indicate that young children are not modality bound, but that they can flexibly shift their attention between the visual and the auditory modality.

Method

Participants Twenty-three 4-year-olds (9 Males and 14 Females, M = 4.52 years, SD = 0.45) and 15 adults (9 Males and 6 Females, M = 19.53 years, SD = 0.43) participated in this experiment.

Materials and Procedure The experiment was the same as Experiment 1 with one exception. The three-shape geometric patterns (visual component of the compound predictor stimulus in Experiment 1) were replaced by a single geometric shape, either a red triangle or a green cross. It was established in a prior calibration experiment that these stimuli were familiar to young children. The triangle and cross were correctly labeled 79% of the time by 4-year-old children.

Results and Discussion

As in the first experiment, participants who correctly predicted where the animal would appear on 4 out of the last 6 training trials or correctly predicted where an animal would appear on the last three training trials were included in the following analyses. Seventeen of the 4-year-olds (74%) and 12 adults (80%) met this criterion. The proportion of correct predictions during the last six training trials was submitted to a one-way ANOVA with age as a between subjects factor to determine if there were differences in accuracy during training. The proportion of correct responses during training did not differ between the 4-year-olds (M = .85, MSE = 0.03) and the adults (M = .89, MSE = 0.05), F < 1.

Overall, 30% of 4-year-olds and 22% of adults provided auditory-based responses, both below chance, both one-sample $ts < -3$, $ps < .01$. A one-way ANOVA with age as a between subjects factor revealed no significant differences in the proportion of auditory-based responses, $F < 1$. This was in contrast to Experiment 1, where 4-year-olds provided mostly auditory-based responses and adults provided mostly visual-based responses.

As in Experiment 1, children and adults were classified as visual responders, mixed responders, and auditory responders. Overall, 12 children were categorized as visual responders, four were mixed responders, and one child was categorized as an auditory responder. Ten of the adults were categorized as visual responders and two adults were categorized as auditory responders. A chi-square analysis on the numbers of visual, mixed, and auditory responders did not differ between the two age groups, χ^2 (2, N = 29) = 3.77, p = .15.

Given that adults used visual information in Experiment 1 and switching to a single geometric shape increased 4-year-olds' attention to visual information, it is not surprising that adults continued to prefer visual information in Experiment 2. It is unlikely that children and adults were using visual information because of an inability to discriminate between the non-speech sounds or an inability to use non-speech sounds to predict where an animal would appear. Two lines of evidence support this claim. First, children had no difficulty using non-speech sounds in Experiment 1, and second, adults had no problem using auditory information in isolation (control from Experiment 1).

General Discussion

The results from the current experiments expand research concerning attention to auditory and visual information in several ways. First, it was hypothesized that, although 4-year-olds are more likely to attend to auditory information (Sloutsky & Napolitano, 2003), they should flexibly shift between auditory and visual information. When children were presented with novel three-shape patterns, they primarily used auditory information to predict where an animal would appear. In contrast, when children were presented with a single familiar geometric shape, they switched their attention and primarily used visual information, and their pattern of results looked very similar to adults. From these results it could be concluded that, depending on visual and auditory stimuli, young children can flexibly shift their attention between visual and auditory modalities.

The current study also expands previous research concerning the role of labels on conceptual development. Although children weigh labels heavier than appearance when performing semantic tasks such as categorization and induction (Gelman & Markman, 1986; Sloutsky, Lo, & Fisher, 2001), very little is understood about the underlying mechanism. It has been argued that labels are important because they are special in that they mark semantic categories (Gelman & Coley, 1991). It has also been argued that labels are important because they contain prosody, and children as early as 9-months of age are more likely to form categories when presented with prosodic information (Balaban & Waxman, 1997). Although no labels were introduced in the current experiments, results demonstrate that, under certain stimulus manipulations, even 4-year-olds are more likely to attend to non-speech sounds over visual information and use this information to make inferences. It is important to note that this does not imply that labels are unimportant. Rather, these results suggest the possibility that 4-year-old children might not be particularly wedded to labels as top-priority auditory information. That is, they appear to be open to accept more general auditory information under certain visual stimulus conditions. Further research is currently examining the specific quality of these conditions and the underlying mechanism(s) that influences modality preference in young children.

Acknowledgments

References

Balaban, M.T., & Waxman, S.R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology*, *64*, 3-26.

Gelman, S. A., & Coley J. (1991). Language and categorization: The acquisition of natural kind terms. In S.A. Gelman, & J.P. Byrnes (Eds.). *Perspectives on language and thought: Interrelations in development* (146-196). New York: Cambridge University Press.

Gelman, S. A., & Markman, E. (1986). Categories and induction in young children. *Cognition*, *23*, 183-209.

Lewkowicz, D. J. (1988a). Sensory dominance in infants: 1. Six-month-old infants' response to auditory-visual compounds. *Developmental Psychology*, *24*, 155-171.

Lewkowicz, D. J. (1988b). Sensory dominance in infants: 2. Ten-month-old infants' response to auditory-visual compounds. *Developmental Psychology*, *24*, 172-182.

Roberts, K., & Jacob, M. (1991). Linguistic vs. attentional influences on nonlinguistic categorization in 15-months-old infants. *Cognitive Development*, *6*, 355-375.

Sloutsky, V. M., Lo, Y., & Fisher, A. V. (2001). How much does a shared name make things similar? Linguistic labels, similarity, and the development of inductive inference. *Child Development*, *72*, 1695-1709.

Sloutsky, V. M., & Napolitano, A. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development*, *74*, 822-833.

Smith, L.B., Jones, S.S., Landau, B. (1996). Naming in young children: A dumb attentional mechanism? Cognition, 60, 143-171.

Woodward, A.L., & Hoyne, K.L., (1999). Infants' learning about words and sounds in relation to objects. *Child Development*, *70*, 65-77.