Title
Flexible attention and modality preference in young children

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 25(25)

ISSN
1069-7977

Authors
Napolitano, Amanda C.
Sloutsky, Vladimir M.

Publication Date
2003

Peer reviewed

Flexible attention and modality preference in young children

Amanda C. Napolitano (napolitano.7@osu.edu)
Center for Cognitive Science & Department of Psychology
Ohio State University, 208 Ohio Stadium East
1961 Tuttle Park Place, Columbus, OH 43210, USA

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

Abstract

It has been previously established (Sloutsky & Napolitano, 2003) that when presented simultaneously with visual and auditory stimuli equated for discriminability and familiarity, 4-year-olds exhibited strong preference for auditory stimuli, failing to process visual stimuli. At the same time, they had no difficulty processing visual stimuli when these were presented without auditory stimuli. The current study examines the possibility that a flexible attentional mechanism underlies modality preference in young children. We specifically examine under which conditions young children are more likely to process auditory stimuli, and under which conditions they are more likely to process visual stimuli, when both stimuli are presented simultaneously. Results indicate that when visual stimuli are well familiar, 4-year-olds are likely to attend to visual stimuli, whereas when neither visual nor auditory stimuli are familiar, they are likely to attend to auditory stimuli.

Introduction

Auditorily presented information plays important role in young children's cognition. For example, auditorily presented linguistic labels have been found to support categorization (Balaban & Waxman, 1997; Sloutsky & Fisher, 2001), inductive inference (Gelman & Markman, 1986; Sloutsky, Lo, & Fisher, 2001), and similarity judgment (Sloutsky & Lo, 1999) in young children. In particular, if two entities share the label, young children often believe that these entities share many properties, belong to the same kind, and are similar.

In an attempt to explain these effects of labels, Sloutsky & Napolitano (2003) hypothesized that the effects of linguistic labels may stem, in part, from the labels being presented in the auditory modality. They further hypothesized that for young children, auditory stimuli may have higher attentional weights than visual stimuli. To test this hypothesis, they selected visual and auditory stimuli that were equated for discriminability. In one of the experiments (Experiment 1), visual stimuli were unfamiliar landscapes (see Figure 1 for an example of such stimuli), and the auditory stimuli were computer generated three tone patterns. 4-year-olds were trained to consistently select a training set VIS_1AUD_1, comprised of a simultaneous visual and auditory component. In the test phase, the trained set was split, and participants were presented with a choice between $VIS_{new}AUD_1$ and VIS_1AUD_{new}. The majority of 4-year-olds reliably selected $VIS_{new}AUD_1$. Furthermore, a follow up experiment (Experiment 2) indicated that these participants did not even encode the visual components. At the same time, in the absence of auditory components (Experiment 1a), they had no difficulty processing the visual components. These findings were replicated with patterns of geometric shapes (Experiments 3 and 4). Note that young children's performance was in a sharp contrast with that of adults, who were more likely to rely on visual information. It was concluded therefore, that when equally novel visual and auditory stimuli are presented simultaneously, young children are more likely to process auditory than visual information.

At the same time, it is well established that humans are flexible attenders, and under different conditions they may attend to different properties of stimuli (Jones & Smith, 2002; Jones, Smith, & Landau, 1991; Nosofsky, 1986; Smith, Jones, & Landau, 1996). Therefore, we deemed it necessary to establish whether this auditory dominance is fixed, such that it exists under all stimuli conditions, or whether it is flexible, such that it exits under some, but not other stimuli conditions. Of course, some visual and auditory stimuli, such as looming objects or sudden loud sounds, are natural "attention grabbers," and participants are likely to automatically attend to these natural "attention grabbers." However, when stimuli are not natural "attention grabbers," it seems plausible that participants should be able to flexibly shift attention between auditory and visual stimuli.

The goal of this research is to find factors affecting these shifts. In particular, although the visual stimuli used by Sloutsky & Napolitano (2003) were equated with the auditory stimuli, the visual stimuli were perceptually complex (i.e., rich in perceptual detail) and unfamiliar entities that were not individuated objects. Therefore, to address the issues of fixedness or flexibility of auditory dominance, we deemed it necessary to control for "objecthood", while manipulating complexity and familiarity of visual stimuli. Because stimuli used in Sloutsky & Napolitano (2003) were complex and familiar, we used three sets of visual stimuli in the current study: (a) simple and familiar (Condition 1); (b) simple and unfamiliar (Condition 2); and (c) complex and familiar (Condition 3). Examples of these stimuli are presented in Figure 2. Note that familiarity of stimuli was estimated in a separate experiment described below, and complexity was judged as amount of perceptual detail, such as the number of identifiable objects and the number of distinct parts. Another marker of perceptual detail could be the number of brightness contrasts per unit of space. Therefore, color photographs of real objects were deemed richer in perceptual detail than monochromatic geometric shapes, and thus the former were judged more complex than the latter.

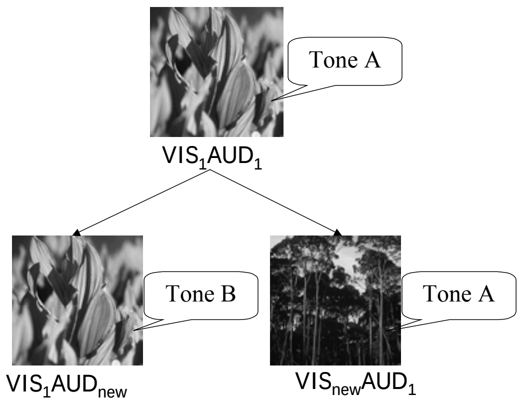

Figure 1: An example of the stimulus sets used in Sloutsky & Napolitano (2003).

Method

The overall study consisted of three between-subjects conditions. Each condition was identical except for the type of visual stimuli used: 1) Condition 1 used simple-familiar shapes, 2) Condition 2 used simple-unfamiliar shapes, and 3) Condition 3 used complex-familiar photographs. In each condition, participants were trained to select a target stimulus set (VIS_1AUD_1) comprised of a simultaneously presented auditory and visual components. If they were able to select the target set to criterion, the target set was broken apart such that the trained image was with a new sound (VIS_1AUD_{new}) and the trained sound was with a new image ($VIS_{new}AUD_1$), and they were asked to continue to pick the "target set". It was argued that selections of the trained auditory stimulus (i.e., $VIS_{new}AUD_1$) would indicate auditory preference, whereas selections of the trained visual stimulus (i.e., VIS_1AUD_{new}) would indicate visual preference.

Two separate calibration studies were conducted for the visual stimuli to determine the discriminability and familiarity of the visual stimuli. Discriminability was established using a same-different immediate recognition task in which a different sample of 14 4 year-olds made same-different judgments after being presented with pairs visual stimuli. Within each trial stimuli were matched by condition and were presented successively for 1 second each in same condition pairs. Participants correctly discriminated visual stimuli in Condition 1 on 91% of trials, in Condition 2 on 86% of trials, and in Condition 3 on 96% of trials. Comparable discriminabilty of auditory stimuli was established previously (Sloutsky & Napolitano, 2003).

Familiarity was established by asking a different sample of 11 4 year-olds two questions about each individual visual stimulus: 1) "Have you ever seen one of these?", and 2) "What is it?". For Condition 1, children recognized the stimuli on 84% of trials and correctly labeled them on 82% of trials. For Condition 2, children recognized the stimuli on 25% of trials and attempted to label them on 23% of trials, although labels differed across participants. For Condition 3, children recognized the stimuli on 96% of trials and correctly labeled them on 96% of trials. Based on these responses, it was concluded that the visual stimuli in Conditions 1 and 3 were familiar, whereas the visual stimuli of Condition 2 were unfamiliar.

Participants

Participants were 45 young children (mean age = 4.41 years, *SD* = 0.346 years; 19 girls and 26 boys) recruited from childcare centers located in middle class suburbs of the Columbus, Ohio area. The overall sample was divided into three groups of 15, and each group participated in only one condition.

Materials

Materials consisted of stimulus sets, each comprised of a visual and an auditory stimulus. The individual auditory and visual stimuli were combined randomly into cross-modal sets. Each set was comprised of a simultaneous presentation visual and auditory component, which was created by pairing an auditory stimulus and a visual stimulus so that they were perceived as one unit. Each stimulus set was presented for 1 second.

For each condition a total of 16 stimulus sets were used. Within each condition, there were four different types of stimulus sets created: 1) the training target set, VIS_1AUD_1, that participants were trained to select, 2) VIS_2AUD_2 that was presented as a distracter during training, 3) VIS_1AUD_{new} that matched the training target's visual component, but had a novel auditory component, and 4) $VIS_{new}AUD_1$ that had a novel visual component, but matched training target's auditory component.

The auditory stimuli were computer generated patterns, each consisting of three unique simple tones. Simple tones varied on timbre (sine, triangle, or sawtooth) and frequency (between 1 Hz and 100 Hz). Each simple tone was 0.3 seconds in duration and was separated by .05 seconds of silence, with total pattern duration of 1 second. The average sound level of auditory stimuli was 67.8 dB (with a range from 66 dB to 72 dB), which is comparable with the sound level of human voice in a regular conversation. These are the same tones patterns used in Sloutsky & Napolitano (2003). Visual stimuli were different in each condition as described below. Examples of visual stimuli for each of the three conditions are presented in Figure 2.

Visual stimuli for Condition 1 (simple and familiar). The visual stimuli for Condition 1 were computer-generated single two-dimensional geometric figures. Each shape was 4 inches x 4 inches in size and was colored green.

Visual stimuli for Condition 2 (simple and unfamiliar). The visual stimuli for Condition 2 were computer-generated two-dimensional figures that were created by randomly coloring in 1 inch x 1 inch squares of a 4 x 4 grid such that: 1) each column had at least one colored square, 2) all colored squares were connected, and 3) no squares that had a colored square on all four sides could be uncolored. Gridlines were removed to create a continuous shape. Each shape was 4 inches x 4 inches in size and was colored green.

Visual stimuli for Condition 3 (complex and familiar). The visual stimuli for Condition 3 were photographs of animals. Each photograph was 4 inches x 4 inches in size and varied in color.

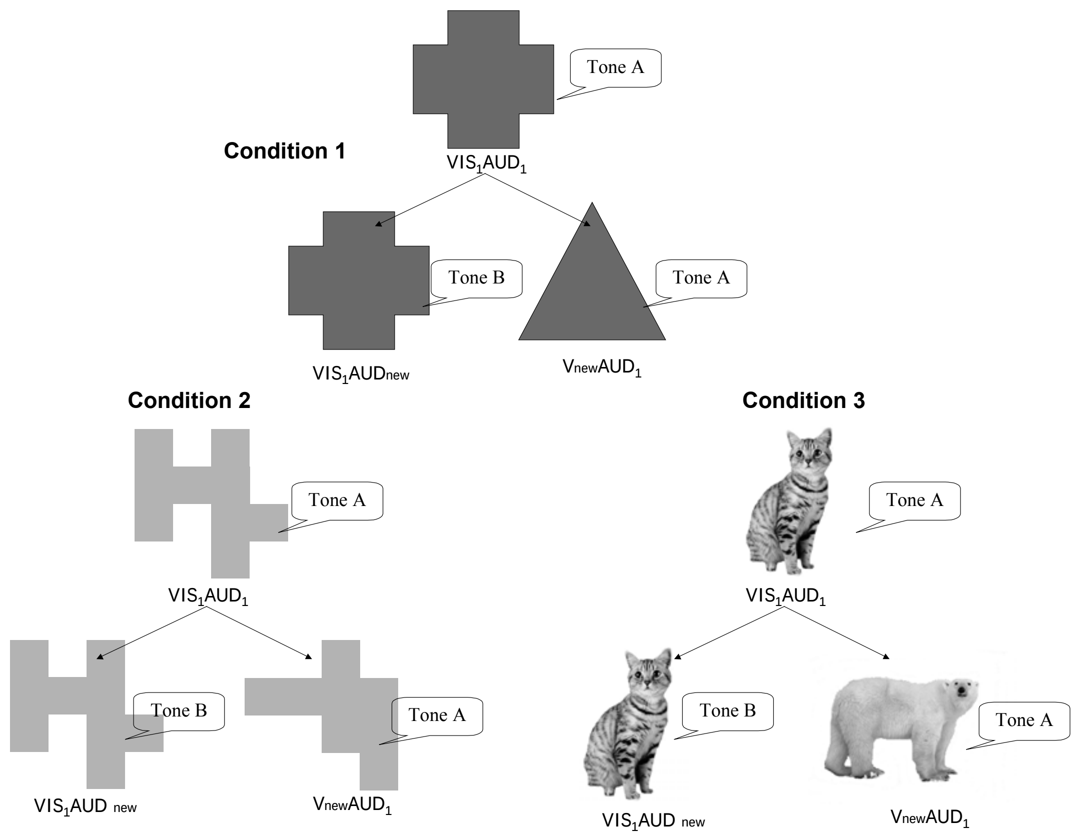

Figure 2: Examples of stimuli for all 3 conditions

Design and Procedure

Participants were tested in a quiet room within their daycare center. They were told that they would play a game, in which they should find the location of a prize, and that they would be rewarded at the end of the game with a prize. Small toys were given at the end of each day for their participation.

For each of the three visual conditions, the overall experiment included 4 blocks, with each block consisting of 8 training trials (a training session) and 6 test trials (a testing session). Participants were presented with 2 blocks per day, and the experiment was spread over a 2-week period. All stimuli were presented on a Dell Inspiron laptop computer, and presentation of stimuli and recording of responses was controlled by a Visual Basic program.

Training session Stimuli were presented in the following manner. First, either VIS_1AUD_1 or VIS_2AUD_2 was presented on one side of the screen, followed by the presentation the remaining stimulus set (i.e., either VIS_1AUD_1 or VIS_2AUD_2) on the other side of the screen. The order of appearance and the side of the screen for which to appear was counterbalanced across training trials for both the two stimulus sets, such that each set could appear either first or second and on either the right or left side of the screen. A white circle icon replaced each set at the end of its presentation. The goal of training was to teach the child to consistently select the VIS_1AUD_1 stimulus set, and, therefore, on each trial the child was provided with "yes" feedback when this stimulus set was chosen, and "no" feedback when the VIS_2AUD_2 stimulus set was chosen. Only participants making correct selections in the final four trials moved into the test session.

Test session The test session followed immediately after the training session, during which participants were presented with two novel stimulus sets. Set VIS_1AUD_{new} matched the training target's visual component, but had a novel auditory component, whereas set $VIS_{new}AUD_1$ had a novel visual component, but matched the training target's auditory component. The participants were asked again to identify the set where a prize was hidden. Again, the positions of the two stimulus sets were counterbalanced across test trials, and a white circle icon replaced each set at the end of its presentation. When the selection was made, the experimenter pressed the keyboard key corresponding to the selection, without giving feedback to the participant. The overall structure of training and testing trials is presented in Table 1.

Table 1: The overall structure of trials in one block

Training Session (n = 8 trials)		Testing Session (n = 6 trials)	
VIS_1AUD_1 (Trained Set)	VIS_2AUD_2 (Distracter Set)	VIS_1AUD_{new} (Test Set A)	$VIS_{new}AUD_1$ (Test Set B)

Results

Proportions of selections for $VIS_{new}AUD_1$ (i.e., selections of auditory trials) were subjected to a one-way ANOVA with condition as a factor. There was a significant main effect of Condition, F (2, 42) > 13.85, p < .0001. A post-hoc Tukey test pointed to significant difference between Condition 2 and Conditions 1 and 3, ps < .0001, such that participants were more likely to select $VIS_{new}AUD_1$ in Condition 2, but not Conditions 1 and 3. Proportions of auditory trials per condition are presented in Figure 3.

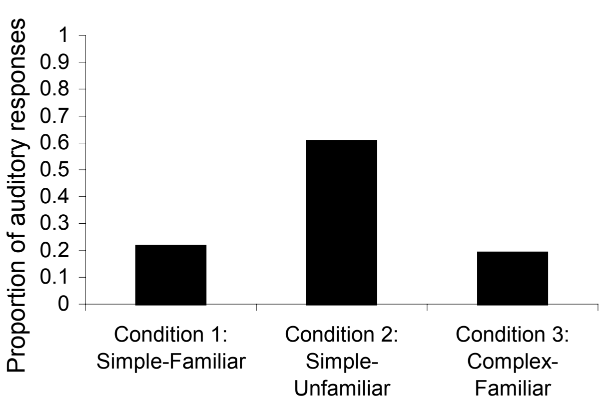

Figure 3: Proportions of auditory responses by condition

As a more conservative analysis of the participants' performance, we calculated the number of blocks with above-chance reliance on auditory stimuli, above-chance reliance on visual stimuli, and chance performance. Performance was considered above-chance if the same choice was made on 5 out of 6 trials (Binomial Test, p = .09), otherwise it was considered at or below chance. Blocks with above-chance auditory responding were judged as exhibiting auditory preference, blocks with above-chance visual responding were judged as exhibiting visual preference.

In Condition 1, out of 60 blocks, participants successfully completed the training phase of 48 blocks, with 11% of successfully completed blocks exhibiting auditory preference, 66% exhibiting visual preference, and 23% were at chance. In Condition 2, out of 60 blocks, participants successfully completed 44 blocks, with 36% exhibiting auditory preference, 30% exhibiting visual preference, and 34% being at chance. In Condition 3, out of 60 blocks, participants successfully completed 55 blocks, with 2% exhibiting auditory preference, 65% exhibiting visual preference, and 33% being at chance. These results indicate that there were significantly more above-chance auditory blocks in Condition 2 than in Conditions 1 and 3, χ^2 (2, N = 147) = 24.1, p < 0001.

We also analyzed the individual responses which fit into one of three distinct patterns: (1) participants who were above chance in relying on auditory stimuli (auditory responders); (2) participants who were above chance in relying on visual stimuli (visual responders); and (3) participants who were at chance (mixed responders). Above chance performance was determined by subjecting the total number of auditory and visual choices made by each individual to the binomial test. Percentages of responders' types across age groups are presented in Table 2.

Table 2: Percentages of responder types by condition.

Condition	Responder Type		
	Visual	Auditory	Mixed
1	73.3%	6. 7%	20%
2	21.4%	42.9%	35.7%
3	73.3%	0%	26.7%

Discussion

Overall, results point to strong familiarity effects: when visual stimuli were familiar, participants were more likely to exhibit auditory preference, whereas when visual stimuli were unfamiliar, participants were more likely to exhibit auditory preference. These results, in conjunction with earlier findings (Sloutsky & Napolitano, 2003), point to a flexible attentional mechanism, underlying processing of auditory and visual stimuli. Of course, this conclusion goes beyond data at hand, because the reported studies did not manipulate auditory stimuli. However, it seems likely that if unfamiliar visual stimuli are paired with familiar auditory stimuli (e.g., bird's calls or dog's bark) participants may be even more likely to rely on the auditory stimuli than they did in Condition 2 and Sloutsky & Napolitano (2003, Experiment 1), where both visual and auditory stimuli were unfamiliar.

These findings further support the auditory dominance explanation of the role of linguistic labels (Sloutsky & Napolitano, 2003). Recall that according to this explanation, some of the effects of linguistic labels on categorization, induction, and similarity may stem from the modality of input. Note that in the majority of previous research, young children were presented either with familiar entities and familiar sounds of human speech (e.g., Gelman & Markman, 1986) or with novel entities and familiar sounds of human speech (e.g., Sloutsky, Lo, & Fisher, 2001) and asked to induce novel biological properties. These inductions could be made either on the basis of perceptual similarity (i.e., one choice stimulus looked like the Target) or on the basis of common label (i.e., another choice stimulus had the same label as the Target). Findings that under both novelty conditions, matching labels had larger effects than perceptual similarity on young children's induction are consistent with the current findings suggesting that when stimuli have comparable familiarity, young children are more likely to attend to auditory stimuli than visual stimuli, whereas if novelty is different, they are more likely to more familiar stimuli.

The reported results have important implications for our understanding of the role of attention in processing. First, attention can be flexibly shifted between auditory and visual stimuli depending on familiarity of stimuli: familiar stimuli are more likely to be attended to than unfamiliar stimuli. Second, given comparable familiarity, young children are more likely to process auditory stimuli than visual stimuli. This property of young children's attention is in a sharp contrast with adults' attention: under the comparable familiarity conditions, adults are more likely to attend to visual stimuli (Sloutsky & Napolitano, 2003). Of course, to ascertain that the familiarity effects hold for familiar auditory stimuli (e.g., human speech, bird calls, or car sounds), additional research manipulating the familiarity of auditory stimuli is needed.

Acknowledgments

This research has been supported by a grant from the National Science Foundation (BCS # 0078945) to Vladimir M. Sloutsky.

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

Jones, S. S., & Smith, L. B. (2002). How children know the relevant properties for generalizing object names. *Developmental Science*, *5*, 219-232.

Jones, S. S., Smith, L. B., & Landau, B. (1991). Object properties and knowledge in early lexical learning. *Child Development*, *62*, 499-516.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39-57.

Sloutsky, V. M., & Fisher, A. V. (2001). Effects of linguistic and perceptual information on categorization in young children. In J. Moore & K. Stenning (Eds.), *Proceedings of the XXIII Annual Conference of the Cognitive Science Society*, 946-951. Mahwah, NJ: Erlbaum.

Sloutsky, V. M., & Lo, Y. (1999). How much does a shared name make things similar? Part 1: Linguistic labels and the development of similarity judgment. *Developmental Psychology*, *6*, 1478-1492.

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. C. (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.