Title
Auditory Stimuli Slow Down Responses and First Fixations:
Support for Auditory Dominance in Adults

Permalink
https://escholarship.org/uc/item/0nc816wd

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 37(0)

Authors
Robinson, Christopher W
Barnhart, Wesley R
Rivera, Samuel

Publication Date
2015

Peer reviewed

Auditory Stimuli Slow Down Responses and First Fixations: Support for Auditory Dominance in Adults

Christopher W. Robinson (robinson.777@osu.edu)
Department of Psychology, The Ohio State University Newark
1179 University Dr., Newark, OH 43055, USA

Wesley R. Barnhart (barnhart.135@osu.edu)
Department of Psychology, The Ohio State University Newark
1179 University Dr., Newark, OH 43055, USA

Samuel Rivera (rivera.162@osu.edu)
Department of Psychology, The Ohio State University
1835 Neil Avenue, Columbus, OH 43210, USA

Abstract

Under some situations sensory modalities compete for attention, with one modality attenuating processing in a second modality. Almost forty years of research with adults has shown that this competition is typically won by the visual modality. Using a discrimination task on an eye tracker, the current research provides novel support for auditory dominance, with words and nonlinguistic sounds slowing down visual processing. At the same time, there was no evidence suggesting that visual input slowed down auditory processing. Several eye tracking variables correlated with behavioral responses. Of particular interest is the finding that adults' first fixations were delayed when images were paired with auditory input, especially nonlinguistic sounds. This finding is consistent with neurophysiological findings and also consistent with a potential mechanism underlying auditory dominance effects.

Keywords: Sensory Dominance, Cross-modal Processing, Attention

Introduction

Most of our experiences are multisensory in nature; however, historically most research has examined processing in a single sensory modality. Over the last forty years there has been a growing body of research examining how sensory modalities interact while processing multisensory information (e.g., sounds and pictures paired together). Under some conditions, presenting information to multiple sensory modalities facilitates processing; whereas, under other conditions, multisensory presentation hinders processing. For example, amodal information such as rate, tempo, etc., can be expressed in multiple sensory modalities (e.g., rate of a hammer tapping can be seen and heard). When processing amodal information, multisensory presentation often speeds up responses and facilitates learning (Fort, Delpuech, Pernier, & Giard, 2002; Giard & Peronnet, 1999; see also Bahrick, Lickliter, & Flom, 2004 for a review).

However, there are many situations where the additional sensory information is arbitrary in nature. For example, tasks such as driving (visual) and talking on the phone (auditory) require people to divide their attention across sensory modalities. Furthermore, due to the arbitrary nature of the input, stimuli in one modality provide little to no details about information presented to another modality (e.g., a phone conversation provides no information about upcoming traffic lights, location of pedestrians, etc.). Research examining processing of arbitrary information often shows that stimuli presented to one modality often interferes with processing in a second modality (see Robinson & Sloutsky, 2010a; Sinnett, Spence, & Soto-Faraco, 2007; Spence, 2009 for reviews). The current paper is primarily interested in these cross-modal interference effects, or modality dominance effects.

There is a clear pattern within the adult literature. When presented with arbitrary, auditory and visual information, visual input often wins the competition (Colavita, 1974; Posner, Nissen, & Klein, 1976; Sinnett, Spence, & Soto-Faraco, 2007). For example, in a typical Colavita task, participants are instructed to press one button when they see a light and press a different button when they hear a tone (Colavita, 1974). The majority of trials are unimodal (only light or sound); however, some trials are cross-modal (light and sound are paired together). On these cross-modal trials, participants often respond incorrectly by only pressing the visual button, as opposed to pressing both buttons or a third button associated with cross-modal stimuli. Over the last forty years visual dominance has been extended to different tasks with a variety of attentional manipulations failing to reverse the visual dominance effect (Ngo, Sinnett, Soto-Faraco, & Spence, 2010; see also Spence, 2009 for a review).

Interestingly, a different pattern can be found in the developmental literature, with auditory information often attenuating processing of visual input (Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003). For example, after familiarizing or habituating infants to auditory-visual pairings, infants increase looking when the auditory component changes at test ($AUD_{new}VIS_{old}$) but often fail to increase looking when only the visual component changes at test ($AUD_{old}VIS_{new}$). This finding is noteworthy because infants discriminate the same visual images when presented in silence; therefore, it was concluded that the auditory input overshadowed or attenuated processing of the visual input. Auditory dominance effects are not limited to infants. When presented with two auditory-visual pairings in a matching game, four-year olds often report that the two pairs are the same when only the visual component changes ($AUD_1VIS_1 \rightarrow AUD_1VIS_2$). In contrast, adults correctly report that the two pairs are different (Napolitano & Sloutsky, 2004; Sloutsky & Napolitano, 2003).

These findings led researchers to posit that auditory input overshadows visual input early in development (Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003). According to this account, sensory modalities share the same pool of attentional resources and compete for these resources (see Robinson & Sloutsky, 2010a for a review). Because auditory stimuli are dynamic and transient in nature and are processed faster than visual input (Green & von Gierke, 1984), attention may automatically be directed to this information. Furthermore, due to competition for resources, processing details of a visual stimulus may not start until the auditory modality releases attention. While this account has received some support in the developmental literature (Robinson & Sloutsky, 2007; 2010b; Sloutsky & Robinson, 2008), there is little support for auditory dominance in adult populations.

How do modality dominance effects change across development? Increased resource capacity and faster processing speed in adults (c.f., Kail & Salthouse, 1994) can explain why under the same stimulus presentation times children only process information in one modality; whereas, adults have ample time to process stimuli in both modalities. However, it is unclear how to reconcile the auditory dominance account with a reversal to visual dominance. One possibility is that the mechanism basically remains unchanged; however, across development, visual stimuli become more salient, automatically engage attention, and attenuate encoding of auditory input. For example, it is well established that the auditory modality develops before the visual modality, with hearing beginning in the third trimester of pregnancy and vision being relatively poor for the first few months of life. It is possible that it might take several years for the visual modality to “catch up” to the auditory modality. Alternatively, it is possible that visual input is less likely to engage attention than auditory input, and adults strategically bias their responses in favor of visual input to compensate for the poor alerting abilities of this class of stimuli (Posner, Nissen, & Klein, 1976). In other words, visual dominance may reflect a response bias rather than visual input attenuating encoding of auditory input (Spence, 2009). Following up on this idea, it is possible that auditory dominance in children (auditory input disrupting visual encoding) and visual dominance in adults (visual response bias) co-exist (Chandra, Robinson, & Sinnett, 2011) and are driven by different mechanisms, with many studies overlooking auditory dominance because adults strategically bias their responses in favor of visual input. The goal of the current study is to test the hypothesis that auditory dominance is still present in adult populations and to test assumptions underlying auditory dominance.

Adults in the current study participated in immediate recognition tasks where they had to determine if two auditory stimuli, two visual stimuli, or two AV pairs were identical or different. In contrast to previous research (Napolitano & Sloutsky, 2004; Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003), images were presented on an eye tracker so we could examine patterns of fixations while participants were discriminating images. Second, rather than examining accuracies, the current study compared response times in the unimodal and cross-modal conditions. Based on previous research and on a proposed mechanism underlying auditory dominance (Robinson, Ahmar, & Sloutsky, 2010; Robinson & Sloutsky, 2010a), it was hypothesized that pairing the pictures with words (Experiment 1) or sounds (Experiment 2) would slow down processing of the visual stimulus and have no negative effect on auditory processing. Furthermore, it was expected that eye tracking variables such as latency of first fixation and mean fixation durations may also account for slower response times in cross-modal conditions.

Experiment 1

Method

Participants Thirty-eight undergraduate students (*M* = 19.52 years, 20 Females) who were enrolled in an Introductory Psychology course at The Ohio State University at Newark participated in this experiment. Completion of the study granted participants with credit that served to fulfill a course requirement. All participants provided informed consent, had normal hearing and vision (self-reported), and were debriefed after completion of the study.

Apparatus Participants were centrally positioned and seated approximately 60 cm in front of an Eye Link 1000 Plus eye tracker with desktop mount and remote camera. The eye tracker computed eye movements at a rate of 500 Hz, and Experiment Builder 1.10.165 controlled the timing of stimulus presentations. Visual stimuli were presented on a BenQ XL2420 24” monitor and auditory stimuli were presented via Kensington 33137 headphones. Eye tracking data were collected and stored on a Dell Optiplex 7010 computer. Gaze fixation positions and durations were identified by the Eye Link system online during the experiment and recorded for offline analysis. The eye tracker, stimulus presentation computer, and eye tracking computer were stationed in a quiet testing room in the High-Tech lab at The Ohio State University at Newark. A trained experimenter oversaw the entire duration of each participant’s study and they manually started each trial when the participants fixated on a central stimulus.

Stimuli Visual stimuli consisted of four pairs of images which were digitally constructed in Microsoft PowerPoint and exported as 600 x 600 bmp files (approximate size), see Figure 1 for examples of visual stimuli and Areas Of Interest (AOI). The stimuli resembled the following real-world objects: cone of cotton candy, tree, globe, and rabbit, and each stimulus pair differed by two or four features. For example, as can be seen in Figure 1, the diamond and circle could be used to differentiate the two trees; thus, these two features/AOIs were considered to be *relevant*. The heart and star were considered *irrelevant* because both trees shared these two features and therefore cannot be used to differentiate the trees. Within an individual trial, one of the items from the pair (i.e., Target) was presented for 1 s, with a 1 s Inter-Stimulus Interval (ISI). The second item (Test) was presented until the participant made a response. Each item in the pair was equally likely to be the Target or Test item.

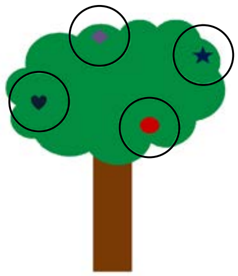

Figure 1: Example of two visual pairs used in Experiments 1 and 2. The circles around each feature denote the AOIs and were not visible during the actual experiment.

As with visual stimuli, auditory stimuli consisted of four word pairs. The auditory stimuli used were one-syllable nonsense words (e.g., paf vs. dax and ket vs. yun) and two-syllable nonsense words (e.g., lapo vs. vika and kuna vs. whonae). Each word was individually spoken by a female experimenter and recorded using Cool Edit 2000. Audio files were saved as 44.1 kHz wav files and presented to participants via headphones at approximately 65-68 dB. Each item in the pair was equally likely to be the Target or the Test item. Stimuli in the cross-modal condition were created by presenting images and words at the same time.

Design Each participant completed three conditions: Unimodal Auditory (UA), Unimodal Visual (UV), and Cross-Modal (CM) conditions. In the UA and UV conditions, participants were either presented with two words or two images, respectively, and they had to determine if the stimuli were exactly the same or different. In the CM condition they had to discriminate the same words and pictures; however, the auditory and visual information were presented at the same time. Discrimination in the cross-modal condition was compared to respective baselines. Visual dominance would be inferred if cross-modal presentation only slows down auditory processing (compared to UA baseline), and auditory dominance would be inferred if cross-modal presentation only slows down visual processing (compared to UV baseline). Increased response times in both modalities in the cross-modal condition would suggest increased task demands with no evidence that one modality dominated the other modality.

Procedure Participants were positioned to face the eye tracker centrally with an approximate viewing distance of 60 cm. At the right side of each participant was the experimenter; s/he began the experiment by calibrating participants' eye measurements, a process that included a 9-point sequence of fixations, which was followed by a 9-point validation. The initial calibration/validation process lasted approximately 1-5 minutes. After calibration, participants were presented with a screen that discussed the experimental instructions. In the unimodal auditory and visual conditions they were told that they would hear two words or see two pictures and they had to press 1 if the stimuli were exactly the same and press 3 if they were different. They were also told to respond as quickly and as accurately as possible. There were 60 trials in each condition, half same trials and half different trials, and each trial began with drift correction (i.e., central fixation stimulus). In the cross-modal condition, participants were told that they would see two picture-word pairs and they were instructed to press 1 if both the pictures and words were exactly the same ($Aud_1Vis_1 \rightarrow Aud_1Vis_1$). They were told to press 3 if the word changed ($Aud_1Vis_1 \rightarrow Aud_2Vis_1$), the picture changed ($Aud_1Vis_1 \rightarrow Aud_1Vis_2$), or if both components changed ($Aud_1Vis_1 \rightarrow Aud_2Vis_2$). There were 60 trials in the cross-modal condition, 15 of each of the trial types listed above, and each trial began with drift correction. Order of condition (auditory, visual, and cross-modal) was randomized for each participant, and as in the unimodal conditions, they were instructed to respond quickly and accurately.

Results and Discussion

Overall, participants exhibited high accuracy throughout the procedure (M = .96, SD = .19); therefore, primary analyses focused on participants' response times on correct trials. In particular, we were primarily interested in how cross-modal presentation affected auditory and visual processing, so we focused on two comparisons. To quantify effects of visual input on auditory processing we compared how quickly participants discriminated words in the cross-modal condition ($Aud_1Vis_1 \rightarrow Aud_2Vis_1$) with discrimination of the same words in the unimodal condition ($Aud_1 \rightarrow Aud_2$). To quantify effects of auditory input on visual processing we compared how quickly participants discriminated visual images in the cross-modal condition ($Aud_1Vis_1 \rightarrow Aud_1Vis_2$) with discrimination of the same images in the unimodal condition ($Vis_1 \rightarrow Vis_2$). The Means and Standard Errors are presented in Figure 2. Log transformed means were submitted to a 2 Modality (Auditory vs. Visual) x 2 Presentation (Unimodal vs. Cross-modal) ANOVA with both factors manipulated within subjects. The ANOVA revealed a main effect of Modality, F (1,37) = 76.89, p < .001, a main effect of Presentation, F (1,37) = 13.97, p < .001, and the predicted Modality x Presentation interaction was also significant, F (1,37) = 10.47, p < .005. How does cross-modal presentation affect processing of visual and auditory input? Paired t-tests with a Bonferonni adjustment (p < .025) showed slower visual response times in the cross-modal condition than in the unimodal condition, t (37) = 4.74, p < .001. The slowdown in the auditory modality was less pronounced, as indicated by the Modality x Presentation interaction, and did not reach significance when adjusting for multiple comparisons, t (37) = 2.20, p = .034.

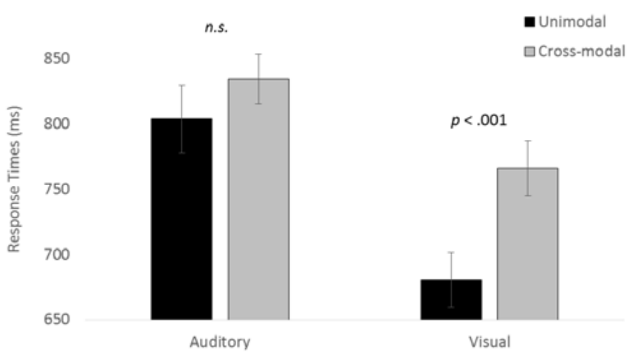

Figure 2: Mean response times and Standard Errors in Experiment 1.

The same visual pairs were used across the whole experiment; thus, it is possible that adults eventually learned to pay attention to the relevant features. However, note that the auditory dominance account (Robinson & Sloutsky, 2010a) argues that auditory input automatically engages attention; therefore, knowledge of the relevant visual features and top-down attentional control should have little effect on how attention is automatically deployed to cross-modal stimuli. To examine if participants could override auditory dominance we focused on visual discrimination in the last half of the cross-modal condition (Trials 31-60). Participants' log transformed visual response times in the cross-modal condition were faster in the last 30 trials compared to the first 30 trials, t (37) = 4.04, $p < .001$, suggesting that some learning occurred. However, despite this learning, the auditory stimuli continued to slow down responses to visual stimuli, t (37) = 3.26, $p < .005$; whereas, visual input had no negative effect on auditory processing in the last half of the study, t (37) = 1.44, p = .16.

According to the proposed mechanism underlying auditory dominance (Robinson & Sloutsky, 2010a), auditory input should slow down or delay the onset of visual processing. Preliminary support for this hypothesis comes from a passive ERP oddball procedure where cross-modal presentation sped up auditory P300s and slowed down visual P300s (Robinson, Ahmar, & Sloutsky, 2010). To further examine this proposal, we directly compared patterns of fixations while participants were discriminating visual stimuli in the unimodal and cross-modal conditions. More specifically, we focused on variables that could potentially account for this slowdown. For example, given increased latency of visual P300 (Robinson, Ahmar, & Sloutsky, 2010), it is possible that latency of first fixation and/or latency of first fixation to a relevant AOI could be delayed. If learning of visual input is disrupted, it is possible there will be relatively less looking to relevant AOIs. We also examined mean fixation times with the assumption that disrupting visual processing would result in longer individual fixations. Finally, given short presentation times, increased fixation durations should be associated with fewer fixations. Latency of fixations, fixation durations, and number of fixations were derived offline from fixations identified by the Eye Link system with custom MATLAB and Python software developed by the third author. Fixations initiated before the stimulus presentation or after responses were excluded. Latencies were defined as the fixation start time relative to the visual stimulus onset time. Relevant fixations were those that occurred within either of the relevant AOIs, as depicted in Figure 1. As can be seen in the Table 1, latencies (delayed) and fixation durations (longer) were in the predicted direction; however, these effects did not reach significance when using a Bonferonni adjustment (p <.01).

Table 1: Means, (Standard Errors), Paired t's, and p's across the unimodal and cross-modal conditions in Experiment 1.

Dependent Measure	Unimodal Condition (SE)	Cross-modal Condition (SE)	Paired t (df)	p value
Mean Fixation Duration	272 (12)	296 (15)	1.56 (37)	.121
Proportion Looking to Relevant AOIs	.12 (.01)	.10 (.01)	1.24 (37)	.223
Number of Fixations	1.20 (.11)	1.21 (.11)	0.13 (37)	.899
Latency of First Look	332 (11)	374 (20)	2.22 (37)	.033
Latency of First Look to Relevant AOI	497 (68)	573 (70)	0.89 (28)	.382

We also looked at correlations between eye tracking variables and costs of auditory input on individual response times. To quantify the cost of auditory input on visual processing we calculated a difference score for each participant (Log transformed RT in cross-modal condition minus Log transformed RT in unimodal condition). Values greater than zero suggest that the words slowed down visual processing and values less than zero indicate that the words sped up response times. We then looked at the correlations between the eye tracking variables reported in Table 1 with this difference score.

Table 2: Correlations between eye tracking variables and Difference score (Diff = RTs for discriminating visual stimuli in cross-modal condition minus RTs in unimodal condition). Note: "*" $p < .05$, "**" $p < .01$.

	Diff	Fix Dur.	% Rel.	Fix #	Lat.	Lat. Rel.
Diff	---	-.34*	-.18	-.44**	.23	-.17
Fix Dur.		---	.05	.05	-.09	.08
% Rel.			---	.08	-.10	.09
# of Fix				---	-.36*	-.45**
Lat. Dur.					---	.18
Lat. Rel.						---

As can be seen in Table 2, the number of fixations was negatively correlated with the difference score, suggesting that adults who responded more slowly to changes in visual images made fewer fixations.

The behavioral findings from Experiment 1 are consistent with auditory dominance, with cross-modal presentation being more likely to slow down visual processing than auditory processing. Is it possible that the effects are specific to human speech, a familiar class of stimuli for adults? To address this issue, we replaced the words with nonlinguistic sounds in Experiment 2.

Experiment 2

Method

Participants, Stimuli, and Procedure Twenty-nine undergraduate students (*M* = 20.15 years, 21 Females) participated in this study in exchange for course credit. The visual stimuli and the procedure were similar to Experiment 1; however, the images in the current experiment were paired with non-linguistic sounds. Four pairs of sounds were created using Audacity software (e.g., tones differing by 200 Hz). As in Experiment 1, the nonlinguistic sounds were one second in duration and the timing and duration of auditory, visual, and cross-modal stimuli were identical to Experiment 1. In contrast to Experiment 1, we focused exclusively on the visual and cross-modal conditions because there were no images to look at in the unimodal auditory condition. Thus, eye tracking data from the auditory condition would have provided no eye tracking information. The nonlinguistic sounds and images used in Experiment 2 have been tested without an eye tracker and cross-modal presentation slowed down visual processing and had no negative effect on auditory processing (Dunifon & Robinson, 2015).

Results and Discussion

As in Experiment 1, we compared how quickly participants discriminated two images when presented in silence with discrimination of the same two images when paired with the same sound. As in Experiment 1, participants were slower at discriminating the images in the cross-modal condition (*M* = 871 ms) than in the unimodal condition (*M* = 752 ms), paired *t* test with log transformed RTs, $t\,(28) = 4.44$, $p < .001$.

We also examined patterns of fixations while participants discriminated pictures in the unimodal and cross-modal conditions. See Table 3 for statistics. Adults in the cross-modal condition were slower to make their first fixations, mean fixation durations were longer, and latency of first look to relevant AOIs were also delayed compared to the unimodal baseline.

Do patterns of fixations predict which adults were most affected by the auditory stimulus? To address this issue we calculated a difference score (Log transformed RT in cross-modal condition minus log transformed RT in unimodal condition) and examined how eye tracking variables correlated with this difference score. See Table 4 for statistics. As can be seen in the Table 4, individuals who were slower at making visual responses in the cross-modal condition made more frequent and longer fixations and were slower to initially fixate on the relevant AOIs; however, these effects were only marginally significant when adjusting for multiple comparisons.

Table 3: Means, (Standard Errors), Paired *t*'s, and *p*'s across the unimodal and cross-modal conditions in Experiment 2. Note: "*" denotes that $p < .015$.

Dependent Measure	Unimodal Condition (SE)	Cross-modal Condition (SE)	Paired *t* (df)	*p* value
Mean Fixation Duration*	296 (13)	347 (24)	2.63 (28)	.014
Proportion Looking to Relevant AOIs	.15 (.02)	.14 (.02)	-0.50 (28)	.619
Number of Fixations	1.10 (.11)	1.07 (.11)	-0.27 (28)	.792
Latency of First Look*	334 (17)	409 (30)	3.13 (28)	.004
Latency of First Look to Relevant AOI*	479 (35)	624 (54)	2.72 (17)	.015

Table 4: Correlations between eye tracking variables and Difference score (Diff = RTs for discriminating visual stimuli in cross-modal condition minus RTs in unimodal condition). Note: "*" $p < .05$, "**" $p < .01$.

	Diff	Fix Dur.	% Rel.	Fix #	Lat.	Lat. Rel.
Diff	---	.41*	.25	.36*	-.04	.49*
Fix Dur.		---	-.21	-.19	-.10	.55**
% Rel.			---	.33	.26	-.12
Fix #				---	-.24	-.17
Lat.					---	.43
Lat. Rel.						---

Discussion

The current study examined how quickly adults could discriminate two pictures that were presented in silence or paired with words or sounds. While the adult literature consistently points to visual dominance (see Sinnett, Spence, & Soto-Faraco, 2007; Spence, 2009 for reviews), the current study found novel evidence that words and sounds both slowed down visual discriminations. At the same time, under similar testing conditions, the visual images did not slow down auditory discrimination (Experiment 1; Dunifon & Robinson, 2015). This asymmetric cost in adults is a novel finding and is consistent with auditory dominance effects reported in the developmental literature (Sloutsky & Napolitano, 2003; Sloutsky & Robinson, 2008).

Nonsense words and nonlinguistic sounds both slowed down behavioral responses; however, eye tracking variables were more predictive in the nonlinguistic sound experiment. In particular, adults who saw images paired with nonlinguistic sounds were slower to make their first fixation, slower to make their first fixation to a relevant AOI, and fixated for longer durations than in the unimodal condition. These findings are consistent with neurophysiological findings where auditory input delayed visual P300s (Robinson, Ahmar, & Sloutsky, 2010) and are consistent with the claim that auditory input slows down or delays visual encoding. Furthermore, consistent with previous research, auditory interference effects are often more pronounced when using unfamiliar auditory stimuli than when using familiar stimuli or a familiar class of auditory stimuli such as human speech (Robinson & Sloutsky, 2010b; Sloutsky & Robinson, 2008, but see Napolitano & Sloutsky, 2004). The underlying idea is that novel auditory stimuli consume more attentional resources, which results in a greater cost on visual processing.

The current study provides support for auditory dominance in adult populations, but two issues need to be addressed in future research. First, slower visual response times in Experiment 1 were associated with fewer fixations, whereas slower responses in Experiment 2 were associated with more fixations (see Tables 2 and 4). One possible explanation is that adults treat words differently than other sounds/features (c.f., Yamauchi & Markman, 2000), and the sounds and words had different effects on visual attention. Second, why was auditory dominance found in this study while other studies show visual dominance effects? We believe one key factor is that we eliminated a potential mechanism underlying visual dominance (i.e., response bias). In contrast to most of the adult studies, auditory and visual discrimination was assessed by making the same response; thus, participants could not bias their response in favor of visual input. Furthermore, using a similar change detection task but requiring separate responses for auditory and visual discrimination resulted in visual dominance (Chandra, Robinson, & Sinnett, 2011).

While much of the adult literature suggests that visual input dominates auditory processing, the current study provides novel support for auditory dominance, with words and sounds slowing down responding to visual input. Furthermore, sounds also delayed the latency of first fixations and increased fixation durations. These findings have implications on a variety of tasks that hinge on the processing of multisensory input.

References

Bahrick, L. E., Lickliter, R., & Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. *Current Directions in Psychological Science*, *13*, 99-102.

Chandra., M., Robinson, C. W., & Sinnett, S. (2011). Coexistence of multiple modal dominances. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2604-2609). Austin, TX: Cognitive Science Society.

Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics*, *16*, 409-412.

Dunifon, C. & Robinson, C. W. (2015, April). *Pay attention to the pictures: Auditory dominance not under attentional control.* Poster presented at the Annual meeting of the Midwestern Psychological Association, Chicago, IL.

Fort, A., Delpuech, C., Pernier, J., & Giard, M. H. (2002). Dynamics of cortico-subcortical cross-modal operations involved in audio–visual object recognition in humans. *Cerebral Cortex*, *12*, 1031–1039.

Giard, M. H., & Peronnet, F. (1999). Auditory–visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, *11*, 473–490.

Green, D. M., & von Gierke, S. M. (1984). Visual and auditory choice reaction times. *Acta Psychologica*, *55*, 231 – 247.

Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica*, *86*, 199 – 225.

Napolitano, A. C., & Sloutsky, V. M. (2004). Is a picture worth a thousand words? The flexible nature of modality dominance in young children. *Child Development*, *75*, 1850-1870.

Ngo, M. K., Sinnett, S., Soto-Faraco, S., & Spence, C. (2010). Repetition blindness and the Colavita effect. *Neuroscience Letters*, *480*, 186–190.

Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review*, *83*, 157-171.

Robinson, C. W., Ahmar, N., & Sloutsky, V. M. (2010). Evidence for auditory dominance in a passive oddball task. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp 2644-2649). Austin, TX: Cognitive Science Society.

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development*, *75*, 1387-1401.

Robinson, C. W., & Sloutsky, V. M. (2007). Visual processing speed: Effects of auditory input on visual processing. *Developmental Science*, *10*, 734-740.

Robinson, C. W., & Sloutsky, V. M. (2010a). Development of Cross-modal Processing. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*, 135-141.

Robinson, C. W., & Sloutsky, V. M. (2010b). Effects of multimodal presentation and stimulus familiarity on auditory and visual processing. *Journal of Experimental Child Psychology*, *107*, 351-358.

Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). Visual dominance and attention: Revisiting the Colavita effect. *Perception & Psychophysics*, *69*, 673–686.

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.

Sloutsky, V. M., & Robinson, C. W. (2008). The role of words and sounds in visual processing: From overshadowing to attentional tuning. *Cognitive Science*, *32*, 354-377.

Spence, C. (2009). Explaining the Colavita visual dominance effect, *Progress in Brain Research*, *176*, 245–258.

Yamauchi, T., & Markman, A. B. (2000). Inference using categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 776-795