

Auditory Overshadowing in Preschoolers: A Preference for the Input, the System, or Both?

Allison O'Leary (oleary.83@osu.edu)
Department of Psychology, 1835 Neil Avenue
Columbus, OH 43210, USA

Vladimir M. Sloutsky (sloutsky.1@osu.edu)
Department of Psychology, 1835 Neil Avenue
Columbus, OH 43210, USA

Abstract

Auditory overshadowing occurs when the presence of an auditory stimulus interferes with visual processing. The current study tested whether this occurs due to a privileged attentional status of auditory input or due to the dynamic characteristics of auditory input. To address these questions, preschoolers completed one of four discrimination tasks. In the sound, motion, and item baseline conditions, children discriminated these single information types by judging whether paired stimuli were the same or different. In the combined condition, children discriminated changing sounds, motions, or items in the face of competing input in the other two dimensions. Although children's discrimination of all information types attenuated in the combined condition relative to baseline, motion and item discrimination attenuated more than auditory discrimination. This provides evidence that early in development auditory information receives privileged processing in the face of competing input.

Keywords: Attention; Cross-modal processing; Cognitive development; Preschoolers

Introduction

To successfully navigate and form an understanding of our environment, we must develop the ability to efficiently integrate information from multiple sensory modalities. There is evidence that some multisensory integration occurs even in neonates, as demonstrated by the structural convergence of input from different sensory modalities (Stein & Meredith, 1993). However, it is clear that the ability to integrate multimodal information is not fully mature at birth and, in fact, exhibits a protracted developmental trajectory. For example, in contrast to normally functioning adults, infants and young children exhibit a phenomenon known as auditory overshadowing (Robinson & Sloutsky, 2004b). This occurs when the presence of an auditory stimulus interferes with one's processing of a visual stimulus. That is, infants and young children more easily discriminate visual stimuli when presented in isolation than when paired with labels or sounds. In some cases, children completely fail to discriminate changes in a visual stimulus when presented simultaneously with a sound (Napolitano & Sloutsky, 2004). Importantly, although visual discrimination is impaired in the presence of auditory information, auditory discrimination does not suffer in the presence of visual

information (Sloutsky & Napolitano, 2003; Robinson & Sloutsky, 2010b).

The importance of audio-visual integration is especially apparent for processes like word learning, where individuals must map verbal labels to objects in visual space. If visual processing is inhibited when an auditory stimulus is present, auditory overshadowing could be a major contributor to the difficulty children face mapping words to objects. Indeed, 10-month-olds encoded only auditory information when presented with visual information and a verbally presented label. However, 16-month-olds demonstrated able processing of both the visual information and the label (as indicated by looking time preferences; Robinson & Sloutsky, 2004a).

Although the ability to process auditory and visual information progresses through infancy, even preschool-aged children show difficulties discriminating static visual stimuli in the presence of sounds and labels. Napolitano and Sloutsky (2004) demonstrated that 4-year-olds were susceptible to auditory overshadowing when asked to discriminate changes in the visual and auditory aspects of a target stimulus. Specifically, this effect was most profound when the visual and auditory stimuli were unfamiliar. Further, the authors demonstrated that this overshadowing was resistant to explicit instruction, in that it still occurred when children were asked to attend exclusively to visual information.

More recent research has aimed to detail the conditions in which auditory overshadowing occurs and to elucidate the basic cognitive mechanisms underlying these effects. Input from the visual and auditory modalities seem to "compete" for processing resources early in development, whereas adults are able to process multimodal information (Robinson & Sloutsky, 2004a). Robinson and Sloutsky (2007a) also argued that overshadowing may occur during initial processing/encoding as well as during response selection. In addition, unfamiliar auditory stimuli slow visual processing, whereas familiar auditory stimuli do not (Robinson & Sloutsky, 2007b). Despite our understanding of the subtleties of the effect, major questions remain regarding the privileged status of auditory input and a number of hypotheses explaining auditory overshadowing effects have been proposed.

One possibility is that auditory and visual information are processed serially, and that auditory stimuli are processed more thoroughly because they are faster to engage attention. Visual processing may be inhibited until attention is disengaged (Robinson & Sloutsky, 2010a). However, there are at least two possible explanations as to why auditory input would more easily engage attention. First, adults process auditory information more quickly than visual information, which may reflect privileged processing of auditory input per se (Green & von Gierke, 1984). Because the auditory system begins maturing before the visual system, it is possible the auditory system processes information more efficiently even during childhood. Alternatively, it is possible that the nature of the input plays a role. For example, auditory stimuli are typically dynamic (i.e., exhibit change over time) and have more abrupt onset than visual stimuli. Perhaps it is this dynamicity that quickly engages attention.

However, previous research could not distinguish between these possibilities because all previous studies of auditory overshadowing have utilized static visual stimuli. The current study aimed to test these explanations by increasing the dynamicity of visual stimuli. This was done by adding motion to the visual stimulus, which is a powerful bottom-up attentional cue (Egeth & Yantis, 1997). If the dynamic nature of auditory input is responsible for auditory overshadowing, one would expect the increased dynamicity of the visual stimulus to attenuate auditory overshadowing, resulting in more thorough visual processing. In this situation, visual information may even overshadow auditory information. On the other hand, if the auditory system itself is privileged, one would not expect such attenuation, even if the visual cue is dynamic. Distinguishing between these possibilities would result in better understanding of the mechanisms underlying auditory overshadowing.

In previous studies where young children were presented with auditory and visual information, visual discrimination attenuated significantly whereas auditory discrimination showed no significant attenuation. The current study aimed to investigate the effect when presented stimuli had an auditory, visual, and motion component. In this “combined” condition, target and test stimuli included all of these components. Children were instructed to say “same” if the target and test stimuli were the same in all 3 aspects. Children were instructed to say “different,” however, if the sound component changed, the motion component changed, or the item appearance component changed. To determine the extent of overshadowing, we assessed children’s discrimination of changes in the sound, image, or motion of these more complex stimuli relative to their ability to discriminate these types of information when presented in isolation (i.e., sound, motion, and item discrimination baselines). If auditory overshadowing stems from the dynamic nature of sound, then the dynamic motion cue should attenuate auditory processing.

Method

Participants

Eighty-two four-year-olds (41 girls and 41 boys, $M = 4.48$ years, $SD = .28$ years) participated in this experiment. Children completed one of four conditions: sound baseline ($N = 20$), motion baseline ($N = 20$), item baseline ($N = 20$), or a combined (sound-motion-item) condition ($N = 22$). Children were recruited through local daycares and preschools located in Columbus, Ohio. The majority of participants were Caucasian.

Stimuli

In the combined condition, children saw two moving “toys” which were each paired with a sound, and were asked to discriminate changes in the sound, motion, or appearance of these toys. Each stimulus in this condition consisted of a sound, motion, and an object, referred to as “item”. There were 8 sound, 8 motion, and 8 item appearance possibilities. All of these were intended to be novel to children, as overshadowing effects are sensitive to stimulus familiarity (Robinson & Sloutsky, 2010b). For example, auditory stimuli consisted of dynamic sounds like camera clicks and notes from an organ. Each auditory stimulus was 1500ms in duration. The 8 items all consisted of a central ‘X’ on which four colored shapes were placed. Each of these items was animated in 8 different motion patterns (e.g., 360° rotation, looming) to produce 512 total sound-motion-item combinations. Each stimulus presentation proceeded in the following way: children viewed the static image in silence for 500ms, after which the motion and sound began simultaneously, lasting for the remaining 1500ms of stimulus presentation.

Each of the baseline conditions aimed to assess children’s discrimination when a single type of information is presented. In the sound baseline, children were asked to discriminate 2 of the 8 sounds used in the combined condition. In this baseline, the only visual stimulus was a small fixation cross presented in the center of the screen. In the motion baseline, children were asked to discriminate 2 of the 8 motion types. Here, the same item was used throughout the task, and these visual stimuli were presented in silence. In the item baseline, children discriminated 2 of the 8 items, presented statically and in silence.

Procedure

In all conditions, children observed pairs of stimuli presented sequentially and were asked to indicate whether the two stimuli were the same or different. Because of their novelty, the experimenter described the stimuli as “toys from outer space.” Children were told they would see two of these toys and that their job was to tell the experimenter whether the two toys were different in any way (i.e., if they noticed a difference in the item, motion, or sound). During each trial, children were first presented with a fixation cross in the center of the screen. Once the child fixated on this

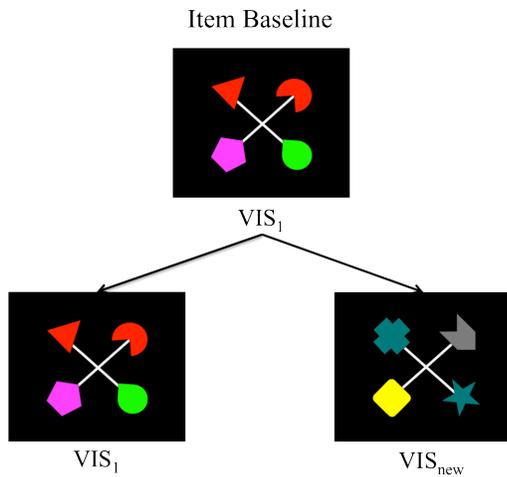
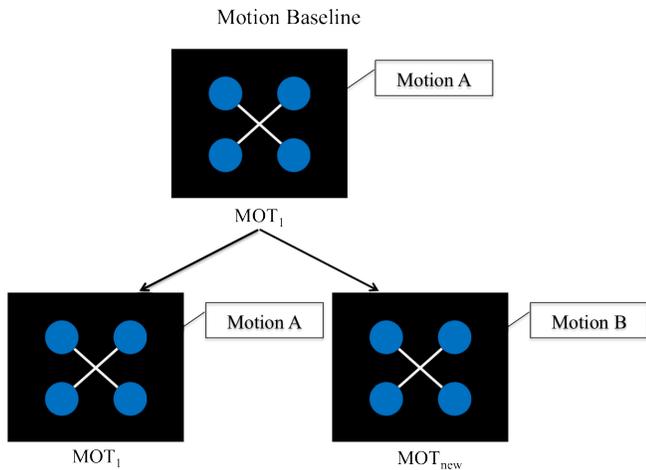
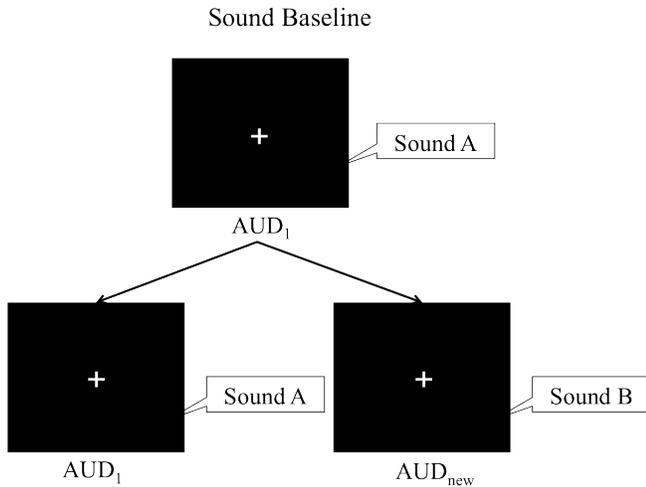


Figure 1. Target and test stimuli used in the sound, motion, and item baseline conditions. The top images depict example target stimuli. Bottom images depict potential test stimuli for that target.

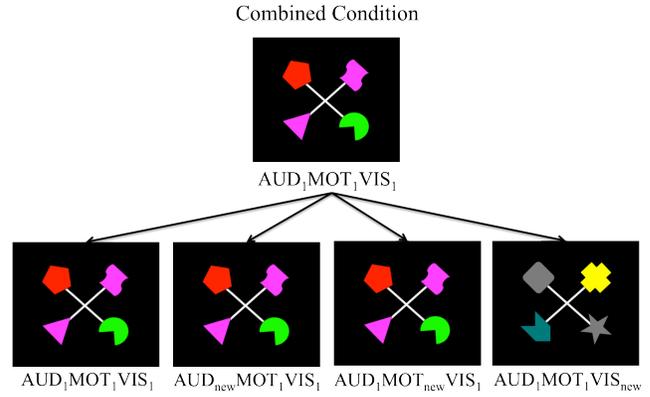


Figure 2. Target and test stimuli used in the combined condition.

cross, the experimenter pressed the spacebar to present the target stimulus. The child then saw another fixation cross.

Upon the child's fixation, the experimenter pressed the spacebar to present the test stimulus. Each target and test stimulus was 4000ms in duration. Following test stimulus presentation, the child was asked to indicate verbally whether the two toys they just observed were the same or different. Half of the trials were "same" trials, in that the target stimulus was identical to the test stimulus. The remaining half of trials were "different" trials, such that the item, motion, or sound presented at test differed from that in the target. Each child completed 4 blocks which each consisted of 12 trials. Children saw a 12-second cartoon between blocks.

In the combined condition, each target trial consisted of a moving toy paired with a sound. Each sound, motion, and item was selected randomly. During the "same" test trials, children were presented with a stimulus identical to the target stimulus. In one third of the "different" trials, only the sound changed at test, and the item and motion remained the same. In one third of these trials, only the motion of the item changed, and the item itself as well as the sound remained the same. In the remaining third of these "different" trials, the item changed, but the motion and sound remained the same. Of interest was whether children showed differential ability to discriminate these 3 types of information when presented simultaneously, relative to when presented in isolation (i.e., performance in the respective baselines).

The baseline conditions assessed children's ability to discriminate sound, motion, and item information individually. In the sound baseline, target trials consisted of an auditory stimulus (selected randomly from the 8 possibilities) presented with only a fixation cross in the center of the screen. During "same" trials, children heard the same sound. During "different" trials, children heard a different sound. The same structure applied to the motion and item baselines.

Results

To ensure that children understood and were engaged with the tasks, only those children whose overall performance was above chance (50% accuracy) were included in subsequent analyses. This eliminated 2 children in the sound baseline condition, 2 children in the motion baseline condition, and 2 children in the item baseline condition.

Baseline conditions

To evaluate the meaning of differential performance (discriminating sound, motion, and item) in the combined condition, it was necessary to first establish children's ability to discriminate each type of information presented independently. Children demonstrated high performance across the sound ($M = .91$, $SD = .08$), motion ($M = .87$, $SD = .10$), and item ($M = .92$, $SD = .07$) baselines. A one-way ANOVA indicated no significant differences between children's performance in these three conditions, $F(2, 51) = 2.03$, $p = .142$.

A one-way ANOVA comparing children's reaction times to "different" trials in the 3 conditions (including only correct responses) revealed a main effect of information type, $F(2, 51) = 8.77$, $p < .005$. Post hoc multiple comparisons using Fisher's LSD revealed that individuals in the item baseline condition ($M = 1486.81$, $SD = 1066.36$) discriminated more quickly than those in the motion ($M = 2409.69$, $SD = 817.29$) and sound ($M = 2653.24$, $SD = 724.60$) baselines. Average RTs in the motion and sound baselines did not differ significantly from one another. These results suggest that item discrimination was somewhat easier than discrimination of sounds or patterns of motion.

Combined condition

In the combined condition, children indeed showed differential accuracy discriminating sound, motion, and item changes, $F(2, 63) = 7.11$, $p < .005$. Post-hoc multiple comparisons indicated children were significantly more accurate when discriminating sounds ($M = .70$, $SD = .25$) than both motions ($M = .48$, $SD = .30$), and items ($M = .39$, $SD = .30$). Children's discrimination of motions and items did not differ significantly. Further, children's reaction times did not differ significantly when discriminating the different information types, $F(2, 63) = 1.45$, $p = .242$.

Comparing baseline and combined conditions

Because the assessment of overshadowing requires the comparison of performance on the between-subjects baseline conditions and the within-subjects combined condition, we calculated difference scores to describe each individual's attenuated performance in the combined condition relative to baseline. For example, we subtracted the mean accuracy of all individuals in the sound baseline from each individual's average accuracy to "different sound" trials in the combined condition.

As indexed by these difference scores, children's performance in the combined condition attenuated significantly across all the information types (i.e., all difference scores differed from zero): sound, $t(21) = -3.83$, $p < .005$; motion, $t(21) = -6.03$, $p < .001$; and item, $t(21) = -8.29$, $p < .001$.

A one-way ANOVA comparing children's difference scores revealed a main effect of condition, $F(2, 63) = 7.472$, $p < .005$. Post-hoc multiple comparisons using Fisher's LSD revealed children's performance in the combined condition attenuated less in the auditory condition ($M = -.20$, $SD = .25$) than in the item condition ($M = -.53$, $SD = .30$) as well as the motion condition ($M = -.39$, $SD = .30$). Difference scores in the item and motion conditions did not differ from one another. Children's reaction times differed amongst the baseline and combined conditions only when it came to discriminating item information, $t(38) = -2.53$, $p < .05$.

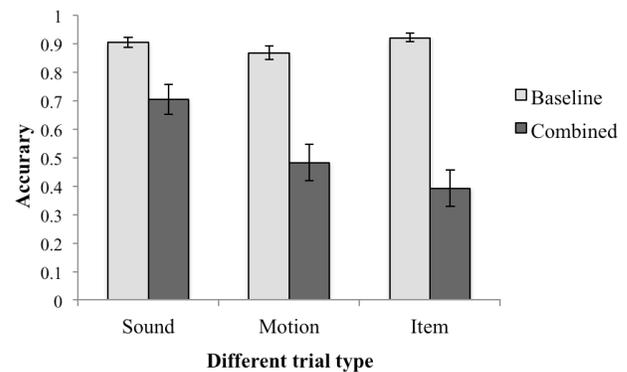


Figure 3. Proportion of accuracy to "different" trials for sound, motion, and item information by condition.

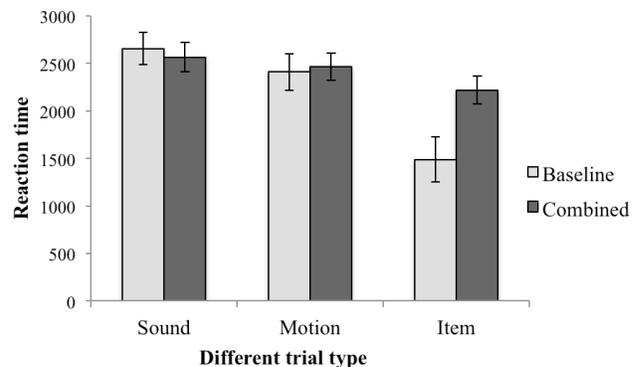


Figure 4. Average reaction time to discriminate "different" trials for sound, motion, and item information by condition. Baseline reaction times include only correct responses.

In sum, children's discrimination of sound, motion, and item information attenuated when presented simultaneously rather than in isolation. Sound discrimination attenuated the least, followed by motion discrimination, and then item discrimination, which attenuated most. These results indicate that auditory information was more likely to overshadow visual motion than the other way around.

These findings, in turn, suggest that processing of auditory information may indeed be privileged compared to visual information, even if visual information is dynamic in nature.

General Discussion

The current experiment compared children's ability to discriminate motion, sound, and item appearance information when presented simultaneously versus in isolation. When combined, children's discrimination of all 3 information types attenuated relative to baseline performance. This indicates that all 3 types of information were somewhat susceptible to interference from one another. It is well established that when auditory and static visual stimuli are presented simultaneously, processing of the visual stimulus suffers whereas auditory processing remains intact. Thus, the additional motion information in the current study contributes to the overshadowing of auditory processing. This provides support to the idea that the dynamicity of visual and auditory input plays a role in the allocation of attention to these two modalities.

Although children in the baseline conditions exhibited high accuracy across the board, the levels of attenuation in the combined condition differed significantly across information types. The most attenuation occurred for item discrimination, followed by motion discrimination, and the least attenuation to sound discrimination. The fact that auditory discrimination was overshadowed least suggests that the auditory system may have privileged status early in development.

Interestingly, children's reaction times in the item discrimination baseline were significantly faster than in the other baseline conditions. Note that accuracy in this condition was high and equivalent to the other conditions. Therefore, it is unlikely that differential attenuation occurred due to differences in baseline discrimination. If this had been the case, one would have expected the least overshadowing for item discrimination.

Overall, the current study demonstrates that there may be multiple influences at play in the manifestation of auditory overshadowing of visual information. Pulls on attention appear to be sensitive to increased stimulus dynamicity. In addition, auditory information receives privileged processing even in the face of competing dynamic visual input. In terms of the directionality of overshadowing effects, however, the current study is limited in its potential conclusions. Children were exposed to one information type in the baseline conditions and all three information types in the combined condition. It is impossible to speculate the direction of overshadowing effects without comparing the extent of children's overshadowing in the combined condition to conditions involving only two types of information. For example, this would allow us to compare overshadowing when sound, motion, and visual information are combined to the traditional overshadowing effect involving only auditory and static visual stimuli. Without this manipulation (which is currently in progress), we are unable to identify whether the additional motion

information contributes to the attenuation of item discrimination.

Further research will also need to elucidate whether increasing the dynamicity of visual input could benefit young children's word learning. Previous work indicates that, in some cases, novel sounds like the ones used in the current experiment overshadow visual input, whereas familiar sounds and words do not. Perhaps the use of more familiar auditory stimuli would result in less overshadowing overall. If this were the case, this increased attention to visual and motion information would provide more optimal conditions for learning object-word mappings. Future work could also investigate whether these mappings occur more easily when visual stimuli are presented dynamically.

In addition, it will be interesting to investigate the developmental trajectory of auditory overshadowing in the context of dynamic visual input. Perhaps there are times throughout development in which motion facilitates visual processing, and times when it hinders or distracts. Further, a developmental investigation can help us identify related processes that could help explain the typically observed reduction in overshadowing with age. For example, perhaps increases in working memory capacity are associated with the developing ability to bind auditory and visual information (Allen, Hitch, & Baddeley, 2009). This research could illuminate ways of presenting particular stimuli to encourage processing between different modalities. Once we understand the developmental process contributing to audio-visual association, we can apply those principles to identify optimal learning conditions in order to teach children more efficiently.

Conclusions

In sum, the current research points to several important findings. When auditory stimuli were presented with dynamic visual stimuli, children's processing of auditory information attenuated. This finding is the first evidence that visual information may interfere with auditory processing early in development. Although auditory discrimination attenuated, it attenuated less than the discrimination of motion patterns or item appearances. The combination of these findings indicates that auditory overshadowing may have multiple underlying causes. Children's attention to dynamicity in the visual modality seemed to pull attention ordinarily devoted to auditory processing. However, the auditory information was still processed better than motion or item appearance information in the face of this competing input. Thus, auditory overshadowing seems to be a function of privileged processing of auditory information.

Acknowledgements

This research was supported by a National Science Foundation (NSF) grant (BCS-0720135) to Vladimir Sloutsky. We thank the children, parents, and preschools that participated in this research.

References

- Allen, R. J., Hitch, G. J., & Baddeley, A. D. (2009). Cross-modal binding and working memory. *Visual Cognition, 17*(1-2), 83-102.
- Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual review of psychology, 48*(1), 269-297.
- Green, D. M., & Von Gierke, S. M. (1984). Visual and auditory choice reaction times. *Acta psychologica, 55*(3), 231-247.
- 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.
- Robinson, C. W., & Sloutsky, V. M. (2004a). Auditory dominance and its change in the course of development. *Child Development, 75*, 1387-1401.
- Robinson, C. W., & Sloutsky, V. M. (2004b). The effect of stimulus familiarity on modality dominance. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the XXVI Annual Conference of the Cognitive Science Society*, 1167-1172 Mahwah, NJ: Erlbaum.
- Robinson, C. W., & Sloutsky, V. M. (2007a). Auditory dominance: Overshadowing or response competition? In D. McNamara & G. Trafton (Eds.), *Proceedings of the XXIX Annual Conference of the Cognitive Science Society* (pp. 605-610) Mahwah, NJ: Erlbaum.
- Robinson, C. W., & Sloutsky, V. M. (2007b). 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.
- 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.
- Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. The MIT Press.