Published in final edited form as:

J Exp Psychol Hum Percept Perform. 2012 December ; 38(6): 1371–1379. doi:10.1037/a0028348.

Spatial attention modulates the precedence effect

Sam London,

University of California, Davis Center for Mind and Brain, 267 Cousteau Place, Davis, CA 95618

Christopher W. Bishop, and

University of California, Davis Center for Mind and Brain, 267 Cousteau Place, Davis, CA 95618

Lee M. Miller

University of California, Davis Department of Neurobiology, Physiology and Behavior, One Shields Avenue, Davis, CA 95616; Also at University of California, Davis Center for Mind and Brain, 267 Cousteau Place, Davis, CA 95618

Abstract

Communication and navigation in real environments rely heavily on the ability to distinguish objects in acoustic space. However, auditory spatial information is often corrupted by conflicting cues and noise such as acoustic reflections. Fortunately the brain can apply mechanisms at multiple levels to emphasize target information and mitigate such interference. In a rapid phenomenon known as the precedence effect, reflections are perceptually fused with the veridical primary sound. The brain can also use spatial attention to highlight a target sound at the expense of distracters. Although attention has been shown to modulate many auditory perceptual phenomena, rarely does it alter how acoustic energy is first parsed into objects, as with the precedence effect. This brief report suggests that both endogenous (voluntary) and exogenous (stimulus-driven) spatial attention have a profound influence on the precedence effect depending on where they are oriented. Moreover, we observed that both types of attention could enhance perceptual fusion while only exogenous attention could hinder it. These results demonstrate that attention, by altering how auditory objects are formed, guides the basic perceptual organization of our acoustic environment.

Keywords

precedence effect; cross-modal attention; spatial attention

Introduction

Perception of auditory space plays a fundamental role in our interactions with the world, but key acoustic spatial cues are often corrupted by noisy interference. Reflections are among the most detrimental forms of acoustic noise because they can cause listeners to perceive a single veridical source (e.g. a talker at a cocktail party) as multiple sources from different locations (e.g. walls, floor). This presents a special challenge for the auditory system as compared to vision because sound reflections from one object commonly mix with those from unrelated events, and because space is not explicitly represented along the sensory epithelium (Recanzone & Sutter, 2008). Fortunately there is a perceptual mechanism for coping with reflections in complex auditory scenes called the precedence effect, whereby the perceived location of an auditory object is determined by the first or “precedent” sound

wave that reaches the ears (Wallach, Newman, & Rosenzweig, 1949). The fusion of the reflections with the primary sound is due to suppression of the spatial information of subsequent, conflicting reflections and the combination of multiple events into a single, coherent object (Litovsky, Colburn, Yost, & Guzman, 1999; Litovsky & Shinn-Cunningham, 2001). The effect is behaviorally consequential, for instance with speech, where the precedence effect appears to improve intelligibility significantly (Brandewie & Zahorik, 2010). This remarkable process can be observed in evoked brain responses soon after sound onset (as early as ~50 msec) (Backer, Hill, Shahin, & Miller, 2010), likely beginning between the midbrain (Litovsky & Yin, 1998a, 1998b) and early auditory cortex (Liebenthal & Pratt, 1997; Liebenthal & Pratt, 1999).

While the peripheral auditory system constructs and initiates basic spatial representations, several higher-level processes also strongly influence the perceptual outcome. Most notably, selective attention has a dominant effect by emphasizing one acoustic source while filtering out the rest (Cherry, 1953). For instance when several people converse, the onset of one talker's voice automatically orients a listener's attention to her location, thus reducing interference from background noise and other talkers. This type of stimulus-driven attentional orienting, called exogenous attention (Corbetta & Shulman, 2002), depends on novelty (e.g. a new talker in the conversation), salience (how loud the talker is), and behavioral relevance (whether the voice is familiar), all crucial attributes for parsing noisy scenes (Bregman, 1994). Conversely with endogenous or goal-directed attention, a listener may deliberately focus on a particular talker or acoustic feature independent of its physical salience. Exogenous attention, in addition to operating involuntarily, acts considerably faster than endogenous attention (Muller & Rabbitt, 1989), but both can improve performance dramatically (Corbetta & Shulman, 2002). Neurally, both forms of attention are coordinated by frontal and parietal or temporo-parietal cortices though they can influence many levels of processing depending on the relevant features and task demands (Hopf, Luck, Boelmans, Schoenfeld, Boehler, Rieger, & Heinze, 2006; Kanwisher & Wojciulik, 2000; Kastner & Ungerleider, 2000; Rinne, Stecker, Kang, Yund, Herron, & Woods, 2007; Salmi, Rinne, Koistinen, Salonen, & Alho, 2009). In the auditory system attentional modulation has routinely been observed in early auditory cortices (Busse, Roberts, Crist, Weissman, & Woldorff, 2005; Woods, Stecker, Rinne, Herron, Cate, Yund, Liao, & Kang, 2009) and with very short latencies (Hackley, Woldorff, & Hillyard, 1987; Woldorff & Hillyard, 1991).

The precedence effect therefore seems to be a process initiated in the early auditory system, while selective attention appears to be initiated in association cortices yet can modulate sensory representations. We first reasoned that exogenous attention should reveal an interaction with the precedence effect since it is both rapid and involuntary. We tested this interaction in a paradigm using lead-lag noise bursts with light flashes as attentional cues. Exogenous attentional effects have been demonstrated cross-modally in a number of studies, both with visual and auditory cues (Spence & Driver, 2004; Koelewijn, Bronkhorst, & Theeuwes, 2010). In our case, non-auditory cues are crucial to avoid altering recent auditory history, which has a strong influence on the precedence effect (Clifton, 1987; Clifton & Freyman, 1989; Clifton, Freyman, Litovsky, & McCall, 1994). We then performed a second experiment to determine whether endogenous attention, though slower and controlled by different neural systems than exogenous attention (Corbetta & Shulman, 2002), had a similar functional effect on the ability to perceive the lagging sound. The results show a significant interaction between the precedence effect and selective attention of both types. This demonstrates how auditory processes, even at very short latencies, can be modulated by attention to improve the perception of complex auditory scenes.

Methods

Experiment 1: Exogenous Attention

Subjects—In accordance with policies approved by the University of California, Davis Institutional Review Board, written informed consent was provided by 33 subjects (22 female, mean age of 21.9 years, age range 18-32 years). Subjects were screened for self-reported good hearing and normal or corrected-to-normal vision. Eight subjects were excused from the study prior to data collection due to an established echo threshold below 2 msec time delay. A 2 msec temporal delay was chosen to ensure accurate lateralization of the sounds to the location of the speakers, based on prior reports suggesting that subjects have difficulty lateralizing stimuli with shorter offsets (Litovsky, et al., 1999; Sanders, Joh, Keen, & Freyman, 2008; Bishop, London, & Miller, 2011). Three other subjects were excluded from data analysis due to poor control performance during the main task. Poor performance was defined as less than 70% *one-location* responses in one or more control conditions (see below). Thus, the findings reported here are based on 22 individuals (16 female).

Temporal integration pilot test—As the exogenous attention experiment would use a visual cue and auditory stimulus separated by only 300 msec, we first confirmed that audiovisual integration played no role in the paradigm. To do this a small pilot study was conducted where 5 subjects who did not participate in Experiment 1 (4 female, mean age of 20.6 years, age range 19-23 years), were presented with light flashes and acoustic noise burst pairings, identical to those used in the main experiment (see below), but with varying temporal offsets. The two stimuli always originated from the same spatial location (left speaker). Training consisted of the presentation of a clear “same time” pair (0 msec offset) and a clear separate pair (350 msec offset), followed by a mock run so that subjects were comfortable with the task. Test stimuli consisted of temporal offsets (visual leading) of 0 msec, 25 msec, 150 msec, and 300 msec. Subjects were instructed to maintain fixation on a small cross in the middle of the screen and respond after it flashed green (roughly a second after stimulus completion) with a button press to indicate whether they perceived the flash and sound simultaneously or separately in time. Both auditory and visual stimuli were presented from the same peripheral locations as during the main experiment (see below).

Experimental setup—The experiment was conducted in a double-walled, sound attenuated chamber measuring 2.64 by 3.4 meters and 2.43 meters high. Though the chamber was not anechoic, the first acoustic reflection was attenuated by approximately 12 dB when it reached the listener's ear approximately 5.8 msec after the first (precedent) sound wave. Instructions were presented on a computer monitor (Dell FPW2407) 120 cm in front of the subject. The subject sat in an acoustically transparent chair (Herman Miller AE500P) and entered his or her responses via keyboard. The auditory stimuli were presented through two Tannoy Precision 6 speakers positioned 18° right and left of midline in the horizontal plane of the ears, 125 cm away from the head. Two white light emitting diodes (LEDs; 7 candelas, 1.7° visual angle) positioned at the top edge of the speaker cones provided exogenous attentional cues. The LEDs were placed inside of white ping-pong balls to increase their size and perceptual salience. Auditory stimuli were 15 msec noise bursts with a 0.5 msec linear rise and fall, intended to match the approximate rise time of the LEDs. The exogenous visual cues lasted 15 msec. All auditory stimuli were presented at approximately 75 dB(A) from each speaker. Identical, lead-lag noise burst pairs were presented to simulate the arrival of a leading (precedent) sound and its corresponding lag; the temporal delay of the lag relative to the lead noise burst was calibrated for each subject prior to the experimental session to evoke fusion with the precedent sound on approximately 50% of trials.

Stimulus conditions—Stimuli were presented in three configurations: lead-lag pairs alone (A_{PE}) or preceded by a visual attentional cue at the lead ($A_{PE}C_{Lead}$) or lag ($A_{PE}C_{Lag}$) location, 300 msec prior to sound onset (Figure 2A). A visual-auditory delay of 300 msec is well outside the temporal window of perceptual fusion, as confirmed in a pilot study we performed (Figure 1), and further substantiated by other reports (Slutsky & Recanzone, 2001; Zampini, Guest, Shore, & Spence, 2005). Additional control stimuli were presented in three conditions: a single noise burst presented without a cue (A) or a single noise burst preceded by a visual cue on the same side (AC_{Lead}) or on the side of the absent lag (AC_{Lag}), 300 msec prior to sound onset. These controls were included to ensure that subjects were performing the task as instructed and that their responses were not simply biased by the lights.

Task—Subjects were instructed to maintain fixation on a small cross in the center of the computer monitor and to ignore the flashing lights. Each trial began with the combined stimulus delivery (the visual cue, 300 msec delay, and the auditory stimulus, together lasting ~320 msec depending on the lead-lag temporal offset) followed by a ~1.4 sec delay, depending on stimulus duration, prior to the subject response (Figure 2B). Subject responses were delayed to emphasize accuracy as opposed to speed of the perceptual report. Subjects were prompted to respond by a color change (white to green) of the fixation cross 1.5 sec after stimulus onset. For each trial the subject responded to two questions. First they indicated with their right index, middle, and ring fingers whether they heard a sound on their *left-side*, *both-sides*, or *right-side*. Second they indicated whether they heard a sound from *one-location* or *two-locations* with their left index and middle fingers. The questions were asked separately in order to distinguish three possible spatial percepts of the stimuli: a reflection with largely unsuppressed spatial information perceived in the opposite hemifield of the leading sound, a reflection with largely suppressed spatial information perceived at the location of the lead and fused with it, or a reflection with partially suppressed spatial information translocated to the same side as the leading sound but remaining a separate object (Litovsky, et al., 1999; Backer, et al., 2010). A new trial began every 2.75 – 3.15 sec. All stimulus presentation and response recording was handled through Neurobehavioral System's Presentation Software. Unique noise samples were dynamically generated by MATLAB using in-house functions for every trial.

The experiment began with a short training session. First subjects were presented with a verbally narrated PowerPoint presentation to familiarize them with the task. Then they underwent a brief, self-paced battery of examples from each of the stimulus conditions, including controls. In an effort to familiarize subjects with the potential percepts for each condition, reflections intended to be clearly fused were presented with a 2 msec temporal offset and those intended to be clearly detected (not fused) were presented with a 20 or 100 msec offset. Once subjects were comfortable with the stimuli, they underwent a three minute mock session of the task.

Following the training session, subjects completed an echo threshold calibration test. The test consisted only of stimuli in the A_{PE} configuration. Each subject's echo threshold (50% one-location responses) was estimated with a one-up/one-down staircase algorithm, described in our previous work (Bishop, et al., 2011). Subjects were instructed to respond to the stimuli just as they would for the main task. The lead-lag temporal offset was initialized at 2 msec and incremented or decremented by 0.5 msec for *one-location* or *two-locations* responses respectively. The stimulus lead side was selected pseudo-randomly such that each side was represented three times in a group of six. The test concluded after the direction of the temporal change reversed ten times. Following the threshold test, the experimenter presented subjects with a series of dynamically generated auditory stimuli in the A_{PE} configuration and solicited verbal reports to the same two questions. The experimenter then

made any small adjustments to the lead-lag delay used for the remainder of the experiment to better target 50% fusion in condition A_{PE} based on manual presentation of sound pairs near the estimated echo threshold. These changes were usually prompted by response saturation (either nearly all suppressed or none suppressed) at the derived echo threshold. In order to ensure adequate numbers of suppressed trials when echo threshold differed by side (Grantham, 1996), thresholds were set to the shorter of the two sides.

The main task lasted approximately 15 minutes and consisted of 50 trials in each of the 6 conditions for a total of 300 trials. The trials were pseudo-randomized such that a single trial type was not repeated for more than 5 consecutive trials. Subjects were offered an optional break every five minutes. The percentage of *one-location* responses was used to quantify echo suppression, since it implies perceptual fusion (Litovsky, et al., 1999; Bishop, et al., 2011).

Statistical analysis—Statistical significance was based on an alpha level of 0.05. Approximate p-values are reported when exceeding three significant digits (e.g. $p < 0.001$). Statistically significant differences in fusion rate were quantified by including the percentage of *one-location* responses as the dependent measure in a repeated measure ANOVA in STATISTICA v.8.0 (www.statsoft.com). As in a recent related report (Bishop, et al., 2011), responses indicative of partial suppression of the reflection's spatial information (e.g. lead side, two locations) were extremely rare, making the side question largely redundant (Table 1). Reported p-values are corrected for non-sphericity with Greenhouse-Geisser correction where appropriate. Post-hoc tests were performed with Fischer's Least Significant Difference (LSD). Unless otherwise noted, means are reported with the standard error of the mean (SEM).

Experiment 2: Endogenous Attention

A second experiment was performed, largely identical to the first, but investigating the influence of endogenous or volitional spatial attention on the precedence effect. Only the differences between the two experiments are described here. Informed consent was provided by 11 subjects (9 female, mean age of 19.8 years, age range 18-25) by the same methods as described in Experiment 1. Subjects underwent virtually identical training and threshold estimation and were subjected to the identical rejection criterion (70% or greater *one-location* responses in control conditions). One subject was excused prior to data collection because they could not adequately perform the task after extensive training. Consequently, the findings reported are based on 10 individuals (8 female). No subjects from Experiment 1 participated in Experiment 2.

Unlike the exogenous, spatially coincident visual cues employed in Experiment 1, Experiment 2 employed central arrows to cue subjects to covertly orient their attention to a spatial hemifield, meaning doing so without intentional movement of the head or eyes. The arrow cues were presented 1.28 degrees above fixation and extended across the central 3.84 degrees of the visual field and pointed to the left, right, or were double-ended. In each trial an arrow appeared on screen for 1.25 sec and its offset preceded audio stimuli by 1 sec (Figure 3). Each arrow cue either pointed to the arriving lead, pointed to the arriving lag, or was neutral (double-ended, similar to those in a classic Posner cueing paradigm (Posner, 1980)). Trials were presented in 6 unique cue-stimulus configurations. Auditory test stimuli comprised three conditions: lead-lag pairs preceded by a double-ended uninformative arrow cue ($A_{PE}C_{Neutral}$), and pairs preceded by either an arrow pointing to the arriving lead ($A_{PE}C_{Lead}$) or lag ($A_{PE}C_{Lag}$) sound source. Control stimuli were presented in three conditions: a single noise burst preceded by a neutral arrow ($AC_{Neutral}$), an arrow pointing to the source (AC_{Lead}), or an arrow pointing to the absent lag (AC_{Lag}). These controls were

included to ensure that subjects were only responding to the sounds they heard and that their responses were not biased by the arrow cues. Leading side was again varied pseudorandomly.

As in Experiment 1, subjects were instructed to maintain fixation on a cross in the center of the monitor, however they were also told to covertly orient their attention to the speaker indicated by the arrow cue. In trials containing the double-ended neutral arrow cue they were told to maintain fixation on the cross but not focus anywhere specifically. Subjects were told that cues would not necessarily indicate where they would hear sound but that they were to orient their attention as directed anyway. Subjects were prompted to respond to the side (left/right/both) and location number (one/two) questions by the changing the fixation cross color from white to green approximately 1 sec after auditory stimulus onset. Subjects were given .75 seconds to respond (Figure 3). The same keys and response types were used as those in experiment 1.

Eye tracking data were collected using the SR-Research Eye Link 1000 (www.sr-research.com) to ensure subjects maintained fixation at a sampling rate of 1000 Hz. The eye tracker was calibrated at the beginning of the experiment and after breaks. An additional automatic fixation check was performed every 15 trials. Specifically, if a subject's gaze exceeded 2.5 degrees from fixation, the experimenter was prompted to recalibrate the eye tracker before continuing the experiment. The main task was approximately 20 minutes and 300 trials long, including 60 trials in each of the 6 test conditions and 40 trials in each of the control conditions. Subjects were given a break every five minutes to ensure a consistent level of alertness.

All eye fixation location data were recorded in pixels. This information combined with the known distance of the head to the monitor was used to calculate visual angle. Due to inherent noise in the eye tracking setup each subject's data were mean-centered prior to analysis. The azimuthal fixation for a trial was calculated by taking all recorded fixation events which began in a 1 second window centered on sound presentation. The mean fixation point of the trial was determined using the events' azimuthal positions weighted by the duration of the fixation event within the window. These weighted means for each condition and percept were used as the dependent measure in a repeated measure ANOVA. Reported p-values are corrected in the same manner reported in experiment 1.

Results

Experiment 1: Exogenous Attention

The average echo threshold used was 6.2 msec with a SEM of 0.45 msec. The manual adjustments made after automatic thresholding, which ensured that subjects understood all potential percepts and accommodated any early practice effects, averaged -2.89 ± 0.65 msec. That is, subjects were typically more conservative in their *two-location* responses during the thresholding procedure than during the main task. When lead-lag noise bursts were presented in the absence of an attentional cue (A_{PE}), subjects reported *one-location* in $43 \pm 5\%$ SEM of trials. In contrast, subjects reported *one-location* in $54 \pm 5\%$ of trials when attention was cued to the lead side ($A_{PEC_{Lead}}$), and $35 \pm 5\%$ of the time with the cue on the lag side ($A_{PEC_{Lag}}$) (Figure 4). The percentage of *one-location* responses was included as a repeated measure in a two-way ANOVA with factors [Lead-Side $\times$ Condition]. The main effect of Condition was significant ($p < 0.001$), while Leading-Side ($p=0.746$; $43 \pm 5\%$ for left-leading; $45 \pm 5\%$ for right-leading) and the interaction ($p=0.293$) were not, suggesting that hemifield did not matter. Post-hoc tests revealed the following pattern of results: $A_{PEC_{Lead}} > A_{PE} > A_{PEC_{Lag}}$. In other words, fusion was enhanced by $11 \pm 3\%$ ($p = 0.002$)

when attention was oriented to the lead sound's location, and hindered by $8 \pm 3\%$ ($p = 0.025$) when oriented toward the lagging location.

A series of control conditions were included to ensure subjects were performing the task and to rule out response bias from the visual stimulation. The control conditions consisted of a single noise burst on one side without an attentional cue (A), or a single noise burst on one side preceded by a visual cue (300 msec leading) on either the same (AC_{Lead}) or opposite side (AC_{Lag}). Subjects exceeded 96% *one-location* responses on average on these conditions. Performance on these conditions suggests that subjects were performing the task well and that responses were not overtly biased by merely the presence of a visual flash.

Experiment 2: Endogenous Attention

The average echo threshold used was 7.5 msec with a SEM of 0.95 msec. As described in the Methods section, individual echo thresholds were initially estimated using a one-up-one-down staircase algorithm (see Methods for details). However, these estimates often required some minor, manual adjustments. Specifically, manual adjustments to the automatically acquired echo threshold averaged -0.88 ± 0.14 msec. During the neutral arrow cue condition ($A_{PE}C_{Neutral}$), subjects reported *one-location* in $43 \pm 8\%$ of trials. In contrast, subjects reported *one-location* on $53 \pm 8\%$ of trials when attention was prompted by the lead arrow cue ($A_{PE}C_{Lead}$), and $44 \pm 10\%$ of trials by the lag arrow cue ($A_{PE}C_{Lag}$) (Figure 5). The percentage of *one-location* responses was included as a repeated measure in a two-way [Lead-Side $\times$ Condition], within-subject ANOVA. As in Experiment 1, the main effect of Condition was significant ($p=0.031$), while the Lead-Side ($p=0.972$; $47 \pm 10\%$ for left-leading; $47 \pm 9\%$ for right-leading) and interaction ($p=0.854$) were not. Post-hoc tests showed that the percentage of *one-location* responses during $A_{PE}C_{Lead}$ trials was greater than on $A_{PE}C_{Neutral}$ trials but did not show a significant difference between the neutral and the lag arrow cue conditions ($A_{PE}C_{Lead} > A_{PE}C_{Neutral}, A_{PE}C_{Lag}$). In other words, the percentage of *one-location* responses increased by $10 \pm 3\%$ ($p = 0.016$) when attention was oriented to the lead sound's location compared to the neutral cueing condition. In contrast, attentional cueing to the lagging sound's location had no significant effect on subject responses ($p=0.751$). Subjects exceeded 94% accuracy on average on control conditions. Performance on these controls suggests that responses were not biased by the arrow cues.

Experiment 2's cue persisted for over a second and required each subject to voluntarily shift his or her attention to a location in space. It is possible that the reported change in *one-location* responses is due in part to changes in gaze position (i.e. subjects look where they are instructed to covertly orient their attention). To rule this out as a possibility, we included the weighted mean of the azimuthal coordinates of eye fixation (Figure 6) as the dependent measure in a two-way within-subject ANOVA with factors [Condition $\times$ Number of Locations (one or two)]. The main effects of Condition and Number of Locations Response were insignificant ($p=0.954$ and $p=0.971$ respectively), while the interaction ($p=0.013$) was significant. Post-hoc tests showed that significance in the interaction was between *one-location* and *two-locations* responses for $A_{PE}C_{Lead}$ and *one-location* responses for $A_{PE}C_{Lag}$. The direction of the effect indicates that subjects were more likely to suppress when fixating in the direction of the arrow. Despite being statistically significant, the mean difference was remarkably small ($0.13 \pm 0.06^\circ$).

Discussion

Our results demonstrate clearly that both exogenous and endogenous attention can enhance the precedence effect, depending on where it is deployed relative to the leading acoustic wave or subsequent reflection. This functional relationship would likely be advantageous in a cluttered natural environment with multiple sources and their spatiotemporally overlapping

reflections. Attention appears to provide additional, top-down evidence for how the auditory system should parse acoustic energy into distinct objects.

Mechanistically, the dominant view of the precedence effect has been of a fast-acting, low-level event, implying that it ought to be insulated from most cognitive processes. Specifically, several studies assert that the precedence effect occurs in the inferior colliculus (IC) (Litovsky & Yin, 1998a, 1998b; Fitzpatrick, Kuwada, Kim, Parham, & Batra, 1999), though strong arguments point to a later stage (Cranford & Oberholtzer, 1976; Cranford, Ravizza, Diamond, & Whitfield, 1971; Liebenthal & Pratt, 1997; Liebenthal & Pratt, 1999). The present study contributes to growing evidence that the precedence effect is not an isolated, fully automatic subcortical process (Backer, et al., 2010; Clifton, et al., 1994; Cranford, et al., 1971; Damaschke, Riedel, & Kollmeier, 2005; Dimitrijevic & Stapells, 2006; Sanders, et al., 2008; Sanders, Zobel, Freyman, & Keen, 2011; Spierer, Bourquin, Tardif, Murray, & Clarke, 2009). Our results alone cannot specify where attention influences fusion, but they inform hypotheses about which neural structures are involved. Anatomically subcortical involvement remains a possibility, as the superior colliculus (SC) has a well-established role in spatial attention (Ignashchenkova, Dicke, Haarmeier, & Thier, 2004), and attentional effects have been seen as early in the lemniscal auditory pathway as inferior colliculus, likely realized through efferent projections from auditory cortex (Winer, 2006; Rinne, Balk, Koistinen, Autti, Alho, & Sams, 2008). However, depending on the task and stimuli, the earliest levels of attentional modulation are more often reported in non-primary auditory cortices (Woods, et al., 2009). Regardless of which neural levels attentional modulation is realized, our results show that it must occur where acoustic information from reflections is still present. That is, attention acts before fusion is complete. Based on this information we might speculate that attention interacts with the precedence effect much as it would with any other auditory spatial process: the supramodal orienting network in the temporo-parietal junction (Corbetta & Shulman, 2002), engaged by a visual or auditory cue, could modulate space-sensitive auditory cortex such as medial planum temporale through a spatial gain function (Deouell, Heller, Malach, D'Esposito, & Knight, 2007). This would increase the perceptual salience of acoustic energy at the cue location and decrease it elsewhere, thereby shifting the echo threshold in the manner observed here.

Interestingly, exogenous and endogenous attention had virtually identical effects when cued to the lead sound location, but only exogenous attention modulated perception when cued to the reflection (i.e. the lag location). Given evidence that the precedence effect is a temporally extended process occurring along the auditory pathway (Backer, et al., 2010; Liebenthal, et al., 1999; Litovsky & Pratt, 1999; Sanders, et al., 2008; Sanders, et al., 2011), perhaps the two forms of attention have access to different levels of representation. Exogenous attention may act at a lower or earlier level where spatial information for the reflection is still present, before it is perceptually translocated or fused with the primary sound. In contrast, endogenous attention may act only after reflection cues have been incorporated or discarded. Thus it can alter the perceptual decision in favor of the lead sound but not in favor of the reflection. Of course, this interpretation presumes that subjects followed directions and attended effectively to the endogenous cues, which we cannot directly confirm in the present paradigm. Nevertheless, building on our neuroanatomical speculation above, our results suggest that the inferior fronto-parietal orienting network may act earlier in the sensory pathway than the superior voluntary control network.

Though the eye tracking collected during the endogenous task confirmed that subjects kept their eyes on the cross, varying in average fixation by a small fraction of a degree, the ANOVA did yield an interaction between condition and percept which required further investigation. We observed that subjects responded *one-location* significantly more often when their gaze shifted slightly in the direction indicated by the arrow, regardless of whether

or not the cue indicated the next arriving sound. It is probable that the slight shift in eye position indicated attentiveness in performing the task, and did not bias the subject to one percept over another. Were this not the case we would expect to see a bias towards responding *one-location* in all conditions when fixation was closer to origin of the leading sound.

This study provides an empirical foundation for several important theoretical questions that could be addressed in future work. Further research could test how attention affects dissociable components of the precedence effect: namely the spatial phenomenon where information about reflection location is modulated (“spatial translocation”), versus the actual fusion of images into a single coherent percept at the lead-sound location (“object fusion”), which have been shown to have different neural correlates (Backer, et al., 2010). Whether attention affects one or both of these precedence effect components would provide further insight into its level of action and specificity. It could also be worthwhile to investigate how the phenomenon evolves in time or across space: for instance whether attention interacts with the well-documented ‘buildup’ of the precedence effect (Clifton, et al., 1994; Grantham, 1996; Litovsky, et al., 1999), or whether its effects differ when both sound sources are located in the same hemifield. Electroencephalography (EEG) or magnetoencephalography (MEG) could be used to characterize how attention affects the neural time course of the precedence effect, which extends over hundreds of msec (Sanders, et al., 2008; Backer, et al., 2010; Sanders, et al., 2011). Functional magnetic resonance imaging would also be informative, as the relevant attentional networks that we presume modulate suppression are well established. Through such convergent approaches, we can characterize the causal, neural mechanisms by which attention affects this crucial real-world perceptual ability.

Acknowledgments

This work was supported by the National Institutes of Health: National Institute on Deafness and other Communication Disorders [R01-DC8171 and R01-DC8171-04S1 to LMM, and F31DC011429 to CWB].

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Deafness and other Communication Disorders or the National Institutes of Health.

References

- Backer KC, Hill KT, Shahin AJ, Miller LM. Neural time course of echo suppression in humans. *Journal of Neuroscience*. 2010; 30(5):1905–1913. [PubMed: 20130199]
- Best V, Ozmeral EJ, Kopco N, Shinn-Cunningham BG. Object continuity enhances selective auditory attention. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105(35):13174–13178. [PubMed: 18719099]
- Bishop CB, London S, Miller LM. Visual influences on echo suppression. *Current Biology*. 2011; 21(3):221–225. [PubMed: 21276724]
- Brandewie E, Zahorik P. Prior listening in rooms improves speech intelligibility. *Journal of the Acoustical Society of America*. 2010; 128(1):291–299. [PubMed: 20649224]
- Bregman, AS. *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, Massachusetts: MIT Press; 1994. p. 1-790.
- Busse L, Roberts KC, Crist RE, Weissman DH, Woldorff MG. The spread of attention across modalities and space in a multisensory object. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102(51):18751–18756. [PubMed: 16339900]
- Cherry EC. Some experiments on the recognition of speech with one and with two ears. *Journal of the Acoustical Society of America*. 1953; 25:975–979.
- Clifton RK. Breakdown of echo suppression in the precedence effect. *Journal of the Acoustical Society of America*. 1987; 82(5):1834–1835. [PubMed: 3693711]

- Clifton RK, Freyman RL. Effect of click rate and delay on breakdown of the precedence effect. *Perceptual Psychophysiology*. 1989; 46(2):139–145.
- Clifton RK, Freyman RL, Litovsky RY, McCall D. Listeners' expectations about echoes can raise or lower echo threshold. *Journal of the Acoustical Society of America*. 1994; 95:1525–1533. [PubMed: 8176056]
- Clifton RK, Freyman RL, Meo J. What the precedence effect tells us about room acoustics. *Perceptual Psychophysiology*. 2002; 64(2):180–188.
- Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*. 2002; 3(3):201–215.
- Cranford J, Ravizza R, Diamond IT, Whitfield LC. Unilateral ablation of auditory cortex in cat impairs complex sound localization. *Science*. 1971; 172:286–288. [PubMed: 5548713]
- Cranford JL, Oberholtzer M. Role of neocortex in binaural hearing in the cat. II. The 'precedence effect' in sound localization. *Brain Research*. 1976; 111(2):225–239. [PubMed: 949598]
- Damaschke J, Riedel H, Kollmeier B. Neural correlates of the precedence effect in auditory evoked potentials. *Hearing Research*. 2005; 205(1-2):157–171. [PubMed: 15953525]
- Deouell LY, Heller AS, Malach R, D'Esposito M, Knight RT. Cerebral responses to change in spatial location of unattended sounds. *Neuron*. 2007; 55:985–996. [PubMed: 17880900]
- Dimitrijevic A, Stapells DR. Human electrophysiological examination of buildup of the precedence effect. *Neuroreport*. 2006; 17(11):1133–1137. [PubMed: 16837841]
- Fitzpatrick DC, Kuwada S, Kim DO, Parham K, Batra R. Responses of neurons to click-pairs as simulated echoes: Auditory nerve to auditory cortex. *Journal of the Acoustical Society of America*. 1999; 106(6):3460–3472. [PubMed: 10615686]
- Galbraith GC, Doan BQ. Brainstem frequency-following and behavioral responses during selective attention to pure tone and missing fundamental stimuli. *International Journal of Psychophysiology*. 1995; 19:203–214. [PubMed: 7558987]
- Grantham DW. Left-right asymmetry in the buildup of echo suppression in normal-hearing adults. *Journal of the Acoustical Society of America*. 1996; 99(2):1118–1123. [PubMed: 8609295]
- Hackley SA, Woldorff M, Hillyard SA. Combined use of microreflexes and event-related brain potentials as measures of auditory selective attention. *Psychophysiology*. 1987; 24(6):632–647. [PubMed: 3438427]
- Hopf JM, Luck SJ, Boelmans K, Schoenfeld MA, Boehler CN, Rieger J, Heinze HJ. The neural site of attention matches the spatial scale of perception. *Journal of Neuroscience*. 2006; 26(13):3532–3540. [PubMed: 16571761]
- Ignashchenkova A, Dicke PW, Haarmeier T, Thier P. Neuron-specific contribution of the superior colliculus to overt and covert shifts of attention. *Nature Neuroscience*. 2004; 7(1):56–64.
- Kanwisher N, Wojciulik E. Visual attention: Insights from brain imaging. *Nature Reviews Neuroscience*. 2000; 1(2):91–100.
- Kastner S, Ungerleider LG. Mechanisms of visual attention in the human cortex. *Annual Review of Neuroscience*. 2000; 23:315–341.
- Koelewijn T, Bronkhorst A, Theeuwes J. Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta Psychologica*. 2010; 134(3):372–384. [PubMed: 20427031]
- Liebenthal E, Pratt H. Evidence for primary auditory cortex involvement in the echo suppression precedence effect: a 3CLT study. *Journal of Basic and Clinical Physiology and Pharmacology*. 1997; 8(3):181–201. [PubMed: 9429985]
- Liebenthal E, Pratt H. Human auditory cortex electrophysiological correlates of the precedence effect: Binaural echo lateralization suppression. *Journal of the Acoustical Society of America*. 1999; 106(1):291–303.
- Litovsky RY, Yin TC. Physiological studies of the precedence effect in the inferior colliculus of the cat. I. Correlates of psychophysics. *Journal of Neurophysiology*. 1998a; 80(3):1285–1301. [PubMed: 9744939]
- Litovsky RY, Yin TC. Physiological studies of the precedence effect in the inferior colliculus of the cat. II. Neural mechanisms. *Journal of Neurophysiology*. 1998b; 80(3):1302–1316. [PubMed: 9744940]

- Litovsky RY, Colburn HS, Yost WA, Guzman SJ. The precedence effect. *Journal of the Acoustical Society of America*. 1999; 106(4):1633–1654. [PubMed: 10530009]
- Litovsky RY, Shinn-Cunningham BG. Investigation of the relationship among three common measures of precedence: Fusion, localization dominance, and discrimination suppression. *Journal of the Acoustical Society of America*. 2001; 109:346–358. [PubMed: 11206163]
- Muller HJ, Rabbitt PMA. Reflexive and voluntary orienting of visual-attention - time course of activation and resistance to interruption. *Journal of Experimental Psychology-Human Perception and Performance*. 1989; 15(2):315–330. [PubMed: 2525601]
- Posner MI. Orienting of attention. *Quarterly Journal of Experiment Psychology*. 1980; 32(1):3–25.
- Recanzone GH, Sutter ML. The biological basis of audition. *Annual Review of Psychology*. 2008; 59:119–142.
- Rinne T, Stecker GC, Kang X, Yund EW, Herron TJ, Woods DL. Attention modulates sound processing in human auditory cortex but not the inferior colliculus. *Neuroreport*. 2007; 18(13):1311–1314. [PubMed: 17762703]
- Rinne T, Balk MH, Koistinen S, Autti T, Alho K, Sams M. Auditory selective attention modulates activation of human inferior colliculus. *Journal of Neurophysiology*. 2008; 100(6):3323–3327. [PubMed: 18922948]
- Salmi J, Rinne T, Koistinen S, Salonen O, Alho K. Brain networks of bottom-up triggered and top-down controlled shifting of auditory attention. *Brain Research*. 2009; 1286:155–164. [PubMed: 19577551]
- Sanders LD, Joh AS, Keen R, Freyman RL. One sound or two? Object-related negativity indexes echo perception. *Perceptual Psychophysiology*. 2008; 70(8):1558–1570.
- Sanders LD, Zobel BH, Freyman RL, Keen R. Manipulations of listeners' echo perception are reflected in event-related potentials. *Journal of the Acoustical Society of America*. 2011; 129(1):301. [PubMed: 21303011]
- Slutsky DA, Recanzone GH. Temporal and spatial dependency of the ventriloquism effect. *Neuroreport*. 2001; 12(1):7–10. [PubMed: 11201094]
- Spence, C.; Driver, J. *Crossmodal space and crossmodal attention*. Oxford, New York: Oxford University Press; 2004. p. 1-323.
- Spierer L, Bourquin NM, Tardif E, Murray MM, Clarke S. Right hemispheric dominance for echo suppression. *Neuropsychologia*. 2009; 47(2):465–472. [PubMed: 18983863]
- Wallach H, Newman EB, Rosenzweig MR. The precedence effect in sound localization. *American Journal of Psychology*. 1949; 62(3):315–336. [PubMed: 18134356]
- Winer JA. Decoding the auditory corticofugal systems. *Hearing Research*. 2006; 212(1-2):1–8. [PubMed: 16555378]
- Woldorff MG, Hillyard SA. Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and Clinical Neurophysiology*. 1991; 79(3):170–191. [PubMed: 1714809]
- Woods DL, Stecker GC, Rinne T, Herron TJ, Cate AD, Yund EW, Liao I, Kang X. Functional maps of human auditory cortex: Effects of acoustic features and attention. *Plos One*. 2009; 4(4):19.
- Wu CT, Weissman DH, Roberts KC, Woldorff MG. The neural circuitry underlying the executive control of auditory spatial attention. *Brain Research*. 2007; 1134(1):187–198. [PubMed: 17204249]
- Zampini M, Guest S, Shore DI, Spence C. Audio-visual simultaneity judgments. *Perceptual Psychophysiology*. 2005; 67(3):531–544.

Figure 1.

Temporal integration test. Five subjects were tested to ascertain the temporal offset between a co-localized visual cue and an auditory stimulus when they no longer integrated the perceptual objects. Data reflect group means and error bars reflect SEM.

Figure 2.

Task and stimuli. A) Auditory stimuli were presented by two speakers, positioned 18° to the right and left of midline. The visual cues were presented by white LED lights (depicted as small white circles) positioned above the speaker cones. Subjects fixated on a small white cross positioned on a monitor 120 cm from their head. B) The combined stimuli (visual cue, visual offset, and auditory stimulus together) varied in duration depending on the acquired threshold, but averaged around 320 msec. Subjects were instructed to delay their response until it was prompted by a color change, white to green, of the fixation cross. C) Graphical depictions of all stimulus conditions are depicted independent of lead side. Precedence effect conditions were presented either unimodally (A_{PE}), or preceded by a visual attentional cue (15 msec light flash) 300 msec before the arrival of the lead sound on either the lead (A_{PE}C_{Lead}) or lag (A_{PE}C_{Lag}) side. Control conditions did not contain a lagging sound and were either unimodal (A) or with the appropriate visual cue on the same side as the single sound (AC_{Lead}) or the opposite side (AC_{Lag}).

Figure 3.

Trial structure of the endogenous cueing task. Trial duration varied only slightly, again due to minor differences in the subject-specific echo thresholds.

Figure 4.

Rate of *one-location* responses, as a function of condition and side of leading sound. Brackets between conditions represent a significant difference in the rate of *one-location* responses (* $p = 0.025$; ** $p = 0.002$; *** $p < 0.001$). Error bars reflect SEM of group means ($N=22$). The mean differences between conditions are plotted in the right panel.

Figure 5.

Rate of *one-location* responses, as a function of condition and side of leading sound. Brackets between conditions represent a significant difference in the rate of *one-location* responses (* $p = 0.016$; ** $p = 0.031$). Error bars reflect SEM of group means ($N=10$). The mean differences between conditions are also displayed.

Figure 6.

Eye fixation as a function of condition and percept. Azimuthal position of fixation is expressed in degrees of visual angle relative to the fixation cross for each condition, both with the *one-location* (O) and *two-locations* (T) responses. The lead and lag speakers are located at 18.0 and -18.0 degrees visual angle respectively.

Table 1

Response breakdown for side and location questions in Experiment 1 and 2. The average proportion of responses is presented for each response combination within each experimental condition along with the SEM.

Experiment 1		One Location			Two Locations		
		Right Side	Both Sides	Left Side	Right Side	Both Sides	Left Side
Ape	Right Leading	43±5%	1±1%	0±0%	4±2%	52±6%	0±0%
	Left Leading	1±1%	1±1%	41±5%	0±0%	52±6%	5±2%
ApeC _{Lead}	Right Leading	52±6%	0±0%	1±1%	7±3%	40±6%	0±0%
	Left Leading	2±1%	0±0%	54±6%	0±0%	39±5%	5±2%
ApeC _{Lag}	Right Leading	33±6%	2±1%	3±2%	3±2%	59±6%	0±0%
	Left Leading	4±2%	1±1%	28±4%	1±1%	63±5%	3±1%

Experiment 2		One Location			Two Locations		
		Right Side	Both Sides	Left Side	Right Side	Both Sides	Left Side
Ape	Right Leading	39±10%	2±1%	1±1%	4±3%	54±10%	0±0%
	Left Leading	1±1%	1±1%	43±9%	0±0%	52±10%	3±1%
ApeC _{Lead}	Right Leading	52±9%	0±0%	1±1%	4±2%	43±9%	0±0%
	Left Leading	0±0%	1±1%	52±11%	0±0%	45±11%	2±1%
ApeC _{Lag}	Right Leading	43±11%	1±1%	1±1%	3±2%	52±10%	0±0%
	Left Leading	2±1%	2±1%	41±11%	0±0%	52±11%	3±1%