Endogenous orientation of visual attention in auditory space

Gaetana Chillemi, Alessandro Calamuneri, Angelo Quartarone, Carmen Terranova, Adriana Salatino, Alberto Cacciola, Demetrio Milardi, Raffaella Ricci

IRCCS Centro Neurolesi Bonino Pulejo, Contrada Casazza, SS113, 98124 Messina, Italy
Department of Biomedical and Dental Sciences and Morphofunctional Imaging, University of Messina, Via Consolare Valeria 1, Gazi, 98125 Messina, Italy
Department of Clinical and Experimental Medicine, Endocrinology, University of Messina, Via Consolare Valeria 1, Gazi, 98125 Messina, Italy
Department of Psychology, University of Turin, Torino 10123, Italy

Highlights
- Facilitation was observed for right-sided auditory stimuli in a new visuo-audio task.
- Auditory space has dynamic nature, which adapts to changes in visual space.
- Sound localization was enhanced by visual cues.
- Crossmodal links in spatial attention were found between audition and vision.
- These findings have theoretical and translational implications for future studies.

Abstract
Visuospatial attention is asymmetrically distributed with a leftward bias (i.e. pseudoneglect), while evidence for asymmetries in auditory spatial attention is still controversial. In the present study, we investigated putative asymmetries in the distribution of auditory spatial attention and the influence that visual information might have on its deployment. A modified version of the Posner task (i.e. the visuo-audio spatial task [VAST]) was used to investigate spatial processing of auditory targets when endogenous orientation of spatial attention was mediated by visual cues in healthy adults. A line bisection task (LBT) was also administered to assess the presence of a leftward bias in deployment of visuospatial attention. Overall, participants showed rightward and leftward biases in the VAST and the LBT, respectively. In the VAST, sound localization was enhanced by visual cues. Altogether, these findings support the
Visuo-audio task

Thirty-eight healthy volunteers (21 women) were recruited from the community through word-of-mouth and from a non-profit private university for older adults located near Messina, Italy (Third Age University). In order to test whether age influenced participants’ performance, they were categorized by age into two groups: the first group (middle-aged adults) included 19 individuals with their age ranging from 38 to 53 years (mean age: 46.94 years), and the second group (older adults) included 19 individuals with their age ranging from 61 to 71 years (mean age: 64.70 years) [29]. All the participants were naïve as to the aims and expected outcomes of the study. The participants’ hand dominance was assessed using the short form of the Edinburgh Handedness Inventory test [30]. Furthermore, eye dominance was determined using a variant of the Hole-in-the-Card test [31]. All participants were right-handers. Twenty-seven of them were right-eye dominant, and had normal/corrected-to-normal vision and normal hearing (audiometrically assessed). The local ethical committee approved the study protocol (approval number UTE-0002) and all the participants signed an informed consent form before examination.

Visuo-audio spatial attention task

Participants were instructed to perform a visuo-audio spatial task (VAST) requiring the spatial localization of auditory stimuli preceded by a visual cue (Fig. 1). The task was prepared using PsychoPy software (release 1.81). Stimuli were presented using a Dell workstation (21” Dell monitor, resolution: 1680 × 1050 pixels, display refresh rate: 60 Hz) and two speakers that were equidistant from the right side of space and provide new evidence for crossmodal links in endogenous spatial attention between vision and audition.

© 2019 The Authors. Published by Elsevier B.V. on behalf of Cairo University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Distribution of attention over space enhances our ability to select relevant objects and anticipate events [1]. The existing literature mainly focuses on visuospatial attention. However, since the external environment is monitored by multiple modalities, there is a growing need for advancing the current knowledge on spatial attention within and between other sensory modalities, including audition [2]. Covert spatial attention, metaphorically represented as a mental spotlight that enhances processing within a selected area [3], allows individuals to attend to peripheral and central locations without shifting gaze, and moving the head and/or body [4,5]. Two types of covert attention can select information and facilitate its processing [3]: ‘endogenous attention’ (i.e. the ability to voluntarily monitor information at a given location), and ‘exogenous attention’ (i.e. automatic orientation of attention to the location of a sudden stimulus).

Visuospatial attention is asymmetrically distributed, resulting in a modest but systematic leftward bias (i.e. pseudoneglect, [5]), as revealed by line bisection [6,7] and other visuospatial tasks [8,9]. The anatomical basis of this asymmetry is thought to lie in the right hemispheric dominance for this cognitive function [10], as supported by evidence that deficits in visuospatial attention are more frequent after right rather than left hemisphere lesion [11,12]. Specifically, the right parietal cortex appears to be of great importance for spatial processing of both hemifields and is thought to be the most frequently affected region in cases of left-neglect, while this is not true for the left posterior parietal cortex [13–16]. Recent studies also show larger connections of the parieto-frontal network, subserving attention in space in the right than in the left hemisphere, and a positive correlation between the degree of right lateralization of attentional networks and pseudoneglect [11].

Less is known regarding auditory spatial attention, although studies of crossmodal spatial attention indicate that auditory stimuli can play a key role in disengaging spatial attention from concurrent visual stimuli [12]. The existing evidence for possible asymmetries in deployment of auditory spatial attention is scant and suggests an opposite rightward bias [17]. Furthermore, sound sources are localized exclusively based on auditory information [18,19], but sound localization is improved by the presence of a visual cue [18,20]. Facilitation of sound localization by visual information [20] has been mainly observed in animal studies [18], although some evidence exists in humans [20]. With regard to the neuroanatomy of auditory processing, central auditory projections have a large ipsilateral component, which is absent in the visual system (characterised by a main contralateral component) [21]. Neuroimaging findings are consistent with the hypothesis of a rightward attentional bias in auditory space. Indeed, they suggest that the left hemisphere mainly responds to sounds originating from the right hemifield, while the right hemifield responds to sounds originating from both hemifields [20,22,23]. Although it is not clear which specific brain regions are mainly involved in auditory spatial attention, some data suggest that this system utilizes distinct spatial coding schemes from those used by visuospatial attention. For instance, auditory spatial attention activates the superior temporal gyrus without affecting visuotopic maps of the intraparietal sulcus [21].

The first aim of this study was to investigate putative asymmetries in deployment of covert auditory spatial attention and the influence of visual cues on its deployment. It employed a new version of the Posner paradigm [4] to investigate spatial processing of auditory targets when endogenous orientation of spatial attention was mediated by central visual cues. The second aim of the study was to investigate whether individuals’ directional biases in auditory space might possibly correlate with biases in deployment of visuospatial attention, as measured using a line bisection task. This task was selected because it is one of the most commonly used tests for the evaluation of asymmetries in deployment of visuospatial attention in both healthy and neurologically impaired individuals [24]. It was hypothesized that, if spatial attention is supramodal, then localization of auditory targets should be more accurate and/or faster in the left than in the right hemispace. In this case, individual differences in deployment of auditory spatial attention are expected to positively correlate with differences in deployment of visuospatial attention. These findings would reflect a right hemispheric dominance for auditory processing, as observed for visuospatial attention [25]. Alternatively, if spatial attention is modality-specific, then the directional bias in the auditory modality might dissociate from the one typically observed in the visual modality [26]. In this case, no significant correlations are expected between performances in auditory and visuospatial tasks. If crossmodal links exist in endogenous orientation of covert spatial attention, they might, nonetheless, involve enhancement of sound localization by central visual cues [27,28].

Subject and methods

Participants

Thirty-eight healthy volunteers (21 women) were recruited from the community through word-of-mouth and from a non-profit private university for older adults located near Messina, Italy (Third Age University). In order to test whether age influenced participants’ performance, they were categorized by age into two groups: the first group (middle-aged adults) included 19 individuals with their age ranging from 38 to 53 years (mean age: 46.94 years), and the second group (older adults) included 19 individuals with their age ranging from 61 to 71 years (mean age: 64.70 years) [29]. All the participants were naïve as to the aims and expected outcomes of the study. The participants’ hand dominance was assessed using the short form of the Edinburgh Handedness Inventory test [30]. Furthermore, eye dominance was determined using a variant of the Hole-in-the-Card test [31]. All participants were right-handers. Twenty-seven of them were right-eye dominant, and had normal/corrected-to-normal vision and normal hearing (audiometrically assessed). The local ethical committee approved the study protocol (approval number UTE-0002) and all the participants signed an informed consent form before examination.

Visuo-audio spatial attention task

Participants were instructed to perform a visuo-audio spatial task (VAST) requiring the spatial localization of auditory stimuli preceded by a visual cue (Fig. 1). The task was prepared using PsychoPy software (release 1.81). Stimuli were presented using a Dell workstation (21” Dell monitor, resolution: 1680 × 1050 pixels, display refresh rate: 60 Hz) and two speakers that were equidistant from the right side of space and provide new evidence for crossmodal links in endogenous spatial attention between vision and audition.

© 2019 The Authors. Published by Elsevier B.V. on behalf of Cairo University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
from the monitor. Subjects were comfortably seated in front of the monitor and speakers at a distance of 60 cm. Before each run, subjects underwent a training session to become familiar with the task. The visual cue was presented at the centre of the screen (exposure time = 500 ms), and consisted of a black arrow (240 × 240 pixels, 120 mm), pointing right, left, or down, against a white background. Participants were instructed to maintain their gaze on the fixation point during each trial. In a neutral, uninformative condition, a black cross (240 × 240 pixels, 120 mm, exposure time = 500 ms) was presented at the centre of the monitor (i.e. no cue was presented). All visual stimuli were presented at the centre of the screen; right and left arrows were oriented along its major axis whereas the down arrows were oriented along its minor axis. The target sound, a sine wave, consisted of a beep (Waveform Audio File format, 440 Hz, 46 dBA, duration: 2000 ms) that was produced by either the right or left speaker (lateral locations, i.e. RIGHT and LEFT conditions, respectively) or simultaneously by both speakers (central location, i.e. BOTH condition). The SOA (i.e. the amount of time between the start of cue stimulus and that of target stimulus) between visual cue and auditory target was 800 ms [32]. Participants were instructed to press a key with the index finger of their right hand as soon as possible when they localized the position of the sound source. They were instructed to use the left arrow key for left position, the right arrow key for right position, and the down arrow key (localized between right and left arrows) for the both speaker/central position. Left, Right, Both, and Neutral cues had an equal representation (i.e. 36 trials each) across the task. Combining congruency and position factors, there were a total of nine possible conditions: congruent-right (arrow pointing rightward and target on the right side), congruent-both (arrow pointing downward and target on both sides), congruent-right (arrow pointing rightward and target on the right side), incongruent-left (arrow pointing rightward or downward and target on the left side), incongruent-both (arrow pointing rightward or leftward and target on the right side), incongruent-right (arrow pointing leftward or downward and target on the right side), uninformative-left (no cue and target on the left side), uninformative-both (no cue and target on both sides), and uninformative-right (no cue and target on the right side). A total of 144 trials were administered, divided into 2 blocks.

Line bisection task

The line bisection task was adopted to measure the presence, direction, and degree of visuospatial attentional bias. The subject was instructed to mark, using a pencil, the middle of a series of lines (100 cm) long, using a white sheet of paper and oriented along its major axis (A4 format). Stimuli were centred on the participant’s sagittal midplane and presented on a table at a distance of approximately 50 cm. Each participant performed 20 trials, using the left hand in half of the trials and the right hand in the other half. The hand order was counterbalanced across subjects.

Statistical analyses

For the VAST, the participants’ accuracy (ACC) and reaction times (RT) were combined to create a new variable termed performance score (PS), as done similarly elsewhere [33–35]. To obtain this score, a two-stage procedure was adopted. First, separately for each subject, RTs were re-scaled to a value between 0 and 100, termed rapidity, according to how close they were to the fastest (100) or slowest (0) RT measured for that subject. Subsequently, the new rapidity score and accuracy rates were combined by using the following formula: performance score = 0.5 × ACC + 0.5 × rapidity. The final measure obtained encodes the ACC-RT trade-off since it assigns higher scores to both correct and fast responses while down-weighting conditions either with low accuracy or slow responses. Analyses were conducted, with PS as a dependent variable, using repeated-measures analysis of variance: congruency (three levels: congruent, incongruent, and uninformative) and position (three levels: right, left, and both) were used as within-subjects factors, and age (two levels: middle-aged and older adults) as a between-subjects factor. For this analysis, sex, ocular dominance, and education were included as covariates in the model. Greenhouse-Geisser degrees of freedom correction was used to account for potential assumption violations in the model. When necessary, Bonferroni correction was applied to post-hoc tests to obtain a global significance threshold of 0.05.

For the line bisection task, the distance (in millimetres) between the subjective and the objective centre of the line was measured. Positive and negative values were assigned to rightward and leftward bisection biases, respectively.

Repeated-measures analysis of variance was conducted using repetition (ten levels) and hand (two levels: right and left) as within-subjects factors, and age (two levels: middle-aged and older adults) as a between-subjects factor.

In addition, one-sample t-tests were performed both at group and individual levels, separately for the left, right, and both hand conditions, to test whether the mean biases were significantly different from 0. Greenhouse-Geisser degrees of freedom correction and Bonferroni correction were applied.

The Pearson correlation coefficient was used to explore possible relationships between scores obtained on the line bisection task and PS on the VAST and age (years). For this analysis, the average scores of line bisections performed using the left, the right, and both hands (AVG_LEFT, AVG_RIGHT, and AVG_BOTH, respectively) was used. Bonferroni correction was applied.

Results

Descriptive statistics are reported in Table 1.

Both position (F(1.67, 40.08) = 15.68, P < 0.001, partial η² = 0.39) and congruency (F(1.94, 46.78) = 13.89, P < 0.001, partial η² = 0.36) factors as well as their interaction (position × congruency) (F(2.96, 71.14) = 7.13, P < 0.001, partial η² = 0.22) were found to be significant. The factor age did not yield statistically significant results. Post-hoc analyses of position revealed that sounds originating from the right speaker yielded a better performance than sounds originating from the left speaker (corrected P = 0.025) and both speakers (corrected P < 0.001). The difference in the performances can be appreciated in Fig. 2a. As expected, post-hoc analyses of congruency showed that subjects had lower PS when incongruent cues were provided than when both congruent (corrected P < 0.001) and uninformative cues (corrected P = 0.024) were provided (see Fig. 2b). Post-hoc analyses of congruency × position interaction revealed that, when sounds were originating from the left speaker, subjects performed the worst during incongruent trials compared with congruent (corrected P < 0.001) and uninformative (P < 0.001) ones (see Fig. 2c).

After applying Bonferroni correction, none of the correlations reached the threshold of statistical significance (P > 0.05); therefore, these results were not included in the manuscript.

Line bisection task

Statistical analyses revealed a significant effect of the factor hand (F(1.36) = 13.713, P = 0.001, partial η² = 0.27) (Table 2). No other significant effects were found. Overall, participants showed a greater leftward bias for the left (mean = –2.543 mm, standard
deviation = 0.463 mm) than for the right (mean = 0.086 mm, standard deviation = 0.502 mm) hand condition. Results of the line bisection task are shown in Fig. 3a and b. At the group level, one sample t-test showed a leftward deviation significantly different from 0 only for the left hand (t(37) = 5.50, P < 0.000) and both hands (t(37) = 3.53, P < 0.001) conditions. At the subject level, 28 out of 38 participants (74%) showed a bias consistently different from 0, with a leftward bias when using both the right (17/28, 61%) and left (27/30, 90%) hands.

No significant correlations were observed between performances in the line bisection task and VAST.

**Discussion**

In line with recent literature [36], the findings of this study support the hypothesis that spatial attention is biased and modality-specific. Moreover, they suggest that visual cues influence the effectiveness of orientation of auditory spatial attention, deepening the knowledge of audiovisual links in endogenous covert spatial attention [37].

---

**Table 1**

Visuo-audio spatial task (VAST). Descriptive statistics (mean and standard deviation [SD]) and P-values of Performance Scores (accuracy and reaction times combined) for target position (Left, Both, and Right) and congruency (Congruent, Uninformative, and Incogruent). Asterisk indicates statistical significance at the 0.01 level. Bonferroni’s adjustment for multiple comparisons was applied.

| VAST | Congruent mean (SD) | P-value | Uninformative mean (SD) | P-value | Incongruent mean (SD) | P-value | Total |
|------|---------------------|---------|-------------------------|---------|-----------------------|---------|-------|
| Speaker position | | | | | | | |
| Left | 77.44 (14.81) | <0.001* | 67.50 (12.61) | <0.001* | 50.37 (19.24) | <0.001* | 63.83 (14.55) |
| Both | 59.12 (17.57) | 1 | 58.57 (17.57) | 0.001* | 55.90 (14.58) | 0.649 | 55.64 (15.12) |
| Right | 76.74 (16.98) | 1 | 70.60 (17.19) | 0.001* | 73.03 (17.73) | 0.08 | 71.78 (16.82) |
| Total | 69.11 (16.94) | 65.39 (8.27) | 59.84 (11.33) |  |

**Table 2**

Line bisection task (LBT). Descriptive statistics (mean and standard deviation [SD]) and P-values of Performance Scores (accuracy and reaction times combined) for hand (Right and Left) and group (Middle-aged and Older adults). Asterisk indicates statistical significance at the 0.05 level. Double asterisks indicate accuracy at the 0.01 level. Bonferroni’s adjustment for multiple comparisons was applied.

| LBT | Hand | Right mean (SD) | P-value | Left mean (SD) | Right vs. left hand (P-value) |
|-----|------|-----------------|---------|-----------------|-------------------------------|
| Middle-aged adults | Right | -0.529 (0.712) | 0.019* | -2.871 (0.659) |  |
| Older adults | Right | 0.358 (0.712) | 0.010* | -2.216 (0.659) |  |
| Middle-aged vs. Older adults (P-value) | | 0.384 | 0.486 |

---

**Fig. 1.** Schematic representation of the visuo-audio spatial task. Example of a congruent trial.

**Fig. 2.** Performance score for the visuo-audio spatial task. Estimated means for position (a) and congruency (b) factors, as well as for the interaction between congruency and position (c). Bars represent standard deviations. Asterisks indicate significant differences between sublevels at the 0.05 level. Left = sounds originating from the left speaker, right = sounds originating from the right speaker, both = sounds originating from both speakers.
In the VAST, participants manifested a rightward bias (i.e., they performed better at detecting sounds originating from the right than from the left speaker). Although evidence for asymmetries in deployment of auditory spatial attention is still controversial, the present data are in line with previous findings showing a rightward preference for auditory spatial processing [28]. It may, therefore, be speculated that a rightward attentional facilitation (i.e., involuntary transient attentional shift [37]) exists for the auditory modality. In the visuospatial task, participants as a group manifested a leftward bisec tion bias, i.e., pseudoneglect [5], when bisecting lines using the left hand, while 45% (17 out of 38) of them exhibited pseudoneglect in the right hand condition.

The leftward bias in the visual task and the rightward bias in the auditory task might be symmetrical manifestations of hemispheric dominance. Thus, these data add evidence to previous findings showing right hemispheric lateralization for visuospatial attention and are in line with the hypothesis of a left hemispheric lateralization for auditory spatial attention. Indeed, while there is strong evidence for a right lateralized fronto-parietal network for visuospatial attention [12], the existence of a left hemispheric-based neural network biasing attention to the right side of auditory space is still debated [17]. Despite the finding of modality-specific attentional bias in the present study, when spatially congruent cues were presented, participants localized the sound more easily. Facilitation of auditory targets by visual cues specifically occurred for sounds originating from the left speaker, indicating that endogenous orientation of visuospatial attention could counteract the disadvantage of the left hemispace for auditory targets. Such a result is consistent with the evidence that performance in detecting or discriminating a visual target is typically better when it appears in the attended location for visual [2] as well as auditory cues [21]. This is known as the ‘validity effect’ [4]. Interestingly, here, the ‘validity effect’ on localization of auditory stimuli was induced by visual cues, suggesting an influence of the visual on the auditory spatial coding scheme [23].

The results of this study are consistent with the evidence that the right space is prominent when gathering auditory information [17] and with the hypothesis that visual stimuli can facilitate sound localization. The rightward bias for auditory spatial attention, together with the evidence of a leftward bias for visuospatial attention [5], support the idea that auditory and visual attentional systems are governed by modality-specific processes [17]. However, the finding that visual stimuli facilitate sound localization suggests an interaction between the two systems, in line with evidence from previous studies. For instance, localization of auditory stimuli improves if their sources are visible to the subject [18]. In the ventriloquism aftereffect, observed in both humans [19] and non-human primates [18], repeated presentation of misaligned visual cues can shift spatial perception of auditory stimuli. Short-term changes in auditory space can also be induced by compression of the visual field (0.5° lenses worn for 3 days [38]). Finally, animal studies show that visual inputs modulate the oscillatory activity of the auditory cortex and enhance its response to related auditory stimuli [21]. These studies highlight the dynamic nature of auditory space, which adapts to changes in visual space. In line with the above evidence, the present data show crossmodal links in spatial attention between vision and audition.

Study limitations

The heterogeneity and small size of the sample population may be considered limitations of the present work. Another limitation that needs to be considered is that a third experimental task similar to the VAST, but without visual information, may have been very informative for assessing the role of vision in the audiostream task. Future studies manipulating and balancing auditory and visual stimuli across structurally identical paradigms may further elucidate the present findings.

Conclusions and future perspectives

Results of the present study support the existence of a facilitation effect for auditory targets originating from the right side of space. Moreover, they provide new evidence for crossmodal links in endogenous spatial attention between vision and audition. Future studies using protocols that overcome the methodologic limitations of the present work are necessary to further validate these initial findings and investigate their potential relevance for enhancement of spatial attention in healthy individuals and/or stroke patients with spatial neglect.

Conflict of interest

The authors have declared no conflict of interest.

References

[1] Mackay A, Juola JF. Are spatial and temporal attention independent? Percept Psychophys 2007:69:972–9.
[2] Hämäläinen JA, Salminen HK, Leppanen P. Basic auditory processing deficits in dyslexia: systematic review of the behavioral and event-related potential/field evidence. J Learn Disabil 2013;46:413–27.
