Auditory spatial discrimination with visual vs. auditory attentional cueing

R. Šebeňa[1], B. Hrebeňárová[2], N. Kopčo[3]

[1]Department of Psychology, Faculty of Arts, PJ Safarik University, Moyzesova 9, 040 59 Košice,
[2] Faculty of Medicine, PJ Safarik University, Trieda SNP 1, 040 11 Košice,
[3]Institute of Computer Science, Faculty of Science, Jesenná 5, 040 01 Košice
rene.sebena@gmail.com, nkopco@gmail.com

Abstract

A previous study found an enhancement of auditory spatial discrimination ability when the listener´s gaze was directed towards the auditory stimulus (Maddox et al., 2014). Here, we examined whether directing spatial auditory attention also affects this cross-modal enhancement when using realistic spatial simulation. Listeners made a judgment about the relative positions of two click-trains following a visual or auditory cue, while fixating on a neutral location. Results show that 1) subjects performed better when visual cue was used, and 2) auditory cue presented from incongruent location resulted in deteriorating performance. These results suggest a complex interaction between attentional and eye-gaze control mechanisms in auditory spatial representation.

1 Introduction

Objects and events in the real world are made up of multimodal sensory attributes.

Our nervous systems process information from different sensory modalities independently, and this information from our senses is at some point combined into one perceptual experience. Perception is a multisensory process where sensory information is integrated both within and across different sensory modalities. Some studies have shown that auditory and visual stimuli can be integrated by bimodal cells, exhibiting spatially overlapping auditory and visual receptive fields. Such neurons have been found in the early sensory cortical areas such as superior colliculus (e.g. Lakatos et al., 2007, 2008; Kayser et al., 2009) and recent study found that multi-sensory effects have been shown to occur in primary sensory areas as well (Lemus et al, 2010).

Multimodal activation has also been found in the human parietal cortex (Bremmer et al. 2001; Bushara et al. 1999, 2003; Cusack et al. 2000; Warren et al. 2002) and Intraparietal sulcus in the areas commonly referred to as LIP (lateral intraparietal sulcus bank) and MIP (medial intraparietal sulcus bank). Neuron cells in this area have been found to be sensitive to the locations of both visual and auditory stimuli (O'Dhaniel et al., 2005; Ben Hamed et al. 2001, 2002; Cohen et al. 2004; Gifford and Cohen 2004; Cohen and Andersen 2000).

Information from one sense has the potential to influence how we perceive information from another.
For example irrelevant visual stimulus can affect the detection of an auditory stimulus (Lovelace et al., 2003) as well as the perceived loudness (Odgaard et al., 2004).

Attention facilitates selection of objects, events, or spatial regions in complex scenes.
Very few studies focused on the effect of attention on sound localization. Even fewer studies looked at whether the effect is modality-dependent. Only a few previous studies asked whether directing automatic (exogenous, involuntary, stimulus-driven) or strategic (endogenous, voluntary, goal-driven) attention by an auditory cue can improve sound localization (Spence & Driver, 1994; Sach et al., 2000; Kopco et al., 2001). The result showed that cueing caused improvements in reaction times (Spence & Driver, 1994), but small (Sach et al., 2000) or no (Kopco et al., 2001) improvements in localization accuracy. Possible reasons were that tested SOAs were too short to orient attention and that auditory cue is not efficient because audition is not primarily a spatial modality.

A recent behavioral study demonstrated enhancement of auditory spatial cue discrimination ability when the listener's gaze was directed towards the auditory stimulus (Maddox et al., 2014). However, such an effect has only been demonstrated for simplistic binaural cues (interaural time and level differences). In the current behavioural study, we are expanding the findings of this paper by utilizing head related transfer functions (HRTFs) and by examining whether spatial auditory attention also affects this cross-modal enhancement.

We hypothesized that automatic attention will improve performance for valid trials and little effect or decrease in accuracy for invalid trials.

We also hypothesized that effect of automatic attention will be modality dependent. Specifficaly based on Maddox paper, we assumed that thare will be better performance in valid visual cue trials compared to valid auditory cue trials. The aim of this study is also to obtain behavioral data for electrophysiological analysis of auditory event related potential changes in cortical brain areas.

2 Methods

13 subjects (9 male, aged 20 - 38 years) participated in the two-session experiment. All participants were without any known hearing deficiencies. Some initial practice trials on each of the different experimental conditions were given prior to data collection. All provided written informed consent as approved by the PJ Safarik University in Kosice.

Auditory and visual stimuli were generated using Matlab (Mathworks, Natick, MA). The experiment was controlled using Matlab with the Psychtoolbox 3 extension (Brainard, 1997). Sound stimuli were presented using Etymotic Research (Elk Grove Village, IL) ER-1 insert headphones connected to a Datapixx system (VPixx Technologies, Saint-Bruno, QC). During the experiment subjects sat in a sound-treated booth (Eckel Laboratories). The experiment consisted of visual cue and auditory cue trials. On each trial there was target consisting of two sounds presented from slightly different locations and listeners task was to discriminate the direction of the target location change.

Visual cue and auditory cue trials

Visual cue trials consisted of a 100 ms white dot presented in horizontal plane at either 0 degrees or +/-25 degrees on the computer screen. Subjects were instructed to fix their gaze at +12.5° or -12.5° (balanced across trials) position during the whole trial, to pay attention to cue and to expect target stimulus from the same position as the presented cue. At 800 ms an auditory target was presented either about 0 or +/-25degrees laterality ipsilateral with the fixation point through insert earphones. The auditory target consisted of two, 100 ms click trains played successively. The first click train was presented at 0 or +/-25 degs, and the second train at a location slightly shifted (4.2° for central and 8.4° for lateral position); relative to the location of the first one. The subject's task was to respond, after the auditory stimuli disappeared, if the target moved left or right (using 1 or 2, respectively, on the keyboard).

Auditory cue trials were similar to the visual primer trials, consisting of an auditory target being played at 800 ms lasting for 200 ms, however the 100 ms click train at 170Hz served as an auditory cue at the beginning of each trial. Subjects were instructed to pay attention to cue, to expect target stimulus from the same position as presented cue and respond in the same manner. Schematic of the trial sequences is presented in Figure 1.

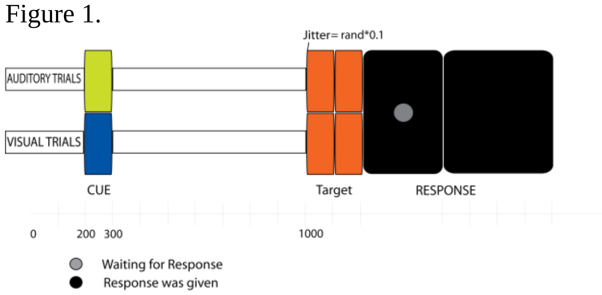

Fig.1: Schematic of the trial sequences and the two experimental conditions.

Conditions

In 'matched' trials (50%), the target stimuli were presented at the same location as the white dot. In 'mismatched' trials (50%), the target was played in the opposite location of the visual cue (0-deg cue for 25-deg target or 25-deg cue followed by a 0-deg target). In the auditory cue conditions, 'matched' trials consisted of the auditory cue and the probe occuring at the same location. 'Mismatched' auditory cue trials had the auditory cue and target occuring in opposite locations (Figure 2).

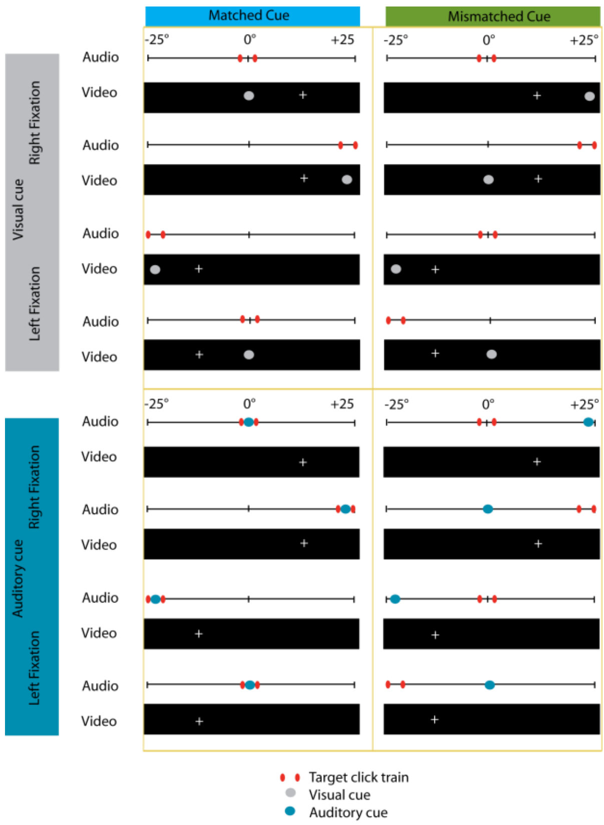

Fig. 2: Experimental conditions

Data analysis

For most of the conditions, there was a left-right symmetry in results. Therefore, data collected with fixation point on the left were mirror-flipped and combined with the data collected with fixation point on the right. And, unles stated differently, the data are presented as if the fixation point was at 12.5°.

3 Results

Data were analyzed for auditory and visual trials and for 'Match' versus 'Mismatch' conditions and for central and lateral target position. The percentage of correct responses for all participants and their means were computed (see Figures 3 and 4).

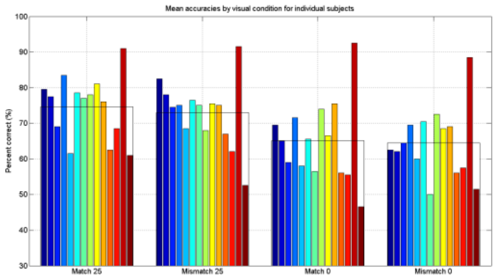

Fig. 3: Percentage of correct responses for individual subjects (color bars) and across-subject mean (open bar) for visual trials and for match, nonmatch cue (M vs. N in labels) and for position of target (0 vs. 25 in labels).

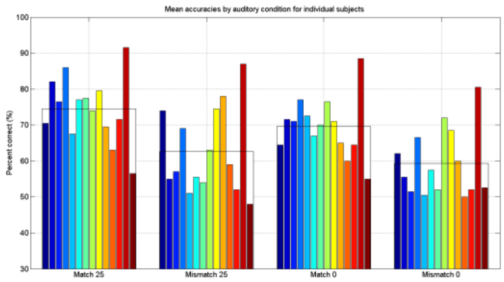

Fig. 4: Percentage of correct responses for individual subjects and their means for auditory trials and for match, nonmatch cue and for position of target

To identify significant differences between experimental conditions, a repeated measures ANOVA was performed on RAU-transformed % correct data (Tab. 1).

	df	F	pValue
Cue	1,13	7.32	0.018
Position	1,13	35.97	< 0.001
Fixation	1,13	3.93	0.069
Matching	1,13	47.13	< 0.001
Cue x Matching	1,13	12.62	0.004
Cue x Position	1,13	6.62	0.023
Fix. x Position	1,13	5.87	0.031
Cue x Fix x Match	1,13	3.84	0.072

Tab. 1: rANOVA table for testing within-subjects effects

The ANOVA showed a significant effect of the Cue type ($F(1,13) = 7.3$; $p < 0.05$), indicating that in auditory cue trials, participants performed worse than in visual cue trials. There was also a main effect of Position (($F1,13$) = 35.9; $p < 0.001$) resulting in a less pronounced decrease in performance in lateral than central position.

ANOVA also indicated a significant Cue by Position interaction ($F(1,13) = 12.6$; $p < 0.01$) (Figure 5). In lateral position performance was better for visual cue compared to auditory cue.

As hypothesized we also found a main effect of Matching (($F1,13$) = 47.13; $p < 0.001$) resulting in a significantly better performance for matched than mismatched trials. There was also a significant Cue x Matching interaction. The difference between 'Match' and 'Nonmatch' condition for visual cue trials was not significant ($F(1,13) = 1.5$; $p = 0.24$), but for auditory cue trials this difference was significant ($F(1,13) = 28.9$; $p < 0.001$), the performance was better for 'Match' experimental conditions (Figure 6).

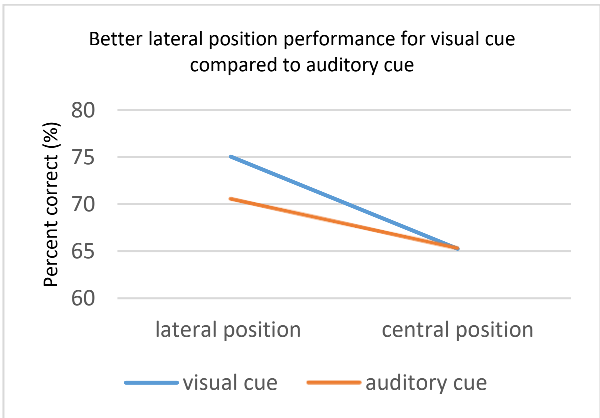

Fig. 5: Cue by target Position interaction

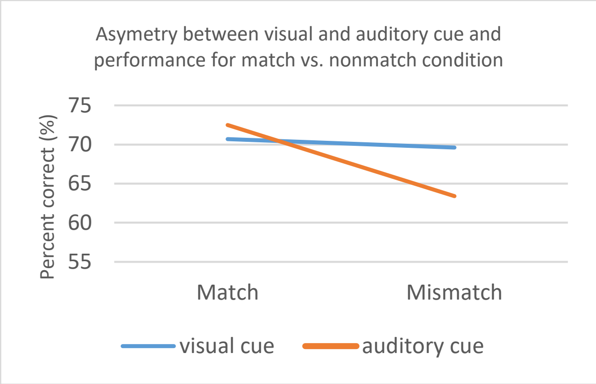

Fig. 6: The difference between 'Match' and 'Nonmatch' condition for visual and auditory cue trials.

Finally we found no significant main effect of Fixation ($F(1,13) = 3.93$; $p = 0.07$). The performance was relatively the same for left fixation and right fixation. On the other hand there was a significant Fixation x Position interaction. In right fixation performance was better for lateral presented target stimuli (Figure 7). Cue by Fixation by Matching interaction was not significant.

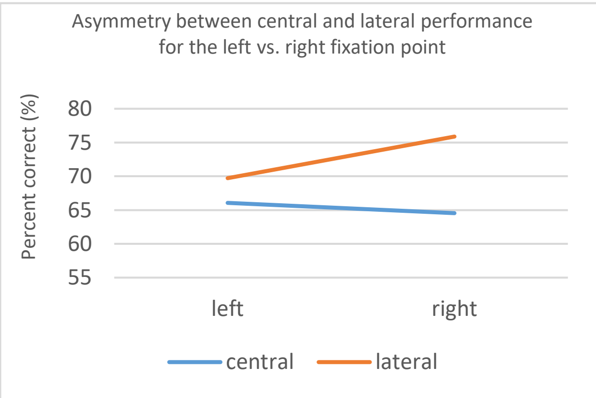

Fig. 7: Fixation by Position interaction

4. Discussion and Conclusions

Performance was better when visual cue was used
We aimed to compare subjects performance when visual cue versus auditory cue was used. Based on our findings it can be concluded that subjects performed better when visual cue was used. This is in line with some other studies showing that visual cues help auditory perception by guiding attention to discriminate target either by enhancing sounds near the threshold of audibility when the target is energetically masked or by enhancing segregation when it is difficult to direct selective attention to the target (Varghese et al, 2012). It seems that visual cues can provide perceptual benefits helping listeners focus selective attention on the target.

Auditory cue presented from incongruent location resulted in deteriorating performance
In our experiment all trials had either ´Matched´ or ´Nonmatched´ cues.We found that only for auditory cue trials performance was better when cue was matching the target position. This result is surprising, partly in contrast to the previous experiments which shown intelligibility and discrimination benefits of knowing where to listen (Best et al, 2007; Maddox et al., 2014). Those gains may come from facilitated selective attention (Mesgarani, Chang, 2012).

This finding is in opposite with Maddox paper, who found for directional auditory primers no benefit for performance compared to uniformative and better performance in spatial discrimination visual directional trials than in visual uninformative trials for ILD at both the center and side positions and for ITD only when stimulus was located on the side.

An important difference between this study and the previous studies is that here only automatic spatial auditory attention was examined since the cue was only informative at 50% of trials, thus making it unlikely that the subjects would use it to direct their strategic attention. However, it is possible that some strategical attention was engaged. Additional experiments need to be performed to distinguish between these two options.

Better performance in lateral than central position
We also found significant main effect of position with better performance in lateral than central position.
This finding is in opposite with Maddox who found center performance better than side performance.
In our case lateral position was much easier to discriminate due to more spatial difference between two target stimuli (8.4°) compared to 4.2° difference for central stimuli as is obvious from initial practice no cue trials on each experimental conditions which were given prior to data collection (Figure 8).

We also observed an asymmetry between central and lateral performance for the left vs. right fixation point (Fig. 7). This asymetry is likely due to the use of non-individualized HRTFs which might have been better matched to the individual subjects' HRTFs on the righthand side compared to the lefthand side. However, as shown in Fig. 8, which shows the nocue baseline performance measured prior to the experiment, this performance was well matched across the locations in the experiment.

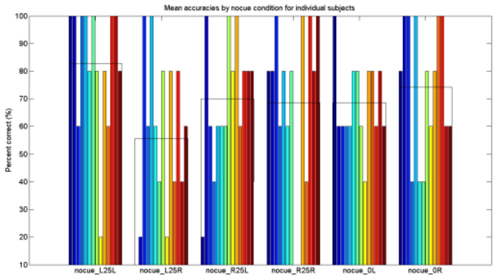

Fig. 8: No cue practice trials on 6 experimental conditions. Each group of bars corresponds to performance of individual subjects (color bars) and across-subject mean (open bar) for one combination of target positoin (L25, R25 or 0) and direction of shift.

In lateral position performance was better for visual cue compared to auditory cue

Maddox study confirmed that gazing leftward would shift the receptive field to the left, resulting in better discrimination of the left-lateralized sounds (Maddox et al., 2014). It is not clear how to relate this result to the current results, given that the subjects were specifically instructed not to move their eyes in the current experiments (and we monitored eye posistion using electro-oculography). It is possible that the presentation of visual cue or auditory cue induced an automatic orienting response or response planning, which then affected performance, in particular for the non-matching cues in the auditory condition.

Future studies

In the future, with regards to these studied experimental conditions and data, we plan to conduct d-prime analysis and analysis of auditory event related potential changes in cortical brain areas.

Acknowledgement

This work was supported by Slovak Research and Development Agency under grant project APVV-0452-12 Spatial Attention and Listening in Complex Environments. Barbara Shinn-Cunningham, Lenny Verghese and Hannah Goldberg provided an early version of the experimental code for this study.

References

Best, V., Ozmeral, E.J., and Shinn-Cunningham, B.G. (2007). Visually guided attention enhances target identification in a complex auditory scene. *J. Assoc. Res. Otolaryngol*. 8: 294–304

Ben Hamed, S., Duhamel, J.R., Bremmer, F., and Graf, W. (2001). Representation of the visual field in the lateral intraparietal area of macaque monkeys: a quantitative receptive field analysis. *Experimental Brain Research.* 140: 127–144

Ben Hamed, S., Duhamel, J.R., Bremmer, F., and Graf W. (2002). Visual receptive field modulation in the lateral intraparietal area during attentive fixation and free gaze. *Cerebral Cortex.* 12: 234–245

Bremmer, F., Schlack, A., Duhamel, J.R., Graf, W., and Fink, G.R. (2001). Space coding in primate posterior parietal cortex. *Neuroimage.* 14: 46–51

Bushara, K.O., Hanakawa, T., Immisch, I., Toma, K., Kansaku, K., and Hallett, M. (2003). Neural correlates of cross-modal binding. *Nature Neuroscience.* 6: 190–195

Bushara, K.O, Weeks, R.A., Ishii, K., Catalan, M.J., Tian, B., Rauschecker, J.P., and Hallett, M. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nature Neuroscience.* 2: 759–766

Cohen, Y.E., Cohen, I.S., and Gifford, G.W. (2004). Modulation of LIP activity by predictive auditory and visual cues. *Cerebral Cortex* 14: 1287–1301

Cohen, Y. and Andersen, R. (2000). Reaches to sounds encoded in an eye-centered reference frame. *Neuron* 27: 647–652

Cusack, R., Carlyon, R.P., and Robertson, I.H. (2000). Neglect between but not within auditory objects. *Journal of Cognitive Neuroscience.* 12: 1056–1065

Gifford, G.W. and Cohen, Y.E. (2004). Effect of a central fixation light on auditory spatial responses in area LIP. *Journal of Neurophysiology*. 91: 2929–2933

Kayser, C., Petkov, C.I., Logothetis, N.K., (2009). Multisensory interactions in primary auditory cortex: fMRI and electrophysiology. *Hearing Research*, doi:10.1016/j.heares.2009.02.011.

Kopco, N., Ler, A., and Shinn-Cunningham, B. (2001). "Effect of auditory cuing on azimuthal localization accuracy," *Journal of the Acoustical Society of America*. 109, 2377

Lakatos, P., Chen, C.M., O'Connel, M.N., Mills, A., Schroeder, C.E., (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53: 279–292

Lemus, L., Hernández, A., Luna, R., Zainos, A., Romo, R. (2010). Do sensory cortices process more than one sensory modality during perceptual judgments? *Neuron.* 67:335-348

Lovelace, C.T., Stein, B.E., Wallace, M.T. (2003). An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Brain Research Cognitive Brain Research*. 17:447–453

Maddox, R. K., Pospisil, D. A., Stecker, G. C., and Lee, A. K. C. (2014). Directing eye gaze enhances auditory spatial cue discrimination. *Current Biology*. 24: 748–752

Mesgarani, N., and Chang, E.F. (2012). Selective cortical representation of attended speaker in multi-talker speech perception. *Nature.* 485: 233–236

O'Dhaniel, A., Mullette-Gillman, Yale, E. Cohen, Y.E., and Groh, J.M. (2005). Eye-Centered, Head-Centered, and Complex Coding of Visual and Auditory Targets in the Intraparietal Sulcus, *Journal of Neurophysiology.* 94: 2331–2352

Odgaard, E.C., Arieh, Y., Marks, L.E. (2004). Brighter noise: sensory enhancement of perceived loudness by concurrent visual stimulation. *Cognitive, Affective & Behavioral Neuroscience*. 4:127–132

Sach, A.J., Hill, N.I., and Bailey, P.J. (2000). Auditory spatial attention using interaural time differences. *Journal of Experimental Psychology: Human Perception and Performance*. 26(2):717-729

Spence, C.J. and Driver. J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance*. 20(3): 555-574

Varghese, L.A., Ozmeral, E.J., Best, V., Shinn-Cunningham, B.G. (2012). How visual cues for when to listen aid selective auditory attention. *J. Assoc. Res. Otolaryngol.* 13:359–368

Warren, J.D., Zielinski, B.A., Green, G.G., Rauschecker JP, and Griffiths, T.D. (2002). Perception of sound-source motion by the human brain. *Neuron* 34: 139-148