

Long-range coupling of prefrontal cortex and visual (MT) or polysensory (STP) cortical areas in motion perception

L.M. Vaina^{1,2}, F. Calabro², F.H. Lin¹, and M.S. Hämäläinen¹

¹ Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, U.S.A.

²Brain and Vision Research laboratory/Biomedical Engineering, Boston University, Boston, U.S.A.

Abstract— To investigate how, where and when moving auditory cues interact with the perception of object-motion during self-motion, we conducted psychophysical, MEG, and fMRI experiments in which the subjects viewed nine textured objects during simulated forward self-motion. On each trial, one object was randomly assigned its own looming motion within the scene. Subjects reported which of four labeled objects had independent motion within the scene in two conditions: (1) visual information only and (2) with additional moving-auditory cue.

In MEG, comparison of the two conditions showed: (i) MT activity is similar across conditions, (ii) late after the stimulus presentation there is additional activity in the auditory cue condition ventral to MT, (iii) with the auditory cue, the right auditory cortex (AC) shows early activity together with STS, (iv) these two activities have different time courses and the STS signals occur later in the epoch together with frontal activity in the right hemisphere, (v) for the visual-only condition activity in PPC (posterior parietal cortex) is stronger than in the auditory-cue condition. fMRI conducted for visual-only condition reveals activations in a network of parietal and frontal areas and in MT.

In addition, Dynamic Granger Causality analysis showed for auditory cues a strong connection of the AC with STP but not with MT suggesting binding of visual and auditory information at STP. Also, while in the visual-only condition PFC is connected with MT, in the auditory-cue condition PFC is connected to STP (superior temporal polysensory) area.

These results indicate that PFC allocates attention to the “object” as a whole, in STP to a moving visual-auditory object, and in MT to a moving visual object.

Keywords— visual search, auditory facilitation, MEG, dynamic Granger causality

I. INTRODUCTION

Current models of brain organization suggest a hierarchical progression of processing from simple to complex aspects of unimodal stimuli within visual and auditory systems [1, 2]. However, recent studies of connection patterns and physiological data suggest that only a small fraction of the cortex is truly unimodal. Convergence between sensory systems occurs even at the early putative unisensory processing stages and at extremely short post-stimulus latencies, arising through either feedback or feed-forward circuits. In

multi-sensory processing [3, 4] recruitment of the early sensory areas into the cognitive-perceptual processes can be partly attributed to feedback mechanisms. This calls for investigating the auditory and visual processing hierarchies in both temporal and spatial dimensions. This is what we have addressed in this study through MEG in healthy human observers performing search tasks in a 3D motion field.

Motion is pervasive in our everyday environments, and perceiving the direction and speed of our own and object motion is essential for functioning effectively, such as walking, coping with traffic, or even detecting objects on our path. These are fundamental everyday tasks which are often helped by additional cues, especially auditory. The combined use of psychophysics and brain imaging techniques is essential to discover the underlying principles governing cross-modal integration during the perception of motion. A grasp of these fundamental principles is not only critical for understanding the function of the healthy brain but, importantly, brings the promise of designing effective strategies to help patients recover and deal with their challenging everyday life environments.

Employing unimodal (visual) and cross-modal (visual and auditory) tasks we investigated psychophysically whether the auditory cues enhance perception of object motion in a search task simulating a forward moving observer searching among a set of moving objects an object moving in depth. Our fundamental hypothesis is that simultaneous congruent information from another sensory modality may facilitate the perception of object motion. Using anatomically constrained MEG in conjunction with dynamic time and frequency domain Granger causality we investigated the direction and dynamics of connectivity between the functional areas involved in selective, unimodal visual attention and cross-modal attention and facilitation for visual perception [5].

II. MATERIALS AND METHODS

A. Subjects

Three experimental conditions were used: vision only (in both psychophysics and MEG), vision and non-spatially localized auditory cue (psychophysics and MEG) and visual

with spatially localized auditory cues (psychophysics only). Subjects participating in MEG also had a high resolution structural MRI and fMRI for localizing the MT+ region.

Eleven subjects, recruited from graduate and undergraduate college students whose age ranged between 19-29 years (7 women, 4 men) participated in the first two experimental conditions and 9 subjects (ages ranging between 20-28 years; 5 women, 4 men) participated in the third experimental condition. All subjects gave Informed Consent according to the Boston University and Martinos Center for Biomedical Imaging Human Subjects Committees.

None of the participants had any history of neurological, psychiatric or ophthalmological/neuro-ophthalmological diseases, or were on any medication, they were all right handed and had normal to corrected vision. All the subjects who participated in the imaging experiments satisfied the inclusion criteria for participation in imaging, according to the Martinos Center for Biomedical Imaging requirements.

B. Psychophysics

To address whether moving auditory cues enhance the perception of object-motion during self-motion, subjects viewed nine computer generated textured spheres (1.5° in diameter) during 1sec simulated forward self-motion. On each trial, one of the spheres was randomly assigned its own looming motion (forward or backward) within the scene. In the last frame of the motion four of the spheres were labeled from 1-4. Subjects were asked to report which of four labelled spheres had independent motion within the scene (4AFC task), with (1) visual information only, (2) an auditory cue localized to the center of the screen, either moving in the same direction as the target (“Auditory moving”) or stationary (“Auditory static”), or (3) an auditory stimulus co-localized to the target object (moving or stationary).

The visual stimuli were displayed on a calibrated 23” LCD screen. Auditory stimuli were presented via Bose QuietComfort noise canceling headphones. Subjects gave responses via a USB number pad and were instructed to maintain fixation throughout the testing block.

C. MEG and MRI acquisition

MEG data were acquired at the Martinos Center using a 306-channel MEG (VectorView, Elekta-Neuromag, Helsinki, Finland) and in a magnetically shielded room. The MEG signals were band-pass filtered to the frequency range 0.03–200 Hz and digitized at 600 samples/s. Electrooculogram (EOG) were recorded to monitor eye-movement artifacts. All epochs exceeding 150 μ V or 3000 fT/cm at any EOG or MEG channel, respectively, were automatically

discarded from the offline averages computed after filtering the data to 0.1...40 Hz.

For stimulus presentation, we employed a DLP projector and a back-projection screen, plastic tubes connected to frequency-compensated loudspeakers outside the magnetically shielded room (ADU1c, Unides Design, Helsinki). Responses were recorded with a fiber-optic response pad.

We also acquired a high-resolution 3D T1-weighted structural MRI using a 32 channel phase array head coil in a 3T scanner (Siemens-Trio, Erlangen, Germany). Using standard functional localizers in fMRI we localized the area hMT+. fMRI data was analyzed using Freesurfer.

D. MEG source estimates

Cortically-constrained depth-weighted l2 minimum-norm estimates were computed using the MNE software (<http://www.nmr.mgh.harvard.edu/martinos/userInfo/data/sofMNE.php>). The sources were restricted to the white matter surface extracted with FreeSurfer software (<http://surfer.nmr.mgh.harvard.edu/>). The dense triangulation of the cortex was decimated to about 9000 sources per hemisphere resulting in an average distance of 4 mm between adjacent sources. The orientations of the sources were approximately constrained to the cortical surface normal direction using the loose orientation constraint approach. For computation of the Dynamic Granger Causality measures of association between cortical regions, the source waveforms were extracted from each anatomical region of interest discussed in Section III.

E. Dynamic Granger Causality (DGC)

We estimated causal influences between the anatomical areas of interest delineated in MEG for each of the search experiment conditions. We obtained time series from regions-of-interest (ROIs) after MEG source localization as discussed above. For each pair of time series from two ROIs (“X” and “Y”), we calculated the Granger causality $G_{X \rightarrow Y}$ and $G_{Y \rightarrow X}$. We used the Schwarz’s Bayesian Criterion to estimate the order of the AR model and stepwise least squares to estimate model parameters [6, 7]. Granger causality also has a frequency-domain representation [8]:

$$G_{X \rightarrow Y}(f) = -\ln(1 - (\Sigma_{XX} - \Sigma_{XY}^2 / \Sigma_{YY}^2) |H_{YX}(f)|^2 / S_{YY}(f)),$$

where the subscripts are entries in the matrix, $S(f) = H(f)\Sigma H^H(f)$ is the spectral matrix,

$H(f) = (\sum_{k=0}^p A_k \exp(j2\pi kf))^{-1}$ is the system transfer function

computed from the coefficient matrix A_k , and Σ is the noise covariance matrix of $[\varepsilon_{y,x}(t) \ \varepsilon_{x,x}(t)]^T$ in the AR model.

To develop a DGC model, similar to the published study by Lin et al. [5], we used a time-window approach in both the time and frequency domains. Granger causality depends on the estimated optimal order of the AR model. This means that the DGC model must incorporate simultaneous estimates of both the AR model order and the window length. For this we will use the SURE metric [5] obtained by formulating the time series as a dynamic AR model. Thus if we use a univariate time series:

$y(t) = \sum_{k=1}^p \alpha_k^w(t)y(t-k) + \varepsilon_y^w(t)$, where $\alpha_k^w(t)$ is the time-varying AR model estimated from the least squares criterion: $\alpha_k^w(t) = \arg \min_w \sum (y(t) - y(t-k))^2$ and w denotes

the length of the temporal window, and p the order of the AR model. The residual variance of this dynamic AR model fit is estimated as $\varepsilon_{p,w}^2 = \sum_{t=1}^n (\sum_{k=1}^p \alpha_k^w(t)y(t-k) - y(t))^2 / T$ and the SURE

metric can be written as $SURE(p,w) = \log \varepsilon_{p,w}^2 + \frac{2p\alpha}{w}$, where

$\alpha = \left(\sum_{t=1}^T y_{t,q}^T M^{-1}(t) y_{t,p} \right) / T$ is a modification term and $M(t) = \left(\sum_{q=0}^{p-1} y_{t-q,p} y_{t-q,p}^T \right) / w$,

with $y_{t-q,p} = [y(t-q-1), y(t-q-2), \dots, y(t-q-p)]^T$. We designate the window length corresponding to minimal SURE metric as optimal. *Since our causality model takes into account both time and frequency domains it would help elucidate the dynamics of cortical spatiotemporal orchestration.*

We used this approach to detect modulation of effective connectivity during visual and visual-auditory search in order to reveal the spatiotemporal cortical network involved in these tasks. We were particularly interested to determine whether the correlations between auditory and visual cortices are mediated by the same connections.

III. RESULTS

A. Psychophysics

In the ‘‘auditory off’’ condition, performance depended on the speed and direction of the object: in both experimental conditions subjects were significantly better when the target object moved towards them (positive speeds) than away, when retinal velocities were larger, suggesting that relative and retinal speed affected performance.

Non-spatial and spatially localized auditory cues were either stationary (25% of trials) or moving in the same direction as the target. With the non-spatial cue (centered at

fixation), subjects showed no statistical difference between the auditory off and auditory static conditions ($p=0.2$, $t=-1.38$, $df=8$, Fig 1 left). For the moving auditory cue, there was a modest but statistically significant improvement of 3.3% compared to the static auditory cue ($p=0.003$, $t=4.18$, $df=8$, min 1.2% and max 8.1, Fig 1 right).

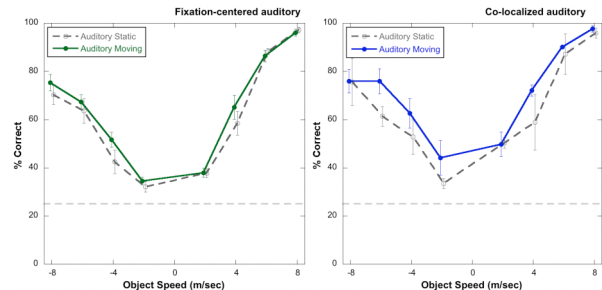


Fig. 1 Results from healthy controls ($n=11$) for the detection of object motion during forward self-motion. Performance significantly improved for an auditory cue moving in the same direction as the target compared to an uninformative, static auditory cue for both fixation-centered auditory cues ($p=0.003$, left), and spatially localized auditory cues ($p=0.04$, right). The co-localized cue was nearly twice as effective as fixation-centered at enhancing detection relative to a static auditory condition ($p=0.06$, $t=2.1$, $df=10$). Dashed line is chance performance (25%), and error bars are s.e.m.

B. Dynamic Granger Causality (DGC)

Pair-wise DGC [5] was computed using the MEG source time courses. Nine regions-of-interest (ROIs) were identified: the left and right auditory cortices, left and right medial prefrontal cortices (MPFC), left and right MT, right posterior IPS complex, right post-central gyrus, and right superior temporal poly-sensory area (STP). In each condition, we calculated the average time courses across source locations in each ROI for each trial separately. Randomly selected 50% of trials were averaged to generate one evoked time course for each ROI. We used a 100 ms time window and 5-th order autoregressive model to calculate dynamic Granger causality using these evoked time courses. The random selection across trials was re-peated 30 times. Statistical significance of each causality estimate was evaluated by (i) averaging all DGC over 30 repetitions, (ii) subtracting the average baseline (-0.2 s to 0 s) DGC, and (iii) dividing each time point with the standard deviation of the DGC in the baseline to transform the DGC time courses calculated from each ROI pair to a t-statistic.

In Figs 2 and 3 we show differences in connectivity between the *Visual-auditory* and *Visual-only* tasks. In Fig 2 we illustrate that in the presence of the auditory cue, the auditory cortex and the polymodal area STP are strongly connected very early on, connection which is not present in the *Visual-only* task (Fig 2D).

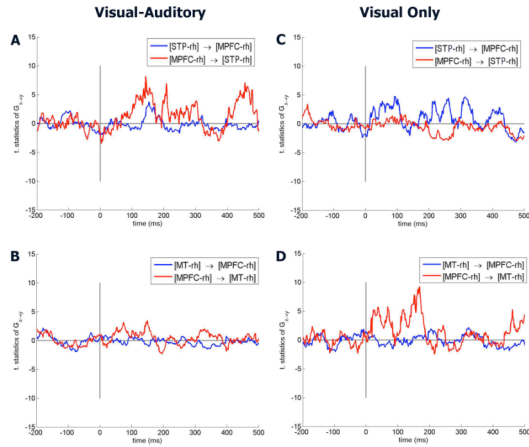


Fig. 2 Feedback connectivity from PFC is directed to MT+ in the visual-only condition (D), but moves to STP when the target is multi-sensory (A).

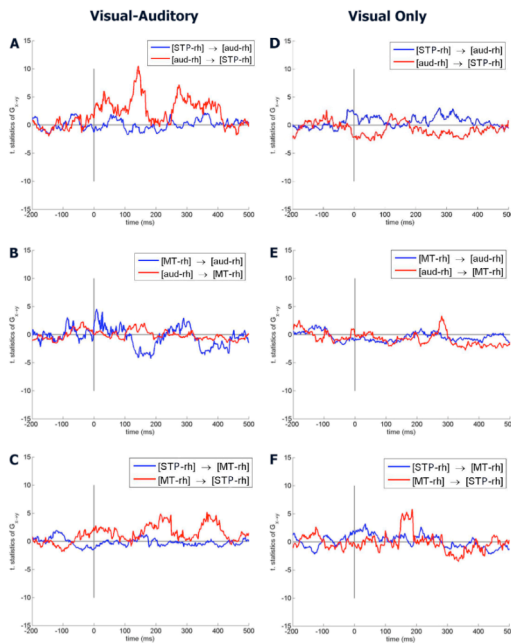


Fig. 3 Dynamic connectivity shows for an illustrative single subject. t-statistics show that both auditory cortex (top, visual-auditory condition only) and MT+ (bottom) input to STP, but there is no direct connectivity between auditory cortex and MT+ (middle)

Figure 3D shows that in the visual only task, the prefrontal cortex provides a very early input to MT+ and a 100 delay input to STP in the visual-auditory tasks, which suggests a top-down influence. It is interesting that the auditory cortex and MT (significantly involved in the motion task) are not connected (Fig. 3) even in the presence of the auditory cue.

IV. CONCLUSIONS

These results suggest that the binding of visual and auditory information may occur in STP, and this area is probably involved in the multimodal task. The visual information from MT appears to arrive to STP later than the auditory information. Figure 2, addresses the question of attention modulation by the prefrontal cortex. Fig 2A shows that the PFC is strongly connected with STP in the Visual-auditory task, where the target is a multimodal object. PFC will modulate STP during the selection of the target. However, this connectivity is not established in the visual-only task where PFC is directly and strongly connected with area MT+ whose neurons are significantly involved in the perception of the stimulus and detecting the target-object. In the visual-auditory task MT+ is not connected with PFC, nor with the auditory cortex. These preliminary results show that the prefrontal cortex allocates attention to the “object” as a whole, in STP to a moving visual-auditory object, and in MT to a moving visual object.

ACKNOWLEDGMENTS

This work was supported by NIH grant R01NS064100 to LMV, and the NCRRT supported Center for Functional Neuroimaging Technologies P41 RR14075 (MSH, FHL).

REFERENCES

1. Felleman, D.J. and D.C. Van Essen, Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex*, 1991. 1(1): p. 1-47.
2. Rauschecker, J.P., et al., Serial and parallel processing in rhesus monkey auditory cortex. *J Comp Neurol*, 1997. 382(1): p. 89-103.
3. Schroeder, C.E., A.D. Mehta, and J.J. Foxe, Determinants and mechanisms of attentional modulation of neural processing. *Front Biosci*, 2001. 6: p. D672-84.
4. Bullier, J., Feedback connections and conscious vision. *Trends Cogn Sci*, 2001. 5(9): p. 369-370.
5. Lin, F.H., et al., Dynamic Granger-Geweke causality modeling with application to interictal spike propagation. *Hum Brain Mapp*, 2009. 30(6): p. 1877-86.
6. Neumaier, A. and T. Schneider, Estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Trans. Math. Softw.*, 2001. 27: p. 27-57.
7. Schneider, T. and A. Neumaier, Algorithm 808: ARfit - A Matlab package for the estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Trans. Math. Softw.*, 2001. 27: p. 58-65.
8. Brovelli, A., et al., Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality. *Proc Natl Acad Sci U S A*, 2004. 101(26): p. 9849-54.

Author: Lucia Maria Vaina
 Institute: Boston University-Biomedical Engineering
 Street: 44 Cummington St
 City: Boston, Ma 02215
 Country: USA
 Email: vaina@bu.edu

