

Temporal Characteristics of Audiovisual Information Processing

Galit Fuhrmann Alpert,^{1*} Grit Hein,^{1*} Nancy Tsai,¹ Marcus J. Naumer,³ and Robert T. Knight^{1,2}

¹Helen Wills Neuroscience Institute and ²Psychology Department, University of California, Berkeley, Berkeley, California 94720, and ³Institute of Medical Psychology, J. W. Goethe University, 60528 Frankfurt am Main, Germany

In complex natural environments, auditory and visual information often have to be processed simultaneously. Previous functional magnetic resonance imaging (fMRI) studies focused on the spatial localization of brain areas involved in audiovisual (AV) information processing, but the temporal characteristics of AV information flow in these regions remained unclear. In this study, we used fMRI and a novel information–theoretic approach to study the flow of AV sensory information. Subjects passively perceived sounds and images of objects presented either alone or simultaneously. Applying the measure of mutual information, we computed for each voxel the latency in which the blood oxygenation level-dependent signal had the highest information content about the preceding stimulus. The results indicate that, after AV stimulation, the earliest informative activity occurs in right Heschl's gyrus, left primary visual cortex, and the posterior portion of the superior temporal gyrus, which is known as a region involved in object-related AV integration. Informative activity in the anterior portion of superior temporal gyrus, middle temporal gyrus, right occipital cortex, and inferior frontal cortex was found at a later latency. Moreover, AV presentation resulted in shorter latencies in multiple cortical areas compared with isolated auditory or visual presentation. The results provide evidence for bottom-up processing from primary sensory areas into higher association areas during AV integration in humans and suggest that AV presentation shortens processing time in early sensory cortices.

Key words: auditory; fMRI; information theory; multisensory; prefrontal cortex; visual

Introduction

Simultaneous processing of auditory and visual information is the basis for appropriate behavior in our complex environment. Previous neuroimaging studies focused on spatially localizing brain regions involved in audiovisual (AV) processing. Various neocortical regions are affected by both auditory and visual inputs (Ghazanfar and Schroeder, 2006), including inferior frontal cortex (IFC) (Belardinelli et al., 2004; Taylor et al., 2006; Hein et al., 2007), the superior temporal sulcus (STS) (Calvert et al., 2000; Beauchamp et al., 2004; Stevenson et al., 2007), and visual (Belardinelli et al., 2004; Meienbrock et al., 2007) and auditory (van Atteveldt et al., 2004, 2007; Hein et al., 2007) areas. However, temporal aspects of information flow across these different brain regions after AV stimulation were not addressed.

Temporal characteristics of AV interactions in the human brain have been explored to some extent by electroencephalography (EEG) studies (Girard and Peronnet, 1999; Fort et al., 2002; Molholm et al., 2004; Teder-Sälejärvi et al., 2005; van Wassen-

hove et al., 2005; Senkowski et al., 2007; Talsma et al., 2007). For example, Fort et al. (2002) investigated AV object detection with scalp current densities and dipole modeling. Their results support AV interaction effects in deep brain structures, possibly the superior colliculus, and frontotemporal regions but earliest in occipitoparietal visual areas. In line with this, several other studies showed that AV interaction effects already occur in early potentials that are probably generated in primary or secondary visual and auditory regions (Giard and Peronnet, 1999; Molholm et al., 2004; Teder-Sälejärvi et al., 2005; van Wassenhove et al., 2005; Senkowski et al., 2007; Talsma et al., 2007; Karns and Knight, 2007). Such results provide insights to the latency of interactions between auditory and visual inputs but do not unravel the flow of AV information. Moreover, they do not permit analysis of AV information processing in specific visual, auditory, or frontotemporal subregions because of the limited spatial resolution of the EEG signal.

In this paper, we use functional magnetic resonance imaging (fMRI) data to explore how AV-related activity propagates across different brain regions during processing of AV stimuli. This goal is addressed by applying a novel information–theoretic approach, which in addition to spatially localizing AV-related activity also provides information about the temporal characteristics of AV information processing.

Materials and Methods

Participants. We report data from 18 subjects (seven female; mean age, 29.8 years; range of 23–41 years) who had normal or corrected-to-

Received Nov. 13, 2007; revised Feb. 29, 2008; accepted March 23, 2008.

This work was supported by German Research Foundation Grant HE 4566/1-1 (G.H.), National Institute of Neurological Disorders and Stroke Grant NS21135, and the University of California President's Postdoctoral Fellowship (G.F.A.). We thank Dr. Daniele Paserman for statistical analysis consulting, Oliver Doehrmann for data acquisition and analysis support, Petra Janson for help preparing the visual stimuli, Tim Wallenhorst for data acquisition support, and Clay Clayworth for help in preparing the figures.

*G.F.A. and G.H. contributed equally to this work.

Correspondence should be addressed to Galit Fuhrmann Alpert, Helen Wills Neuroscience Institute, University of California, Berkeley, 132 Barker Hall, Berkeley, CA 94720. E-mail: galit@berkeley.edu.

DOI:10.1523/JNEUROSCI.5039-07.2008

Copyright © 2008 Society for Neuroscience 0270-6474/08/285344-06\$15.00/0

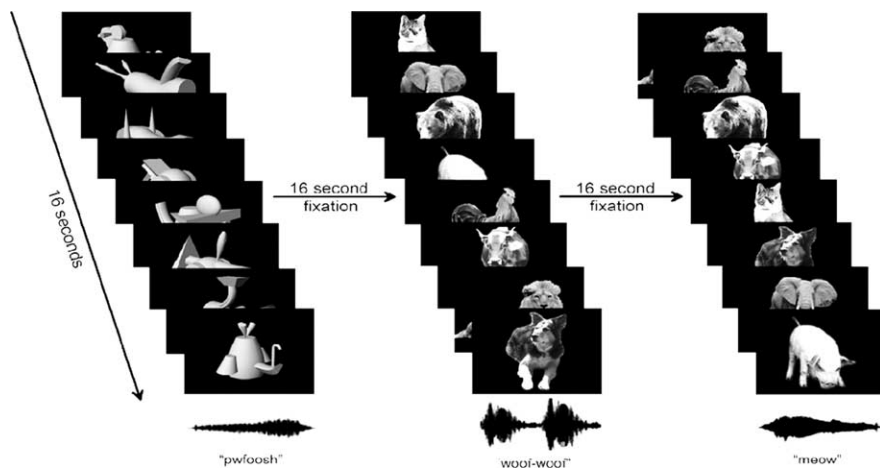


Figure 1. Experimental protocol. AV stimuli of complex sounds (representative examples) were presented simultaneously with images of three-dimensional objects. Besides AV stimulation, there were unimodal VIS and AUD stimulation blocks.

normal vision and hearing and signed an informed consent in accordance with the declaration of Helsinki.

Stimuli and procedure. Subjects perceived sounds and images of three-dimensional objects, separately (AUD and VIS, respectively) or simultaneously (AV), involving animal sounds and images, abstract images (<http://alpha.cog.brown.edu:8200/stimuli/novel-objects/fribbles.zip> view), and abstract sounds, created by distortion (played backwards and with underwater effect) of the animal sounds (Fig. 1) (for details, see Hein et al., 2007). Inspired by previous studies (Calvert et al., 2000; van Atteveldt et al., 2004, 2007), we choose a passive paradigm to investigate neural correlates of “pure” AV information processing, independently of task effects.

Stimuli were presented in blocks of eight events, each lasting 2000 ms, with a fixation cross constantly present, followed by a fixation period of equal length. Sixteen seconds of fixation were added at the beginning and end of each block. The experiment had a total of 320 AV trials, 160 AUD trials, and 160 VIS trials, presented in five runs.

MRI data acquisition. fMRI data were acquired on a 3 tesla Siemens (Erlangen, Germany) scanner. A gradient-recalled echo-planar-imaging sequence was used with the following parameters: 34 slices; repetition time (TR), 2000 ms; echo time (TE), 30 ms; field of view, 192 mm; in-plane resolution, $3 \times 3 \text{ mm}^2$; slice thickness, 3 mm; gap thickness, 0.3 mm. A magnetization-prepared rapid-acquisition gradient-echo sequence was used (TR, 2300 ms; TE, 3.49 ms; flipped angle, 12° , matrix, 256×256 ; voxel size, $1.0 \times 1.0 \times 1.0 \text{ mm}^3$) for anatomical imaging.

Preprocessing. The functional data were preprocessed using BrainVoyager QX (version 1.7; Brain Innovation, Maastricht, The Netherlands), including the following steps: (1) three-dimensional motion correction, (2) linear-trend removal and temporal high-pass filtering at 0.0054 Hz, (3) slice-scan-time correction with sinc interpolation, and (4) spatial smoothing using a Gaussian filter of 8 mm (full-width half-maximum).

Event-related information-theoretic analysis. To study event-related blood oxygenation level-dependent (BOLD) activation, we used the information-theoretic approach and related equations described in detail in Methods of Fuhrmann Alpert et al. (2007). In particular, we used the measure of mutual information (MI) to quantify the reduction in uncertainty about a reference task condition (REF) (e.g., the AV condition), achieved by knowing the magnitude of BOLD response a fixed time delay (dt) later. All time series of BOLD signals are z -normalized to have 0 mean and unit variance. Therefore, the analysis focuses on changes of responses relative to baseline activity.

If $H(\text{BOLD})$ is the unconditional entropy of BOLD magnitudes, and $H(\text{BOLD}|\text{REF})$ is the conditional entropy of BOLD magnitudes given a preceding task condition (REF), then MI is computed as the difference between the two entropies: $I(\text{BOLD}; \text{REF}) = H(\text{BOLD}) - H(\text{BOLD}|\text{REF})$. The difference defines MI as a reduced uncertainty about the preceding task condition attributable to the knowledge of the magnitude of the BOLD response.

Here the considered reference task condition (REF) is an AV event. For each voxel in the scanned volume, we use MI to quantify the information contained in its BOLD response about the preceding stimulus condition. Voxels with high MI are more involved in coding the task condition.

Additionally, the information content of a given voxel about the task condition depends on the considered latency (dt) between the task condition and the BOLD response. Typically, for a given voxel, the dependence of MI on different considered latencies is a peaky function, suggesting a preferred latency of response for which information content is maximal. MI of each voxel is therefore computed for different possible latencies (dt), and each voxel is attached with its maximal MI at whatever latency. The range of considered latencies (dt) was 1–5 TRs (2–10 s).

For spatial localization of task-related activity, we construct brain maps of maximal MI values. Significantly informative voxels are chosen as those with maximal MI values that exceed an MI-significance threshold.

For each task-related voxel, we determine its preferred latency as the latency (dt) that maximizes the information contained in its BOLD response about the AV condition. Different voxels have different preferred latencies. Whole-brain temporal maps are constructed, in which the color at each voxel indicates its preferred latency. Because analysis is voxel based, i.e., performed for each voxel independently of other voxels, the spatial resolution of the temporal characteristics is limited only by the scanning resolution.

We note that, in theory, preferred latencies are independent of activation strength, i.e., short and long latencies could accompany either a weak or a strong activation. In practice, weak activations are accompanied by low signal-to-noise ratios; therefore, temporal accuracy could depend on the strength of activation and on the threshold criteria for activation.

Here MI-significance thresholds for activation maps in individual subjects' native space were computed to match the number of significantly activated voxels for a $p < 0.001$ (uncorrected) statistical threshold, determined in the corresponding standard general linear model (GLM) analysis (see below).

The results presented in this study were robust to several considered activation thresholds, including $p < 0.05$, false discovery rate corrected.

Localization using GLM. Individual subjects' activation thresholds was determined by a standard GLM analysis, using BrainVoyager. Predictors of the experimental conditions (AUD, VIS, and AV) were convolved with a canonical gamma function (Boynton et al., 1996). Statistical parametric maps (t statistics) of contrasts were generated for individual subjects with significance values of $p < 0.001$, uncorrected.

Group average of spatiotemporal maps. Group average activation maps were constructed both for MI-spatial activation maps and for preferred latency-temporal maps. For both types of maps, individual subjects' maps were first normalized to standard Montreal Neurological Institute (MNI) brain coordinates using SPM2 (www.fil.ion.ucl.ac.uk/spm), with nearest neighbor algorithm. Group average spatial activation maps were computed from the MNI normalized maps of all subjects. We used the MI-significance thresholds as determined in individual subjects native space. For illustration purposes, the figures depict only voxels that were significantly activated in at least half of the subjects. Group average temporal maps were constructed as the mean of the MNI-normalized temporal maps. In each voxel, only latency values corresponding to suprathreshold activations in individual subjects were considered for the average.

Information flow across ROIs in individual subjects' space. To determine information flow after AV stimulation, we used latency maps obtained in individual subjects' native space. For each subject, we defined anatomical ROIs (see below) in brain regions known to be involved in AV process-

ing, including Heschl's gyrus (van Atteveldt et al., 2004, 2007), superior temporal gyrus (STG) (Hein et al., 2007), middle temporal gyrus (MTG) (Beauchamp et al., 2004), IFC (Belardinelli et al., 2004; Taylor et al., 2006; Hein et al., 2007), secondary visual cortex [inferior occipital cortex (IOC)] (Belardinelli et al., 2004), and primary visual cortex (V1) (Meinenbrock et al., 2007).

For each subject and each ROI, we computed the mean preferred latency, averaged over all significantly activated voxels in that ROI. We then computed the mean preferred latency at that ROI, averaged over all subjects, and the corresponding SD. Subsequently, we explored which regions are preferentially activated earlier than others.

Information flow across regions of interest for unimodal stimulation (individual subjects' space). As a complimentary issue, we also performed whole-brain analysis to compute preferred latencies of all voxels after the unimodal stimuli (AUD/VIS), using the procedure described above. Average latencies of each ROI were computed in the same voxels significantly activated in the AV condition.

Anatomical definition of regions of interest. Anatomical ROIs were defined in individual subjects' space following protocols of the Laboratory of Neuro Imaging Resource at University of California, Los Angeles (<http://www.loni.ucla.edu>). In accordance with those guidelines, STG was subdivided in anterior (aSTG) and posterior (pSTG) portions.

Results

Single-subject results

In a first step, spatiotemporal activation maps for the brain response after AV stimulation were constructed for each of the 18 subjects, using tools from information theory. Figure 2 depicts temporal maps, in which the color indicates the preferred latencies of response to AV stimulation at significantly activated voxels, for several sample subjects.

The results showed a consistent sequence of AV-related brain activations across single subjects (Fig. 2). The shortest latencies were observed in temporal regions along the STS (as short as ~4 s after AV stimulation; blue), whereas frontal and visual regions had a late AV-related activity (as late as ~10 s after AV stimulation; red).

Spatiotemporal information (group average)

Because a consistent temporal pattern of activation was observed in single subjects, we next constructed group average spatiotemporal maps for AV information flow. Group average results are depicted in Figure 3. The left panel of Figure 3 shows the spatial localization of AV-related activity. AV-related activity was observed in frontal and lateral temporal regions and in visual cortex, as expected from the results of the single-subject analysis.

More importantly for the focus of this paper, the right panel of Figure 3 provides temporal information about the sequence of activation flow in these regions. In line with the single-subject results, the earliest component of AV information processing was found in lateral temporal cortex with an average latency as short as 4 s after AV stimulation (Fig. 3, blue green), including superior and middle temporal regions and Heschl's gyrus. A relatively early activation was found in low-level visual regions of the posterior temporal and inferior occipital lobes. Informative AV activity in frontal regions seemed to occur with a long latency, as late as 10 s after AV stimulation (Fig. 3, red).

Information flow (individual subjects' space)

To quantify these differences in preferred latencies between regions of interest and to explore which brain regions are activated earlier than others after AV stimulation, we performed a detailed ROI analysis of temporal activation in the individual subjects' native space (as in Fig. 2).

Because in individual subjects preferred latencies of voxels

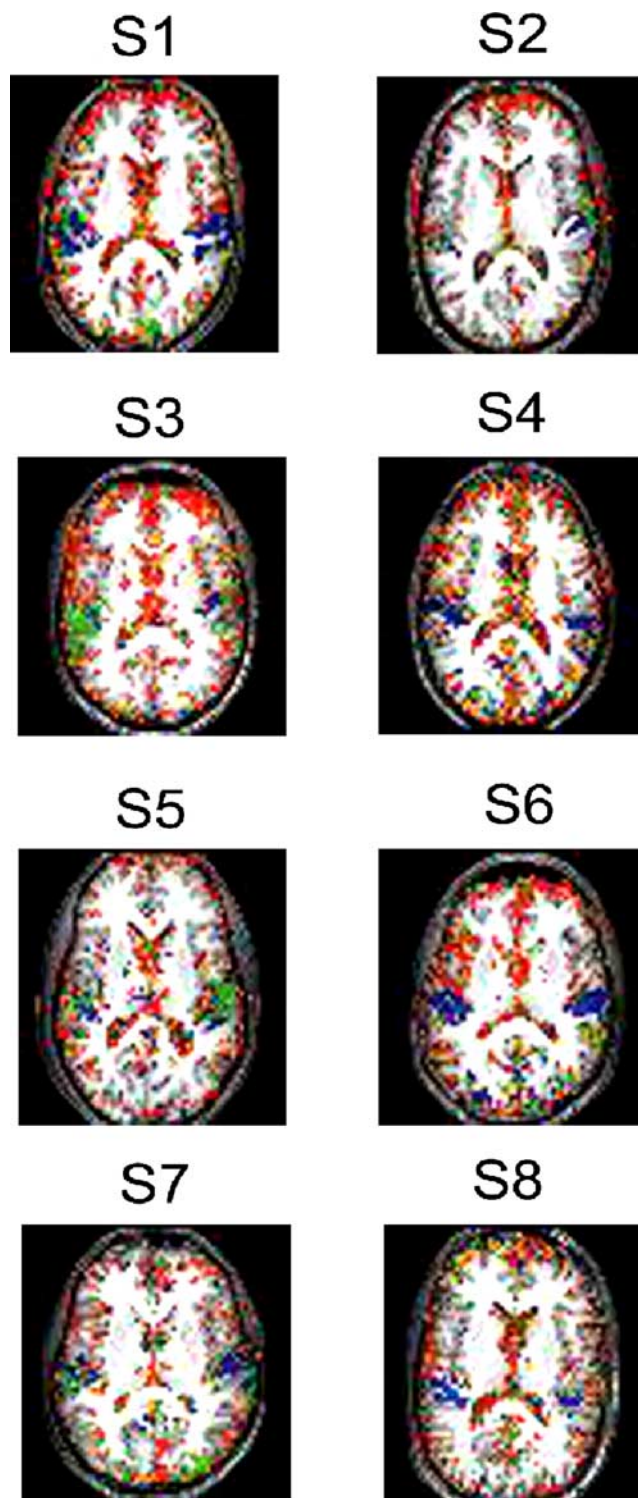


Figure 2. Temporal maps of preferred latencies for AV stimulation of eight representative subjects. Similar brain slices for all subjects are shown. Color coded is the time from AV stimulation, until the BOLD signal in a voxel becomes most informative about the preceding stimulus. Blue, Early AV stimulation response; red, late response. Only voxels in which the mutual information value was suprathreshold are shown.

within each ROI were similar (small SEM; 98% SEMs were <0.15 of mean; 2% were <0.36), for each ROI we computed the mean preferred latency, averaged over all voxels and all subjects. The results are summarized in Figure 4 (top). We found the earliest informative activity in right Heschl's gyrus, left primary visual

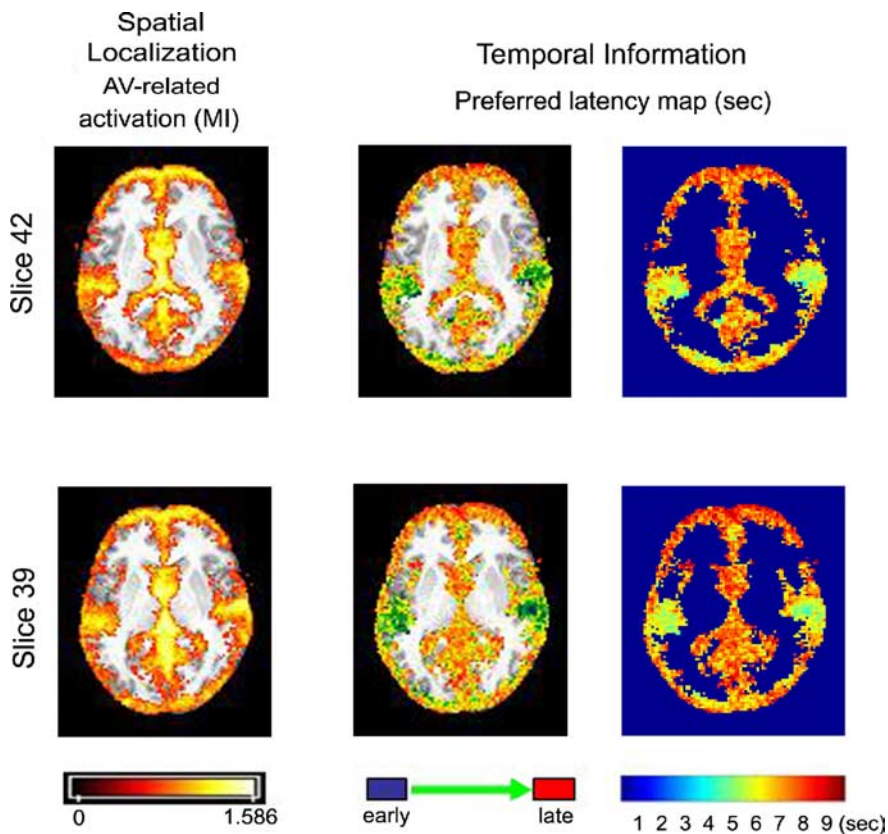


Figure 3. Spatiotemporal group average maps for the AV-related activation. Two brain slices are depicted (slices 42 and 39). Left, Spatial localization map showing MI values for significantly activated voxels for the AV condition on an MNI-normalized brain. Right, Temporal activation map showing the group average preferred latencies for significantly activated voxels. Both maps were overlaid on an MNI brain.

cortex, and pSTG. Latencies for left Heschl's gyrus were significantly longer than for right Heschl's gyrus: $F_{(1,17)} = 4.6$; $p < 0.05$. Informative activity in right V1, aSTG, MTG, IOC, and IFC occurred significantly later than in STG: right pSTG versus right V1; right/left pSTG versus right/left aSTG; right/left pSTG versus right/left MTG; right/left pSTG versus right/left IOC; right/left pSTG versus right/left IFC, all $p < 0.001$. Additionally, we analyzed AV latencies for animal and abstract AV material separately. The results revealed no significant effect of material (natural vs abstract; region \times hemisphere \times material, $F_{(1,17)} < 1$; regions \times material, $F_{(1,17)} = 1.2$, $p > 0.3$; hemisphere \times material, $F_{(1,17)} = 1.2$, $p > 0.27$).

We also computed the preferred latencies of activation in response to unimodal auditory and visual stimulation in the same voxels that were significantly activated for the AV stimuli, in auditory, visual, and frontal regions. Latencies after unimodal auditory stimulation were significantly longer compared with AV stimulation in bilateral Heschl's gyrus, right pSTG, right MTG and right IFC, right Heschl's gyrus: $F_{(1,17)} = 5.5$, $p < 0.03$; left Heschl's gyrus, $F_{(1,17)} = 7.9$, $p < 0.01$; right pSTG, $F_{(1,17)} = 14.1$, $p < 0.002$; right MTG, $F_{(1,17)} = 7$, $p < 0.02$; right IFC, $F_{(1,17)} = 9.6$, $p < 0.01$ (Fig. 4, bottom left). No other ROI showed significant differences between unimodal auditory and AV latencies: right aSTG, $F_{(1,17)} < 1$; left aSTG, $F_{(1,17)} = 2.3$, $p > 0.15$; left pSTG, $F_{(1,17)} = 2.4$, $p > 0.14$; left MTG, $F_{(1,17)} = 1$, $p > 0.3$; left IFC, $F_{(1,17)} = 1.7$, $p > 0.2$.

As in early auditory regions, we also found in primary visual cortex longer latencies after unimodal visual stimulation compared with the processing of AV events: right V1, $F_{(1,17)} = 5.7$,

$p < 0.03$; left V1, $F_{(1,17)} = 11.3$, $p < 0.01$ (Fig. 4, bottom right). Moreover, there was a tendency for longer unimodal latencies in right IFC: $F_{(1,17)} = 3.5$, $p = 0.07$. No other differences in latencies were observed: right and left IOC, $F_{(1,17)} < 1$; right IFC, $F_{(1,17)} = 3$, $p > 0.1$.

Discussion

We applied a novel information-theoretic approach to study the flow of AV information in the human brain. We computed for each voxel in the brain its preferred latency after AV stimulation, i.e., the latency that maximizes the information contained in the BOLD signal about the preceding AV stimulus, in a whole-brain analysis and a detailed ROI analysis in the individual subjects' native space. The results indicate that the earliest AV-informative activity occurs in right Heschl's gyrus, bilateral pSTG, and left primary visual cortex. Latencies of informative activity in other lateral temporal regions such as aSTG and MTG were longer, comparable with latencies in higher-level visual and inferior frontal cortices.

Our finding of shortest AV-related latencies in primary sensory auditory and visual regions, compared with frontal regions, is in line with previous EEG results (Girard and Peronnet, 1999; Fort et al., 2002; Teder-Sälejärvi et al., 2005; Senkowski et al., 2007). Other EEG studies

showed that auditory information reaches primary auditory regions faster than visual information reaches the visual cortex, also supported by our results in secondary visual cortex (Picton et al., 1974; Celesia, 1976; Di Russo et al., 2003). The spatial resolution of the fMRI recording allowed us to study the flow of AV information processing in well defined anatomical substructures, which cannot be specified with the scalp-recorded EEG signals because of the well known inverse problem. Previous fMRI studies have focused on the spatial localization of brain regions involved in AV processing, showing AV-related activity in pSTG (Hein et al., 2007), MTG (Beauchamp et al., 2004), visual (Belardinelli et al., 2004; Meienbrock et al., 2007), and inferior frontal regions (Taylor et al., 2006; Hein et al., 2007). Here we were able to uncover the temporal sequence of informative activity in those AV-related regions.

Previous reports on temporal aspects of information flow in spatially well defined brain region came from intracranial recordings in monkeys (Schroeder et al., 1998, 2001). Although it is possible to estimate timing in the human cortex by extrapolating latencies in the monkey cortex [monkey values are approximately three-fifths of corresponding values in humans (Schroeder et al., 1998)], interspecies comparability remains a critical issue. In our study, we analyze information flow directly in the human brain. Our findings suggest bottom-up processing of AV information from primary sensory areas such as pSTG into association areas in middle temporal, occipital, and inferior frontal cortex. In agreement with our results, monkey data suggest a flow of information from primary sensory areas to middle temporal regions

around the STS and prefrontal cortex (Schroeder et al., 2001; Foxe and Schroeder, 2005).

Moreover, we found that neighboring brain regions in the human temporal lobe such as pSTG, aSTG, and MTG process AV stimuli with different latencies. This suggests functional differentiation between these subregions in the human brain, in line with previous reports regarding differences in the processing of object sounds between aSTG and pSTG (Altmann et al., 2007) and differences in AV integration properties between pSTG and MTG (Hein et al., 2007).

We also studied the preferred latencies of responses to unimodal stimulation in the same AV-related voxels. Previous fMRI studies reported amplitude modulations of AV activations compared with unimodal AUD/VIS activity (Calvert et al., 2000; Beauchamp et al., 2004; van Atteveldt et al., 2004, 2007; Taylor et al., 2006; Stevenson et al., 2007). Our findings add temporal information to this line of research, showing shorter latencies for the processing of AV stimulation compared with AUD/VIS alone in IFC and most strongly in primary auditory and visual regions, pointing to a facilitation of early sensory processing during multisensory stimulation. Similar findings were reported by Martuzzi et al. (2007), analyzing BOLD response dynamics and showing a facilitation of the event-related BOLD response peak latencies for AV stimuli compared with unimodal conditions. Additional evidence for multisensory effects within primary sensory cortices, carried by either direct intracortical connections or nonspecific thalamic inputs, is provided by electrophysiological data in humans (van Wassenhove et al., 2005) and nonhuman primates (Schroeder and Foxe, 2005; Lakatos et al., 2007). Such low-level multisensory interactions may be part of a global multisensory network spanning multiple neocortical regions (Ghazanfar and Schroeder, 2006). A mechanism underlying modulation within this network might rely on phase resetting of spontaneous neural activity by inputs of one modality on the phase of incoming inputs of the other modality, relative to the oscillatory cycle, as shown for auditory–somatosensory integration in the primary auditory cortex by Lakatos et al. (2007).

As in many other studies (Beauchamp et al., 2004; Belardinelli et al., 2004; van Atteveldt et al., 2004, 2007), we investigated AV processing with static images and sounds. There is evidence that different neural networks are involved if dynamic visual stimuli are matched with auditory dynamics (Campbell et al., 1997; Munhall et al., 2002; Calvert and Campbell, 2003). Natural and behaviorally relevant stimuli, as well as active tasks, may elicit a different temporal flow of information processing than passive perception of static objects used in this study.

Given that fMRI does not directly record neural activity but rather the hemodynamic response linked to it, the absolute values of BOLD response latencies are longer than those recorded electrophysiologically (e.g., monkey data). However, the relative la-

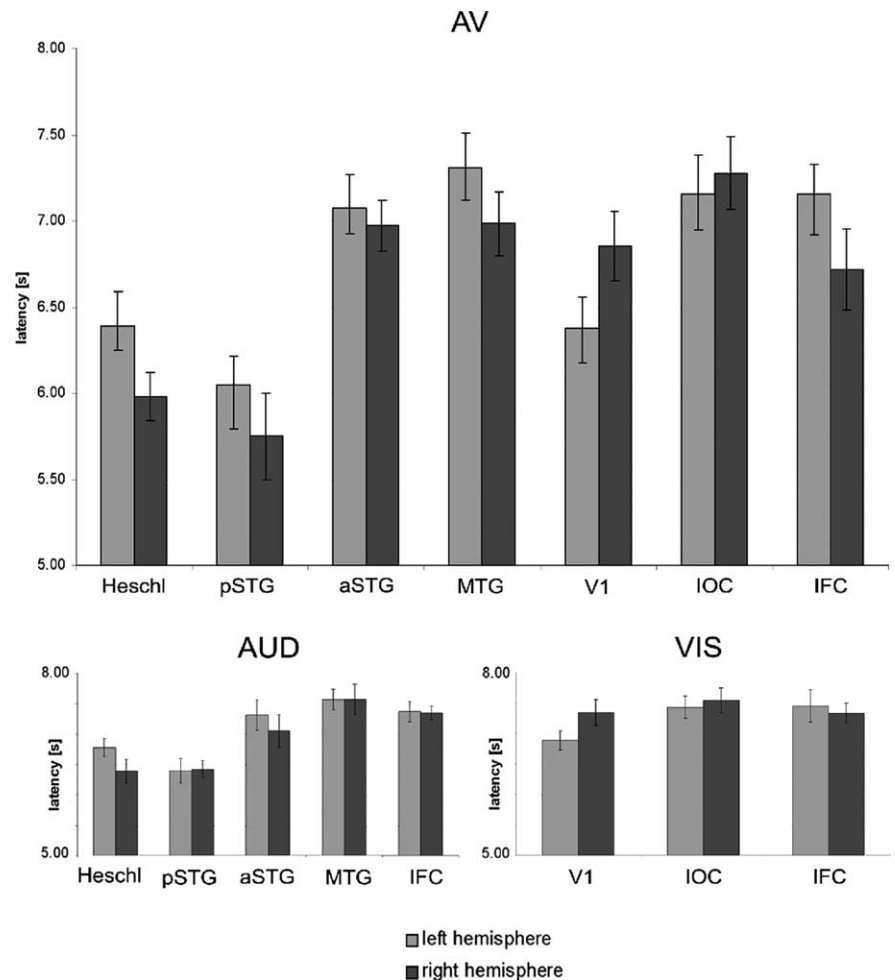


Figure 4. Information flow in anatomical ROIs defined in native subject space after AV stimulation (top), unimodal AUD stimulation (bottom left), and unimodal VIS stimulation (bottom right).

tencies and sequence of ROI activation reported here are supported by extensive animal and human electrophysiological literature (Picton et al., 1974; Celesia, 1976; Girard and Peronnet, 1999; Fort et al., 2002; Di Russo et al., 2003; Foxe and Schroeder, 2005; Teder-Sälejärvi et al., 2005; Senkowski et al., 2007). Moreover, the latencies are different for the various stimulus conditions (AV, AUD, and VIS), suggesting functional validity of the observed flow pattern.

References

- Altmann CF, Bledowski C, Wibral M, Kaiser J (2007) Processing of location and pattern changes of natural sounds in the human auditory cortex. *NeuroImage* 35:1192–1200.
- Beauchamp MS, Lee KE, Argall BD, Martin A (2004) Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron* 4:809–823.
- Belardinelli MO, Sestieri C, Di Matteo R, Delogu F, Del Gratta C, Ferretti A, Caulo M, Tartaro A, Romani GL (2004) Audio-visual crossmodal interactions in environmental perception: an fMRI investigation. *Cogn Process* 5:167–174.
- Boynton GM, Engel SA, Glover GH, Heeger DJ (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16:4207–4221.
- Calvert GA, Campbell R (2003) Reading speech from still and moving faces: the neural substrates of visible speech. *J Cogn Neurosci* 15:57–70.
- Calvert GA, Campbell R, Brammer MJ (2000) Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr Biol* 10:649–657.

- Campbell R, Zihl J, Massaro DW, Munhall K, Cohen MM (1997) Speechreading in the akinetopsic patient, L.M. *Brain* 120:1793–1803.
- Celesia GG (1976) Organization of auditory cortical areas in man. *Brain* 99:403–414.
- Di Russo F, Martinez A, Hillyard SA (2003) Source analysis of event-related cortical activity during visuo-spatial attention. *Cereb Cortex* 13:486–499.
- Fort A, Delpuech C, Pernier J, Girard MH (2002) Dynamics of cortico-subcortical cross-modal operations involved in audio-visual object detection in humans. *Cereb Cortex* 12:1031–1039.
- Foxe JJ, Schroeder CE (2005) The case for feedforward multisensory convergence during early cortical processing. *NeuroReport* 16:419–423.
- Fuhrmann Alpert G, Sun FT, Handwerker D, D'Esposito M, Knight RT (2007) Spatio-temporal information analysis of event-related BOLD response. *NeuroImage* 34:1545–1561.
- Ghazanfar AA, Schroeder CE (2006) Is neocortex essentially multisensory? *Trends Cogn Sci* 10:278–285.
- Girard MH, Peronnet F (1999) Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J Cogn Neurosci* 11:473–490.
- Hein G, Doehrmann O, Müller NG, Kaiser J, Muckli L, Naumer MJ (2007) Object familiarity and semantic congruency modulate responses in cortical audiovisual integration areas. *J Neurosci* 27:7881–7887.
- Karns CM, Knight RT (2007) Electrophysiological evidence that intermodal auditory, visual, and tactile attention modulates early stages of neural processing. *Soc Neurosci Abstr* 33:423.24.
- Lakatos P, Chen CM, O'Connell MN, Mills A, Schroeder CE (2007) Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 18:279–292.
- Martuzzi R, Murray MM, Michel CM, Thiran JP, Maeder PP, Clarke S, Meuli RA (2007) Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cereb Cortex* 17:1672–1679.
- Meienbrock A, Naumer MJ, Doehrmann O, Singer W, Muckli L (2007) Retinotopic effects during spatial audio-visual integration. *Neuropsychologia* 45:531–539.
- Molholm S, Ritter W, Javitt DC, Foxe JJ (2004) Multisensory visual-auditory object recognition in humans: a high-density electrical mapping study. *Cereb Cortex* 14:452–465.
- Munhall KG, Servos P, Santi A, Goodale MA (2002) Dynamic visual speech perception in a patient with visual form agnosia. *NeuroReport* 13:1793–1796.
- Picton TW, Hillyard SA, Krausz HI, Galambos R (1974) Human auditory evoked potentials. I. Evaluation of components. *Electroencephalogr Clin Neurophysiol* 36:179–190.
- Schroeder CE, Foxe J (2005) Multisensory contributions to low-level “unisensory” processing. *Curr Opin Neurobiol* 15:454–458.
- Schroeder CE, Mehta AD, Givre SJ (1998) A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cereb Cortex* 8:575–592.
- Schroeder CE, Mehta AD, Foxe JJ (2001) Determinants and mechanisms of attentional modulation of neural processing. *Front Biosci* 6:672–684.
- Senkowski D, Saint-Amour D, Kelly SP, Foxe JJ (2007) Multisensory processing of naturalistic objects in motion: a high-density electrical mapping and source estimation study. *NeuroImage* 36:877–888.
- Stevenson RA, Geoghegan ML, James TW (2007) Superadditive BOLD activation in superior temporal sulcus with threshold non-speech objects. *Exp Brain Res* 179:85–95.
- Talsma D, Doty TJ, Woldorff MG (2007) Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cereb Cortex* 17:679–690.
- Taylor KI, Moss HE, Stamatakis EA, Tyler LK (2006) Binding crossmodal object features in perirhinal cortex. *Proc Natl Acad Sci USA* 21:8239–8244.
- Teder-Sälejärvi WA, Di Russo F, McDonald JJ, Hillyard SA (2005) Effects of spatial congruity on audio-visual multimodal integration. *J Cogn Neurosci* 17:1396–1409.
- van Atteveldt NM, Formisano E, Goebel R, Blomert L (2004) Integration of letters and speech sounds in the human brain. *Neuron* 43:271–282.
- van Atteveldt NM, Formisano E, Blomert L, Goebel R (2007) The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cereb Cortex* 17:962–974.
- van Wassenhove V, Grant KW, Poeppel D (2005) Visual speech speeds up the neural processing of auditory speech. *Proc Natl Acad Sci USA* 4:1181–1186.