Audio-Visual Multisensory Integration in Superior Parietal Lobule Revealed by Human Intracranial Recordings

Sophie Molholm,1,5 Pejman Sehatpour,1 Ashesh D. Mehta,2 Marina Shpaner,1,5 Manuel Gomez-Ramirez,1,5 Stephanie Ortigue,4 Jonathan P. Dyke,3 Theodore H. Schwartz,2 and John J. Foxe1,5

1The Cognitive Neurophysiology Laboratory, Program in Cognitive Neuroscience and Schizophrenia, Nathan S. Kline Institute for Psychiatric Research, Orangeburg; 2Department of Neurological Surgery and 3Department of Radiology, Weill Cornell Medical College, New York Presbyterian Hospital, New York, New York; 4Dartmouth Functional Brain Imaging Center, Center for Cognitive Neuroscience, Dartmouth College, Hanover, New Hampshire; and 5Program in Cognitive Neuroscience, Department of Psychology, The City College of the City University of New York, New York, New York

Submitted 15 March 2006; accepted in final form 6 May 2006

INTRODUCTION

The bulk of our knowledge regarding multisensory processing in the parietal cortex comes from intracranial recordings in animals (Andersen et al. 1997; Barth et al. 1995; Brett-Green et al. 2004; Cohen et al. 2004; Di et al. 1994; Mazzoni et al. 1996; Schlack et al. 2005; Wallace et al. 1993, 2004). Single unit recordings in nonhuman primates, with the greatest across-species anatomical correspondence to humans, have revealed multisensory neurons in the inferior parietal sulcus (IPS) that are responsive to combinations of visual, auditory, and tactile stimuli (Andersen et al. 1997; Cohen et al. 2004; Mazzoni et al. 1996; Schlack et al. 2005). However, the homologies between primate and human multisensory parietal regions remain to be fully established (Astafilev et al. 2003; Sereno and Tootell 2005). Human functional imaging studies have shown that multiple sensory inputs are indeed co-localized to regions of the parietal lobe, including the IPS and the superior parietal lobule (SPL) (Bremmer et al. 2001; Calvert et al. 2001; Lewis et al. 2000; Macaluso and Driver 2001). A subset of these studies also shows nonlinear interactions of multisensory inputs, suggesting that this information is integrated (Calvert et al. 2001; Lewis et al. 2000; Miller and D’Esposito 2005). That is, these studies have shown that the response to a bisensory stimulus differs from the sum of the responses to its unisensory constituents (so-called super- or subadditivity; see Stanford et al. 2005).

While hemodynamic imaging has provided excellent spatial localization of multisensory processing in humans, the temporal resolution of this method precludes the study of dynamic information processing, where meaningful distinctions are seen on the order of 10s and 100s of milliseconds. Hence, it is not possible to resolve whether this multisensory processing represents direct sensory-perceptual level interactions or if it reflects later cognitive processes (Foxe and Schroeder 2005; Schroeder and Foxe 2005). This lack of timing information may be the reason that imaging data can lend itself to alternate and equally plausible interpretations. For instance, Ojanen et al. (2005) attribute SPL activation for conflicting auditory-visual speech compared with matching auditory-visual speech to increased attentional processing, a function more often associated with superior parietal regions than multisensory processing.

Here, we took advantage of the excellent spatial and temporal resolution provided by intracranial electrical recordings in humans to directly study multisensory processing in parietal cortex. Using a simple reaction-time task (Molholm et al. 2002) in which subjects responded to visual and auditory stimuli that were presented simultaneously or alone, we identified a highly localized region of parietal cortex, in the region of the lateral superior parietal lobule, that responded to both auditory and visual stimulation. Auditory and visual inputs to this region occurred early in time (≤80 ms), and multisensory integration processes were evident shortly thereafter. The timing of the inputs and ensuing multisensory interactions are consistent with sensory-perceptual level processing in the SPL.

METHODS

Subjects

Data from three individuals with epilepsy are reported (ages 29, 35, and 45 yr). The patients were implanted with subdural electrodes for the costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Address for reprint requests and other correspondence: S. Molholm or J. Foxe, Cognitive Neurophysiology Lab., Program in Cognitive Neuroscience and Schizophrenia, Nathan S. Kline Inst. for Psychiatric Research, 140 Old Orangeburg Rd., Orangeburg, NY 10962 (E-mail: molholm@nki.rfmh.org; foxe@nki.rfmh.org).
evaluation of the foci of pharmacologically intractable epilepsy. They were all men, and two were right-handed. Recordings were made after all clinical procedures related to seizure localization were completed. During localization procedures, subjects were removed from their antiepileptic medications until sufficient numbers of seizures were recorded, but when these clinical measures were completed, they were immediately returned to their regular doses. All three were receiving a combination of levetiracetam (keppra) and zonisamide (zonegran). All recordings for this study were made after subjects had been restarted on their medications. Clinical data are presented in Table 1. Two of the three subjects had left hippocampal foci with the third showing a multifocal right hemisphere neocortical disorder. Two of the three had relatively early seizure onsets at ages 1 and 5, whereas the third had relatively late onset at age 16. This latter subject was one of the left hippocampal patients. In studies examining electrical responses in patients with long-term epilepsy, there is always the possibility that reorganization has occurred. Nevertheless, given the different histories of the patients, there is no reason to expect that subjects would have undergone similar cortical reorganization. As such, patterns of activity seen consistently across the three subjects quite likely reflect “normal” information processing.

All subjects provided written informed consent after the procedures were fully explained to them before entering into this study. The Institutional Review Boards at both Nathan S. Kline Institute and Weill Cornell Medical College approved all experimental procedures.

Stimuli and task

AUDITORY ALONE. Tone pips (1,000-Hz; 60-ms duration; 10-ms rise-fall; 75 dB SPL) were presented over headphones (Sennheiser HD600).

VISUAL ALONE. A disk (60-ms duration), subtending 1.2° in diameter (140-cm viewing distance) and appearing red on a black background, was presented centrally on a cathode ray tube (CRT) computer monitor.

BISENSORY AUDIO-VISUAL. The auditory and visual stimuli described above were presented simultaneously.

PROCEDURE. Subjects were instructed to make button press responses as quickly as possible using their right index finger when a stimulus in either or both stimulus modalities was detected. The purpose of this task was to ensure that subjects attended the stimuli. Subjects were instructed to maintain fixation on a centrally located cross. They were visually monitored by two experimenters throughout recordings to ensure that fixation was maintained and were verbally prompted to re-engage when/if necessary. Stimulation was immediately aborted if subjects became fatigued or found difficulty in fixation. Interstimulus interval (ISI) varied randomly between 690 and 2,940 ms. The three stimulus conditions were presented with equal probability in random order, such that subjects could not predict either when or what would occur next. Stimuli were blocked into sequences of 150 trials. Frequent breaks were provided to maintain concentration and prevent fatigue.

EEG recordings

Continuous EEG from 75–118 subdurally placed electrodes was recorded using BrainAmp amplifiers (Brain Products, Munich, Germany). The data were band-pass filtered on-line from 0.05 to 100 Hz and digitized at 1,000 Hz. A frontally placed intracranial electrode was used as the reference. The continuous EEG was divided into pre- to poststimulus onset epochs and baseline corrected over the full epoch. An artifact criterion of ±300 μV was applied to electrodes within the region of interest to reject trials with excessive noise transients. An average of 542 trials was accepted per stimulus condition. When clean averages had been obtained, baseline was redefined as the epoch from −100 to 0 ms before stimulus onset.

For all subjects, the electrode site(s) from which data were analyzed was chosen based on the following criteria: 1) it was over parietal cortex; 2) both auditory and visual stimuli elicited a robust unisensory response at the site; and 3) both the auditory and the visual responses were larger than the corresponding responses from the surrounding electrodes. We reasoned that if the maximum auditory and visual responses were at different electrode sites, this would indicate that largely different sets of neurons were responding to the two stimulus types. No analyses were performed on data from nonideal sites.

EEG epochs were sorted according to stimulus condition and averaged for each subject to compute the event-related potential (ERP). Statistics were performed on individual subject data. For each subject, the EEG epochs for a given condition were used to calculate the SE of the ERP response at each time-point. Differences between conditions (or from baseline) that fell outside the error of the mean were considered significant (for a similar application to human intracranial data (see Rizzuto et al. 2005). Performing tests on multiple time-points increases the probability of a false positive. We therefore only considered significant differences that were present for 10 consecutive time-points, because the likelihood of getting 10 false positives in a row is considerably low (Molholm et al. 2002; Murray et al. 2001). Multisensory interactions were defined as a significant difference between the audio-visual (AV) response and the summed auditory (A) and visual (V) responses (Foxe et al. 2000; Molholm et al. 2002, 2004; Murray et al. 2001). Individual epochs of A and V were summed and used in the calculation of the SE for the summed response (A + V).

Electrode placement and localization

In two left hemisphere cases, an 8 × 8 grid of electrodes was placed that spanned frontal, parietal, and temporal lobe neocortex. Several 6-contact strips were slid under the dura to extend the coverage toward the interhemispheric fissure and beneath the temporal lobe. A depth electrode was placed stereotactically in the body of the hippocampus. In the third case, bilateral strip electrodes were placed in a starburst pattern over both frontal, parietal, temporal, and occipital lobes. Bilateral depth electrodes were placed in the hippocampus. Interelectrode spacing within a grid or strip was 1 cm. High-resolution presurgical MRIs were coregistered with postsurgical MRIs using a method that lines up the anterior and the posterior commissures. These were reconstructed into three-dimensional (3D) images. Next, the 3D coordinates of each electrode were calculated from the postsurgical MRI and mapped to the presurgery MRI. The BrainVoyager software suite (Brain Innovation, Maastricht, The Netherlands) was used for coregistration and reconstruction of the MRIs, as well as to calculate the electrode coordinates. The BrainVoyager-BESA (Version 5.0.2) software package (MEGIS Software, Munich, Germany) was used to project the electrode coordinates onto the presurgical MRI.

The localization of electrodes within parietal cortex was done with respect to the IPS. Because the central sulcus is a prominent landmark, we first identified this sulcus (black arrows in Fig. 2). We identified the postcentral sulcus (POCS) as a sulcus that runs parallel and posterior to the central sulcus. The intraparietal sulcus (IPS, yellow arrows in Fig. 2) was identified as a sulcus that runs somewhat perpendicular to the plane of the POCS with an inclination toward the midline (Ebeling and Steinmetz 1995).
TABLE 1. Subject clinical data

<table>
<thead>
<tr>
<th>Subject</th>
<th>Age at Onset, yr</th>
<th>Type of Seizures</th>
<th>Location of Epileptic Foci</th>
<th>Neuropsych</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>KK</td>
<td>16</td>
<td>CPS</td>
<td>L. hippocampus and lateral temporal lobe</td>
<td>No major deficits</td>
<td>Left</td>
</tr>
<tr>
<td>DM</td>
<td>5</td>
<td>CPS, GTC</td>
<td>R. multifocal neocortical</td>
<td>No major deficits</td>
<td>Left</td>
</tr>
<tr>
<td>VH</td>
<td>1</td>
<td>CPS, GTC</td>
<td>L. hippocampus</td>
<td>Poor memory</td>
<td>Left</td>
</tr>
</tbody>
</table>

CPS, complex partial seizure; GTC, generalized tonic clonic.

Behavioral measures

Button press responses to the three stimulus conditions were acquired during the recording of the EEG and processed off-line. For this simple reaction-time task, a hit was recorded when a response after any of the stimuli fell within a predefined response window of 100- to 800-ms poststimulus onset. This window was used to avoid the double categorization of a response. The percent hits and average response time were calculated for each stimulus condition for each subject. Misses were simply the remainder of the percent hits and thus were not analyzed. Faster reaction times for the multisensory compared with each of the unisensory stimuli was followed by a test of Miller’s race model (Miller 1982), to determine if response facilitation could be accounted for by simple probability summation of the fastest responses triggered by two independently operating inputs (i.e., the race model). When probability summation cannot account for the observed response facilitation (i.e., the race model is violated), response facilitation is unquestionably caused by interactions between the auditory and visual information during neural processing (see Molholm et al. 2002 for a detailed description of the race model and its implementation).

The race model places an upper limit on the cumulative probability (CP) of reaction-time (RT) at a given latency for stimulus pairs with redundant targets. For any latency, t, the race model holds when this CP value is less than or equal to the sum of the CP from each of the single target stimuli (the unisensory stimuli) minus an expression of their joint probability \((CP_{t}^{AV} \leq \left[CP_{t}^{Aud-unisensory} + CP_{t}^{Vis-unisensory}\right] - \left(\left[CP_{t}^{Aud-unisensory} \times CP_{t}^{Vis-unisensory}\right]\right))\). For each subject, the RT range (100–800 ms) was calculated over the three stimulus types (auditory-visual-bisensory, auditory-unisensory, and visual-unisensory) and divided into quantiles from the 5th to 100th percentile and for DM, it was the case across quantiles 4 through 14, with an average speeding of 26.0 ms, a large part of the distribution. For KK, this was seen across quantiles 1 through 13, with an average speeding of 26.0 ms, and for DM, it was the case across quantiles 4 through 14, with an average speeding of 4.1 ms. In the case of VH, for whom the race model was not systematically violated, there was a very close match between the combined unisensory RT distribution and the bisensory distribution. Nonetheless, across five of the actual recorded bisensory RT data, fully 7 and 9% of the RTs to the bisensory AV condition fall to the left of the predicted distributions (i.e., are faster). For subject VH, only 1% of the RTs to the multisensory condition exceeded the predictions of probability summation. Race model violations are shown in Fig. 1, where positive values indicate that the probability of the RTs to the AV trials exceeded that predicted by the race model.

To further elaborate on these results, we used the following strategy to determine if this speeding of RT was statistically significant on a within-subject basis. 1) For each subject, all individual RTs to each of the unisensory stimuli were collapsed into a single distribution and arranged from the fastest to the slowest. 2) The slower half of this combined distribution was discarded because there were twice as many unisensory RTs as bisensory. 3) The faster half of the unisensory RT distribution was binned into 20 quantiles as defined by the total number of trials (e.g., if there were 400 trials, a quantile would include 20 RTs). The same binning procedure was performed on the entire bisensory RT distribution. 4) RTs within each quantile were compared between the unisensory and bisensory conditions using a two-tailed Student’s \(t\)-test (Student 1908).

Not surprisingly, for the two subjects where the race model was clearly violated, significantly faster RTs were seen across a large part of the distribution. For KK, this was seen across quantiles 1 through 13, with an average speeding of 26.0 ms, and for DM, it was the case across quantiles 4 through 14, with an average speeding of 4.1 ms. In the case of VH, for whom the race model was not systematically violated, there was a very close match between the combined unisensory RT distribution and the bisensory distribution. Nonetheless, across five of the

RESULTS

Behavioral data

Across the three subjects, response times were fastest for the multisensory condition (average = 279 ms), intermediate for the auditory condition (334 ms), and slowest for the visual condition (355 ms). This pattern of mean RT data is equivalent to that observed in our earlier study on a larger sample of 12 subjects (Molholm et al. 2002). The race model was violated in all three subjects, strongly in two and only weakly in the third, indicating that when presented together, the auditory and visual information were integrated causing speeded RTs. The race model violation was particularly pronounced for subjects DM and KK. That is, if one compares the distribution of RTs that is predicted by a simple probability summation account to the

FIG. 1. Violation of the race model. Positive values indicate that the probability of the reaction-times to the audio-visual trials exceeded predictions of the race model.
early quantiles (2 through 6), bimodal RTs were also significantly faster than unisensory, albeit by a modest 1.4 ms.

For individual subjects, the pattern of RTs was for the most part paralleled by percent hits. The individual performance data (reaction times and percent hits) are presented in Table 2.

**Electrophysiological data**

**Parietal electrode locations.** In two of the subjects, in which left hemisphere grids were placed, the electrode that met the predetermined criteria was located just anterior to the intraparietal sulcus over the superior parietal lobule. In the third subject, the location was located on a somewhat more lateral and anterior portion of the left parietal lobe. However, the sparser electrode coverage in this subject prevented as precise a localization of this activity as was possible for the first two subjects (strips rather than grids were used). Nonetheless, the highly similar morphology and timing of the responses to those of the other subjects strongly suggests a similar origin. The middle column of Fig. 2 displays the location of the electrode for each of the subjects. The specificity of the response is shown in Fig. 3, in which the response of interest and the responses at the surrounding electrodes are shown, for subject VH. To facilitate the comparison of these data to group data acquired from functional neuroimaging studies, the talairach coordinates of the electrode of interest (Fig. 2) are given for each of the subjects in Table 3, determined following talairach normalization of their brains. It should be kept in mind that this highly useful talairach normalization procedure nonetheless distorts the anatomy of the individual brains, and thus reference to electrode placement with respect to the surrounding sulci and gyri in the untransformed individual brains is highly informative given the cortical geometric variability we observe.

The sites from which data were analyzed did not exhibit interictal epileptiform activity, and in at least two of the three cases, the sites were far from the epileptic foci (for comparison, see Table 1 for the location of epileptic foci and Fig. 2 for the electrode location).

**TABLE 2. Behavioral data**

<table>
<thead>
<tr>
<th></th>
<th>Audio-Visual</th>
<th>Auditory</th>
<th>Visual</th>
</tr>
</thead>
<tbody>
<tr>
<td>KK</td>
<td>359</td>
<td>471</td>
<td>409</td>
</tr>
<tr>
<td>VH</td>
<td>243</td>
<td>270</td>
<td>341</td>
</tr>
<tr>
<td>DM</td>
<td>236</td>
<td>260</td>
<td>314</td>
</tr>
</tbody>
</table>

Values are reaction times and percent hits.

![Fig. 2.](http://jn.physiology.org/). Auditory (gray trace) and visual (red trace) responses are shown in column 1. Location of electrodes that met criteria are indicated by red dots on the reconstructed brains of individual subjects in column 2; for landmarks, we put black arrows to indicate estimated location of the central sulcus and yellow arrows to indicate estimated location of the inferior parietal sulcus (IPS). Audio-visual (red) and summed unisensory (auditory alone plus visual alone; in blue) responses are shown in column 3. For the event-related potentials (ERPs), SE is represented by thickness of trace.
AUDITORY RESPONSE IN THE SPL. In all subjects, the response to the auditory condition exhibited an initial positive deflection followed by a negative deflection. This positive-then-negative complex was largest at the site of interest compared with the surrounding sites. The initial deflection was statistically different from baseline starting between 24 and 34 ms across the three subjects. Auditory onset latencies for the individual subjects are presented in Table 4, and the auditory responses are shown in gray in the first column of Fig. 2.

VISUAL RESPONSE IN THE SPL. Similar to the auditory response, in the visual response there was an initial positive deflection followed by a negative deflection for all subjects. Again, this positive-then-negative complex was largest at the site of interest compared with the surrounding sites. The initial deflection was statistically different from baseline starting between 62 and 90 ms across the three subjects. Visual onset latencies for the individual subjects are presented in Table 4, and the visual responses are shown in red in the first column of Fig. 2.

MULTISENSORY RESPONSE IN THE SPL. The multisensory response was very similar to the auditory alone response, with an initial positive deflection followed by a negative deflection (see red trace in the last column of Fig. 2). This positive-then-negative complex was largest at the site of interest compared with the surrounding sites. The multisensory response initially behaved in a linear fashion; that is, it was identical to the artificially summed A + V response. However, beginning between 120 and 160 ms, multisensory interactions were clearly evident, with the multisensory response differing significantly from the algebraic sum of the unisensory responses.

TABLE 3. Talairach coordinates of the electrodes

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>KK</td>
<td>-44</td>
<td>-14</td>
<td>46</td>
</tr>
<tr>
<td>VH</td>
<td>-40</td>
<td>-20</td>
<td>57</td>
</tr>
<tr>
<td>DM</td>
<td>-52</td>
<td>-18</td>
<td>50</td>
</tr>
</tbody>
</table>

TABLE 4. Electrophysiological response onsets

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>V</th>
<th>AV - (A + V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KK</td>
<td>30</td>
<td>74</td>
<td>130</td>
</tr>
<tr>
<td>VH</td>
<td>24</td>
<td>62</td>
<td>118</td>
</tr>
<tr>
<td>DM</td>
<td>34</td>
<td>90</td>
<td>160</td>
</tr>
</tbody>
</table>

A. auditory; V, visual.
The difference onsets for the individual subjects are presented in Table 4. The multisensory (red) and summed (blue) responses are presented in the rightmost column of Fig. 2. In all three cases, the multisensory response seems to return more quickly to baseline than the summed response, although in the first case (VH), this response rebounds and significantly exceeds baseline in the positive direction. It is tempting to interpret this as evidence for so-called subadditivity, but it is important to point out that event-related potential measures cannot distinguish between inhibition, disinhibition, or excitation. For example, in this case, it is perfectly possible that an additional positive going generator has become active [e.g., excitatory postsynaptic potentials (EPSPs) in the supralaminar layers].

**Discussion**

These intracranial data provide the first direct electrophysiological evidence of multisensory processing in the human SPL. Auditory and visual sensory information projected to the same highly localized parietal region with auditory inputs arriving considerably earlier (30 ms) than visual inputs (~75 ms). These multisensory inputs were also clearly integrated in this same region, showing a consistent nonlinear response across all three subjects that onset between 120 and 160 ms. The morphology and time-course of the responses were also remarkably similar across subjects, strongly suggesting that the recordings were from a functionally homologous region of cortex. In the two subjects for whom we had extensive coverage of the surrounding region and were therefore able to precisely localize the source of the activity, the electrodes were placed 1) on the anterior aspect of the IPS (KK) and 2) just superior to the IPS (VH). Thus these recordings likely reflect activity from a region just anterior to the IPS on the lateral SPL. The results are consistent with the human neuroimaging literature, which consistently shows co-localization of auditory and visual and tactile and visual processing in these regions (Bushara et al. 1999; Calvert et al. 2001; Macaluso and Driver 2001; Sereno and Huang 2005), and that these inputs also show nonlinear multisensory interactions within this region (Calvert et al. 2001).

**Role of a multisensory SPL**

Human SPL is well known for its role in spatial attention and in particular in shifts of spatial attention. (Bushara et al. 1999; Corbetta et al. 1993; Pardo et al. 1991; Vandenbergh et al. 2001) Also, considerable evidence now indicates that spatial attention is supramodal (Driver and Spence 1998; Eimer 2001), and further that SPL is involved in both auditory and visual spatial localization (Bushara et al. 1999). As such, the coregistration of visual and auditory spatial processing in SPL could well account for oft-seen multimodal spatial attention effects. For example, a number of studies have shown that when subjects attend to stimuli in a given modality, responses to colocalized inputs in another modality are also enhanced, even when this second modality information is completely task-irrelevant (Hillyard et al. 1984; Macaluso and Driver 2001; McDonald and Ward 2000; Teder-Salejarvi et al. 1999). Shomstein and Yantis (2004) have also found SPL involvement in switching attention between auditory and visual streams of information, although it should be noted that their SPL activations were somewhat more medially focused than those in this study. Further evidence for a role in spatial representation comes from a very recent imaging study by Sereno and Huang (2005) who found both visual and somatosensory inputs to a similar region of the SPL as seen here, where these inputs showed both retinotopic and somatotopic organizations. Although spatial attention was not an overt factor in our study, we would nevertheless argue that attending to and processing a stimulus almost always involves processing its location. Indeed, in her well-known model of feature integration, Treisman argues that attention to an object necessarily involves attention to its location (Evans and Treisman 2005; Treisman 1982; Treisman and Gelade 1980). At the same time, it should be pointed out that because of the adverse conditions inherent in making recordings in a hospital room, the stimuli were not actually presented to identical locations insofar as visual stimuli were presented on a monitor placed in front of the subject and the auditory stimuli were presented over headphones. Spence and Driver (1997; also Driver and Spence 1998) have argued the importance of spatial register for spatial attention effects, and it is possible that this mismatch would preclude spatial attention effects.

The SPL also plays an important role in the transformation of multisensory inputs into the same spatial reference frames (e.g., body centered or head centered), and in sensory–motor integration (Andersen et al. 1997; Cohen and Andersen 2004; Iacoboni et al. 1998; Lacquaniti and Caminiti 1998; Strickanthe et al. 1996; Tanabe et al. 2005). For example, it was shown using functional MRI (fMRI) that regions in and around the SPL, including the anterior IPS, were activated during sensorimotor transformations for both eye and finger movements when these movements were triggered by either somatosensory or visual cues (Tanabe et al. 2005). Similarly, a positron emission tomography (PET) study has shown colocalization of SPL activation for auditory or visual cueing of motor responses (Iacoboni et al. 1998). From our data, it is tempting to speculate that, because subjects responded with the right index finger, the left hemisphere SPL multisensory activity reflects sensory–motor integration processes. However, this remains to be explicitly tested with subjects alternating the hand of response and electrode coverage over the equivalent region in the right hemisphere.

While these particular functions associated with the SPL would greatly benefit from the coregistration and integration of multisensory inputs, it is obvious that further work is necessary to specify the exact role of the multisensory zone of the SPL that we have identified here. The ultimate goal of any such future work will not only be to characterize the zone in terms of its functionality, but also to determine the underlying principles for multisensory convergence and interactions, as has been done for the superior colliculus in the seminal work of Stein and colleagues (Stein and Meredith 1993).

The timing of the onset of multisensory processing is key to understanding the functional role of these SPL activations. Of course, integration cannot begin until the later visual input arrives in SPL, which occurs here at an average latency of 75 ms. However, the convergent auditory and visual inputs sum linearly for a considerable period thereafter, with multisensory integration lagging visual onset by an average of 60 ms (ranging from 54 to 70 ms). The average onset of multisensory
processing in SPL at 135 ms is therefore noticeably late. Previous studies have shown that multisensory processing can onset within the time frame of early sensory processing, revealing the possibility of very early multisensory influences on sensory processing. For example, in our previous study, using surface recorded potentials and using a nearly identical paradigm (Molholm et al. 2002), we found a series of multisensory interactions with the earliest effects beginning at just 45 ms (also Giard and Peronnet 1999). However, this was over right parieto-occipital scalp and clearly not caused by neural generators in the SPL. Later modulations, however, could very conceivably have included contributions from SPL generators. Thus the SPL integrations seen in this study would have to be considered relatively late in processing and might well have been caused by feedback processes from higher-order regions (see Foxe and Schroeder 2005).

SPL versus IPS?

One might wonder why it is in the SPL that we find colocalization of the maximal auditory and visual responses and ensuing multisensory interactions, when the bulk of the monkey intracranial data points to more inferior and posterior regions within the IPS? Here we present two potential explanations. First, the grid electrodes used here are nestled directly against the surface of the cortex but do not extend down into the sulci. As such, they will be much more sensitive to responses coming from the gyri directly beneath them than to deeper regions within the sulci. Second, it is possible that regions of SPL in the human are in fact the same as those that have been identified in the monkey IPS as these homologies are still in dispute. For example, Watson et al. (1994) point out that Brodmanns area 7, a well-established multisensory region in nonhuman primates (Bremmer 2005; Hyvarinen 1981; Leinonen 1980; Schlack et al. 2005), is below and within the IPS of the monkey, whereas it is significantly more dorsal within the SPL of the human. Bushara et al. (1999) also found co-localization of auditory and visual responses within the SPL during a spatial localization task and similarly reasoned that the SPL might in fact be a more dorsal homologue of monkey IPS. At the same time, although cytoarchitectonic homology is typically thought to be consistent with functional homology (Geyer et al. 2000), they may not necessarily go hand-in-hand (Sereno and Tootell 2005). Brodmann areas 39 and 40 in the inferior parietal lobule, which are unique to humans, correspond more closely in anatomical space to monkey area 7, and it could be argued that in this case human-monkey functional homology is based on anatomical rather than cytoarchitectonic organization.

Latency and amplitude considerations

In all three subjects, auditory responses clearly onset earlier than visual responses and were also greater in amplitude, although it should be pointed out that no systematic manipulation of stimulus intensity was made. Rather, a single moderate intensity level was chosen for each sensory stimulus. Nonetheless, these rather large differences in responses can help us to make some inferences about putative homologies between monkeys and humans. First, the presence of large auditory responses would appear to argue against the lateral intraparietal region (LIP) as a realistic candidate. A number of studies have shown that auditory responsiveness is only found in LIP neurons after animals have learned that specific auditory stimuli are important for oculomotor tasks (Grunewald et al. 1999; Linden et al. 1999; Stricanne et al. 1996). A refinement of this position was made by Gifford and Cohen (2004) who showed that stimulus driven responses could in fact be seen in LIP to unlearned auditory stimuli when monkeys fixated and sounds were presented in a darkened environment such that they were the only salient information present (i.e., there was no distracting visual input). Even when these auditory responses are elicited in LIP, they tend to be weakly stimulus-driven, generally showing lower firing rates than visual neurons do in this region (Cohen et al. 2004; Mullette-Gillman et al. 2005). Also, the proportion of cells in LIP that show auditory stimulus-driven responsiveness is lower than those showing visual responsiveness (Grunewald et al. 1999; Mazzoni et al. 1996; Mullette-Gillman et al. 2005). As such, our finding of more robust auditory responses in this human SPL region, in a task where subjects were not required to make any eye movements at all, renders it somewhat unlikely that these responses originate from a homolog of monkey LIP. A better case for homology can be made for the neighboring ventral intraparietal region (area VIP) where a far greater proportion of cells have been found to be responsive to auditory stimuli and where auditory responses tend to be more robust (Schlack et al. 2005). Schlack et al. found that fully 80% of neurons tested in VIP responded to auditory stimuli, stimuli that had no particular behavioral significance, with 92% of neurons showing visual responsiveness. While there was a tendency for visual responses in bivisual neurons to be stronger, no significant differences in amplitude were seen in the majority of cells recorded from. Nonetheless, the fact that the auditory response was stronger in all three subjects than the visual response also casts some doubt on VIP as a potential homolog and leaves open the distinct possibility that this SPL region is not a homolog of any of the well-characterized monkey IPS regions.

Turning now to the rather substantial difference in onset latencies for auditory and visual responses that we find in the SPL (a mean of 29 vs. 75 ms, respectively). The first thing to note is that these differences seem to largely preserve transmission time differences from the periphery to the primary auditory and visual sensory cortices (A1 and V1). Intracranial recordings of primary auditory cortical responses to tone pips in humans have shown an initial cortical response at just 10 ms (Celesia 1976; Celesia and Puletti 1969, 1971). On the other hand, the onset of visual responses in V1 is seen considerably later, typically between 45 and 60 ms (Clark and Hillyard 1996; Foxe and Simpson 2002; Molholm et al. 2002). Thus there is a 35- to 50-ms delay between the arrival of these two sources of information into their respective primary sensory cortices, much the same as the latency difference observed here in the multisensory zone of the SPL, where there was a mean difference of ~46 ms across the three subjects (range from 38 to 56 ms).

These latency differences also seem to be largely consistent with onset times seen in a number of monkey intracranial studies. Schlack et al. (2005) established a lower limit of 15 ms for auditory inputs to VIP and this agrees very well with our own recordings from monkey LIP where we found both local field potential (LFP) and multiunit onsets in layer IV at just 15
ms (Schroeder et al. 2004). We have also shown average visual onsets in the IPS of awake-behaving macaques at ~28 ms (see Schroeder and Foxe 2002; Schroeder et al. 1998, 2001). A similar onset latency difference is observed between somatosensory and visual input latencies to the IPS with somatosensory inputs arriving at just 10 ms and visual inputs, as above, arriving at ~28 ms (see Fig. 7 of Schroeder and Foxe 2002). Avillac et al. (2005) also found very similar onset latency differences for somatosensory and visual inputs to VIP (10 vs. 40 ms, respectively). In LIP, the story seems to be much the same. Mazzoni et al. (1996) observed auditory and visual onset latency differences of some 30 ms, with a lower limit of about 30 ms for auditory responses and 60 ms for visual responses. It should be mentioned that they also found a very wide range of onset latencies for both sensory inputs, so although the fastest neurons were auditory, many other auditory neurons had quite late onsets such that the median response latency of auditory neurons was somewhat slower than that to visual neurons (155 vs. 125 ms, respectively). However, because the ERP measure used in our study samples net activity from all neurons in a given region, onset latency will be determined by the fastest inputs; thus the differential onset found here seems to be quite consistent with recordings from both LIP and VIP.

Hemispheric specialization?

The electrode sites that met our criteria were all from the left hemisphere. This does not, however, necessarily imply a left lateralized function. Rather this may reflect our restricted sampling of cortical space: In two of our subjects, electrodes were only placed in the left hemisphere, whereas in the remaining, there was bilateral coverage, but it was relatively more sparse. Furthermore, in a high-density scalp ERP study in which an identical paradigm was used, multisensory AV integrations were observed at very similar latencies to those found here. Scalp mapping of these effects showed a strong focus over left parieto-central scalp but also a second focus, albeit weaker, over the right centro-parietal scalp (see Fig. 4, bottom panel of Molholm et al. 2002).

Reaction times and SPL multisensory integration

As in our previous study (Molholm et al. 2002), all three subjects showed reaction times that were faster for multisensory AV stimuli compared with either of the unisensory stimuli. In two of the three subjects, this speeding of reaction time exceeded the so-called race model, indicating that multisensory interactions must necessarily have contributed to the production of responses. However, in the third subject (VH), no violation of the race model was observed. This is of particular interest because VH’s electrophysiological responses in the SPL were identical to the other two subjects, showing the same extent of multisensory integration. Although this is only a single subject, these data seem to suggest that SPL may not be playing a role in the speeding of RT’s typically seen in such paradigms. Clearly, this needs further testing in a larger sample.

In conclusion, clear auditory and visual convergence was seen in a region of the lateral superior parietal lobule while subjects performed a simple reaction time task to randomly intermixed stimuli in both modalities. Auditory inputs were of somewhat greater amplitude and were substantially earlier, by ~40 ms, than visual inputs. Nonetheless, visual responses were also highly robust. Because no manipulation of stimulus intensity was undertaken, the fact that auditory inputs appeared stronger should not be overemphasized. Auditory and visual inputs also showed multisensory integration, relatively early in processing (120–160 ms) within ~60 ms of the initial visual input to this region.

ACKNOWLEDGMENTS

We thank Dr. John Smiley for informative discussions regarding the anatomy of the parietal cortices and J. Mahoney and B. Higgins for expert data collection.

GRANTS

This work was supported by National Institute of Mental Health Grants MH-65350 to J. J. Foxe and F32-MH-068174 to S. Molholm. S. Ortigué was supported by a grant from the Swiss National Foundation for Research in Medicine Biology (FSMB, 1223/PASMA/111563).

REFERENCES


MULTISENSORY INTEGRATION IN HUMAN SPL


