Behavioral/Systems/Cognitive

Spatial Organization of Multisensory Responses in Temporal Association Cortex

Christoph D. Dahl,[1] **Nikos K. Logothetis,**[1,2] **and Christoph Kayser**[1]

[1]Max Planck Institute for Biological Cybernetics, 72076 Tübingen, Germany, and [2]Division of Imaging Science and Biomedical Engineering, University of Manchester, Manchester M13 9PL, United Kingdom

Neurons in sensory cortices are often topographically organized according to their response preferences. We here show that such an organization of response preferences also exists in multisensory association cortex. Using electrophysiological mappings, we probed the modality preference to auditory and visual stimuli of neurons in the superior temporal association cortex of nonhuman primates. We found that neurons preferring the same modality (auditory or visual) often co-occur in close spatial proximity or occur intermingled with bimodal neurons. Neurons preferring different modalities, in contrast, occur spatially separated. This organization at the scale of individual neurons leads to extended patches of same modality preference when analyzed at the scale of millimeters, revealing larger-scale regions that preferentially respond to the same modality. In addition, we find that neurons exhibiting signs of multisensory interactions, such as superadditive or subadditive response summation, also occur in spatial clusters. Together, these results reveal a spatial organization of modality preferences in a higher association cortex and lend support to the notion that topographical organizations might serve as a general principle of integrating information within and across the sensory modalities.

Introduction

Neurons in sensory cortices are often spatially organized according to their response preferences. Such topographical organizations are not only evident in primary cortices but are also present in higher and association regions (Tsunoda et al., 2001; Tanaka, 2003; Op de Beeck et al., 2008). In inferotemporal cortex, for example, the response preferences of visual object-selective neurons are spatially organized in a columnar-like organization (Perrett et al., 1984; Wang et al., 1996; Tamura et al., 2005; Zangenehpour and Chaudhuri, 2005), suggesting that such an organization might be one of the principles underlying the representation and integration of sensory information within a given sensory modality. However, whether this principle also applies to multisensory association areas involved in merging information across different sensory modalities remains an open question.

Electrophysiological mappings in rats have demonstrated that multisensory neurons occur preferentially near the intersections of unisensory cortices (Wallace et al., 2004), and studying the human brain, a high-resolution functional imaging study found that activations to auditory and visual stimuli cluster in separate unisensory patches in the temporal lobe (Beauchamp et al., 2004a,b). Although functional imaging probes neural activity only indirectly (Laurienti et al., 2005; Logothetis, 2008), this nevertheless promotes the hypothesis that a topographical layout of modality preferences might serve as a principle underlying sensory integration. If this was indeed the case, one should be able to

find a spatial organization of modality preferences of individual neurons in typical multisensory association cortices. Here we demonstrate such an organization by investigating the spatial layout of response preferences to auditory and visual stimuli in the macaque monkey superior temporal sulcus.

The upper bank of the superior temporal sulcus (uSTS) is one of the regions in the primate brain frequently implicated in integrating sensory information across modalities (Jones and Powell, 1970; Calvert, 2001; Ghazanfar and Schroeder, 2006; Beauchamp et al., 2008). The uSTS contains an anatomically defined multisensory region, the so-called temporal polysensory area, that receives sensory-related inputs from visual, auditory, and somatosensory cortices (Seltzer and Pandya, 1989; Barnes and Pandya, 1992; Padberg et al., 2003). Neurons in this region respond to stimulation of several modalities, supporting a role in sensory integration (Benevento et al., 1977; Bruce et al., 1981; Baylis et al., 1987; Hikosaka et al., 1988; Barraclough et al., 2005). We assessed the modality preference of neuronal responses in uSTS to auditory and visual stimuli at different spatial scales using systematic electrophysiological mappings and naturalistic stimuli. Our results demonstrate that neurons preferring the same modality preferentially occur in spatial clusters and that these unisensory regions are interspersed with multisensory regions in which neurons show signs of multisensory processing. Altogether, our findings suggest that spatial organizations of neuronal response preferences govern feature integration not only within but also across the sensory modalities.

Materials and Methods

Recording procedures. Two adult rhesus monkeys (*Macaca mulatta*) participated in this study. All procedures were approved by the local authorities (Regierungspräsidium) and were in compliance with the European Community guidelines (EUVD 86/609/EEC) for the care and use of lab-
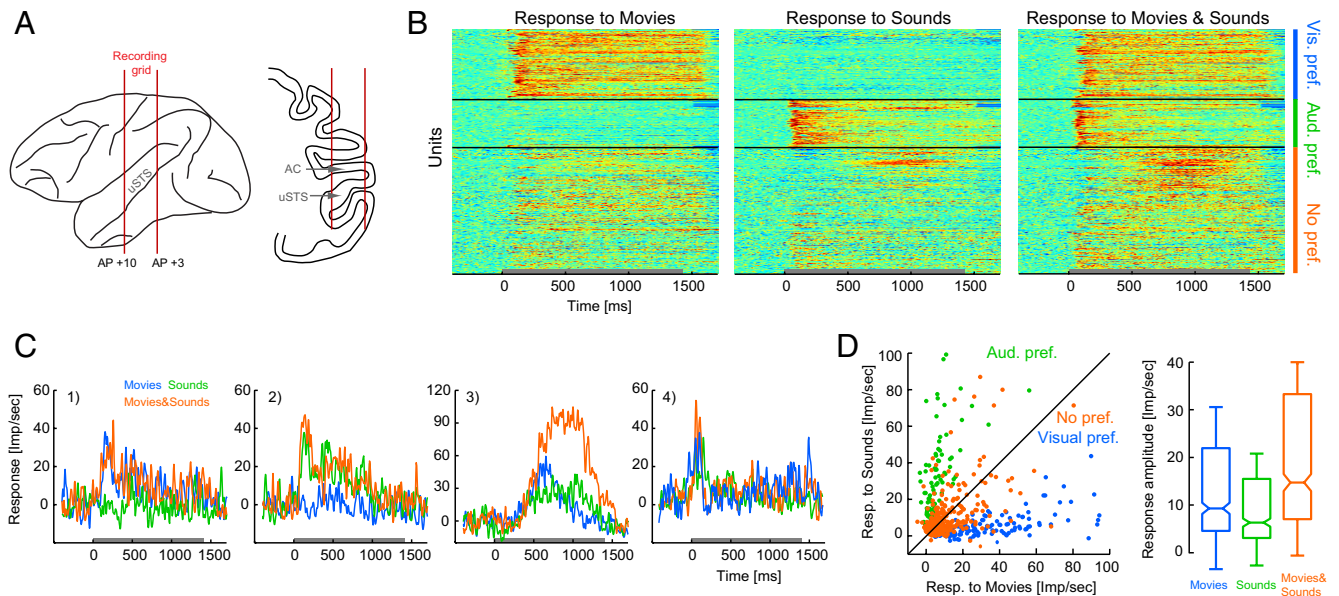
Figure 1. Responses to audiovisual stimuli in the uSTS. ***A***, Recording sites were located in the midportion of the uSTS and were approached vertically using a recording grid. Electrodes first passed through auditory cortex (AC) before reaching uSTS. ***B***, Response time courses of all units to each of the three modality conditions. Responses were normalized by the strongest of the three conditions for display purposes. Stimuli started at 0 ms and lasted 1400 ms (gray bar). Units were sorted according to their modality preference (right color code). ***C***, Example responses of four units (averaged across stimuli). Unit 1 responds to visual but not to auditory stimuli, whereas unit 2 responds to auditory but not to visual. Units 3 and 4 respond to both modalities (bimodal units). ***D***, Left, Response amplitudes to auditory and visual stimuli for all units. Units falling far from the diagonal show a significant preference to either modality, whereas units close to the diagonal are bimodal (response preferences are color coded). Right, Response amplitudes to all three conditions. Boxes denote the 25th and 75th percentile, and the middle line denotes the median. Note the different meaning of the color code in each of the panels. Also note that, to compute response amplitudes, the baseline firing rate of each unit has been subtracted.

oratory animals. All surgical interventions were conducted under general anesthesia and analgesia during an aseptic and sterile procedure (Logothetis et al., 1999). Recording chambers were positioned based on anatomical magnetic resonance (MR) images and stereotaxic coordinates (anteroposterior, +6 mm; mediolateral, +22 mm) (Fig. 1*A*) and were equipped with a plastic grid for systematic electrode positioning. A custom-built electrode drive was used to lower up to six microelectrodes (6 MΩ impedance; FHC) to the STS. Signals were amplified using an Alpha Omega system and digitized at 20.83 kHz. Recordings were performed in a dark and sound attenuating booth (Illbruck Acoustic), while the animals performed a visual fixation task for juice rewards (fixation window: 2° for animal 1, 6° for animal 2). Auditory, visual, or audiovisual stimuli were presented in pseudorandom order for 1.4 s and were preceded by a 500 ms baseline period (silent, neutral gray screen). Visual stimuli were presented by a monitor on a visual field of 24 × 18°. Acoustic stimuli (average intensity of 65 dB sound pressure level) were presented using a Yamaha amplifier (AX-496) and using two free-field speakers (JBL Professional) positioned at ear level 70 cm from the head and 50° to left and right. For additional details on experimental procedures, see Kayser et al. (2007, 2008).

Sensory stimuli. Stimuli were chosen from three categories (four examples each): CS, conspecific communication signals consisting of movies/sounds of other rhesus monkeys producing different vocalizations; NS, scenes of different animals making noises in their natural settings; and AM, artificial motion stimuli consisting of (1) uniformly moving random dots and pulsed (100 ms on, 50 ms off) broad-band (100 Hz to 20 kHz) noise translating in space [intensity linearly increasing from left to right (or vice versa) speaker, 100% intensity modulation in 700 ms], and (2) random dots expanding/contracting and corresponding looming/receding sounds (compare Fig. 4*A*). Stimuli were presented either in a single modality (auditory, visual) or as synchronous bimodal pair (audiovisual). The different modalities and 12 stimulus exemplars were repeated in a pseudorandom order, and each exemplar was usually repeated five times.

Assignment of recording sites. In previous experiments (Kayser et al., 2007, 2008), the auditory cortices of both animals (mostly auditory field A1 and caudal fields CL and CM) have been located, and their frequency organization was systematically mapped. Auditory cortices were located

using MR images, sound frequency maps were constructed for each animal, and core and belt regions were distinguished using the responsiveness for tones and band-passed noise (Recanzone et al., 2000; Lakatos et al., 2005). Having established the depth of auditory cortex along the anteroposterior axis within the recording chamber, we then determined the approximate depth of STS sites from MR images. Functionally, STS sites were identified in each recording sessions based the depth of the electrodes (with regard to the already known depth of auditory cortex), the systematic occurrence of several millimeters (usually >2 mm) of white matter between auditory cortex and STS, and longer response latencies and the prominence of visual responses in the STS. In addition, we performed postmortem histological analysis in one animal, which confirmed the proper location of the electrode tracks in the upper bank STS.

Data analysis. All analysis (except spike sorting) was performed in Matlab (MathWorks). The spike-sorted activity of single units and multiunit clusters was extracted using commercial spike-sorting software (Plexon Offline Sorter) after high-pass filtering the raw signal at 500 Hz. Peristimulus time histograms were obtained using bins of 5 ms and Gaussian smoothing (10 ms full-width at half-height). For many sites, spike sorting could extract single-unit activity. However, for the present analysis, we did not distinguish single and multiunit clusters.

The modality preferences of individual units were computed in several steps. First, the "response amplitude of individual" units was computed using a (100 ms) window centered on the peak response of each unit. The peak response was defined as the time at which the (trial-averaged) response reached its maximum and was computed separately for each stimulus modality. This method was chosen because we used time-varying naturalistic stimuli, and variable preferences of individual units can easily lead to different response time courses (Fig. 1). Results did not change qualitatively when varying this window from 75 to 200 ms. Then, we determined those units that actually responded to sensory stimuli. "Significantly responsive units" were defined as those for which, in at least one modality condition, the response amplitude during stimulation differed significantly from the amplitude in a baseline window (established using a t test, $p < 0.05$; baseline window defined as a 100 ms window during the prestimulus period, starting 400 ms before stimulus onset). This criterion was found to perform similarly to one based on SDs
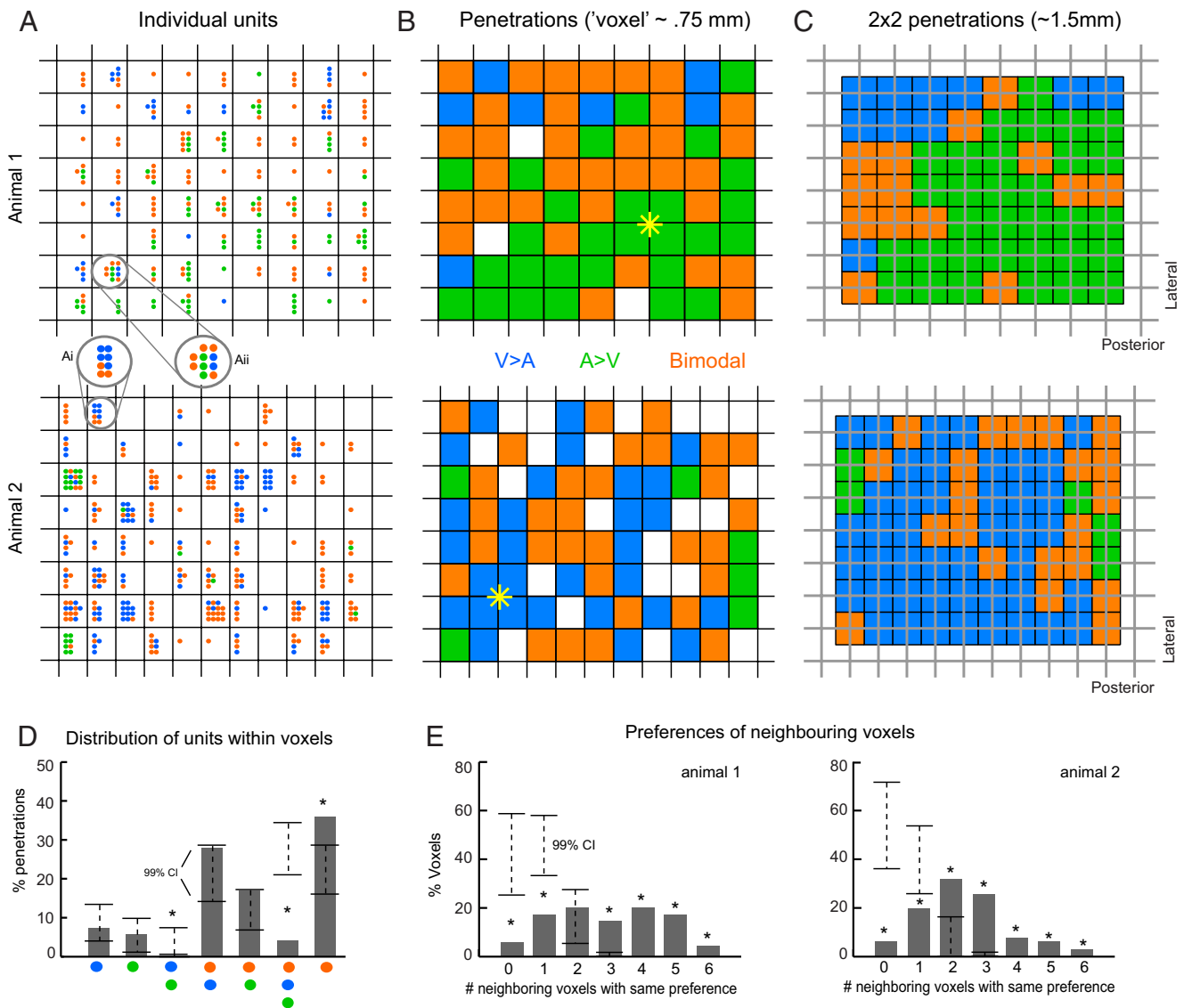
Figure 2. Spatial organization of modality preferences. ***A***, Individual units are color coded according to their modality preference and are shown superimposed on the recording grid of each animal. In the grid, each box denotes one electrode penetration (750 μm spacing), and open boxes did not yield sensory-responsive units. ***Ai*** provides an example of a typical penetration, in which unimodal and bimodal were encountered together. ***Aii*** denotes one of the five penetrations in which units preferring visual and units preferring auditory stimuli were encountered together. ***B***, Modality selectivity at the spatial scale of penetrations (also called "voxels"). One selectivity value was obtained by pooling the responses of all units along each penetration using an ANOVA. Penetrations with a significant modality preference ($p < 0.01$) are labeled by their preferred modality or as bimodal. Yellow asterisks denote continuous regions of same preference. ***C***, Modality selectivity at the spatial scale of 2 × 2 penetrations (1.5 mm spacing). One selectivity value was obtained by pooling the responses of all units along 2 × 2 neighboring penetration using an ANOVA, similar as in ***B***. These virtual 2 × 2 voxels were centered on the edge of the recording grid. ***D***, Frequency of different combinations of modality preferences encountered along the same penetrations. Colored dots denote the different combinations of unit preferences, bars denote their frequency across animals, and error bars denote the 99% confidence intervals obtained from a randomization test (assuming a random distribution of units). Bars falling outside the confidence intervals (marked by *) indicate modality distributions that deviate from chance and hence indicate a true spatial pattern. This analysis reveals that combinations of different modality-selective units (blue and green) are less frequent than expected, whereas penetrations with only bimodal units are more frequent than expected. ***E***, Frequency of common modality preferences of neighboring voxels. Bars denote the frequency for different numbers of neighbors with same preference, and error bars denote the 99% confidence interval obtained from a random distribution of units. This reveals that neighboring voxels are more likely to share the same preference as expected.

from baseline variability, as used in previous studies (Kayser et al., 2007, 2008). For additional analysis, only units with significant responses to at least one modality were included. Then, the "modality preference" of individual units was determined by comparing the response amplitudes to visual and auditory modalities using an ANOVA (including stimuli as repeats of the same modality and modalities as factor). Units with a significant ($p < 0.01$) modality effect were labeled according to the modality eliciting the stronger response (Fig. 2A), whereas units with insignificant modality effect were labeled bimodal. In addition, there were some units that did not respond significantly to either visual or auditory stimuli but responded only in the audiovisual condition compared with baseline. These units were included in the group of bimodal

units. The modality preference at larger scales was determined by pooling the responses of units in an ANOVA (i.e., treating units and stimuli as repeats of the same modality, using modalities as factor). In this way, the modality preference at larger scales was computed from the mean response of all included neurons but taking the variability between neurons into account. This analysis was repeated by pooling all units along each penetrations (scale of 750 μm) or by pooling all units of groups of 2 × 2 neighboring penetrations (scale of 1.5 mm).

"Multisensory influences" in the responses of individual units were characterized using an established criterion (Stanford et al., 2005; Avillac et al., 2007; Kayser et al., 2008): the linearity index compares the bimodal response with a linear superposition of the two unimodal responses [here

done using a randomization procedure (Stanford et al., 2005)]. Given the large number of tests performed (>500 units), we used the false discovery rate to correct for multiple comparisons (Benjamini and Hochberg, 1995). The strength of the multisensory interaction was determined using the normalized difference between the bimodal and the sum of the unimodal responses: $[AV - (A + V)]/[AV + (A + V)] * 100$ (Kayser et al., 2008), where A and V are audio and visual unimodal responses, and AV is the audiovisual bimodal response.

The "stimulus type preference" of individual units was determined by comparing the peak responses (averaged over the four exemplars of each stimulus type) and labeling each unit by the type eliciting the strongest response (see Fig. 4C). The "stimulus selectivity" of individual units was determined as follows: for each unit, we sorted the three stimulus types for increasing responses and normalized the resulting graph by the strongest response (see Fig. 4A). The resulting graph indicates how much responses differ between optimal and non-optimal stimuli. An interaction of modality selectivity and stimulus type was assessed using an ANOVA (with modalities and stimulus type as factors).

Statistical analysis of spatial patterns. To determine whether the spatial patterning of modality preferences (or location of units with significant multisensory influences) in the actual data could arise by chance, we performed statistical randomization tests as follows (Efron and Tibshirani, 1994; Sokal and Rohlf, 1995). First, we established an index of interest. This could be the frequency of how often units with the same or different modality preferences occur along the same penetrations (Fig. 2D), the number of neighboring penetrations with same modality preference (Fig. 2E), and the number of units with significant multisensory effects in neighboring penetrations (see Fig. 5C). Then, we randomized the assignment of units to individual penetrations, by sampling without replacement from the total set of units. Importantly, this procedure not only preserved the distribution of units with different properties but also the number of units per penetration. This randomization was performed 1000 times, and, for each randomized dataset, we computed the index of interest. Finally, we computed the 99% confidence intervals from the randomized dataset and compared the actual values with these.

Results

Sensory responses: examples and response amplitudes
We recorded from a total of 532 sensory responsive (single and multi) units in the uSTS of two animals (311 in animal 1, 221 in animal 2). Individual microelectrode penetrations were systematically arranged using a grid along the mediolateral and anteroposterior axes and were focused on the midportion of the uSTS region (Fig. 1A). Sensory responses to naturalistic time-varying audiovisual stimuli were recorded while the animals performed a visual fixation task.

Naturalistic audiovisual stimuli elicited robust responses throughout the sampled region, with some units responding to only one and others responding to both modalities. This is visible in the graphs showing the (normalized) response time courses for each of the units (Fig. 1B) and in the example units (Fig. 1C). Of these examples, units 1 and 2 showed a clear preference toward one modality (unit 1 for visual and unit 2 for auditory stimuli), whereas units 3 and 4 responded to both modalities with comparable strength. To quantify such modality preference, we first computed the response amplitudes for each unit and sensory condition (Fig. 1D) and then compared the responses to both modalities using an ANOVA: units for which responses (across trials and stimuli) to auditory and visual stimuli differed significantly ($p < 0.01$) were labeled modality selective, whereas units with comparable responses to both modalities were labeled bimodal (no effect of the factor stimulus modality). In Figure 1B, units have been sorted according to this modality preference, which is also marked by the same color code used in D. Overall, this revealed that the large majority of units responded similarly

to both modalities (i.e., bimodal units, 53.2%), whereas fewer preferred visual (28.4%) or auditory (18.4%) stimuli. Across the population of neurons, the response to visual [9.3 impulses per second (Imp/s) above baseline, median value] (Fig. 1D) was stronger than to auditory (6.3 Imp/s; Wilcoxon's rank-sum test, $p < 10^{-4}$) stimuli, but responses to audiovisual stimuli were significantly stronger than both unimodal responses (14.7 Imp/s; Wilcoxon's rank-sum tests, both $p < 10^{-9}$), demonstrating the multisensory nature of the uSTS region.

Spatial organization of modality preferences: individual units
To determine whether the modality preference of individual neurons is spatially organized, we compared preferences of units recorded along the same or neighboring electrode penetrations. Individual penetrations were systematically spaced (750 μm) on a recording grid, and often several units were recorded at different depths of the same penetration. Figure 2A displays the modality preferences for all units in each of the two animals.

At first sight, there seems to be little organization of modality selectivity at the scale of individual units (Fig. 2A). Along many penetrations, bimodal units (orange) intermingle with modality-preferring units (green or blue, as exemplified in Ai). However, closer inspection revealed a striking result: of 136 penetrations (70 in animal 1, 66 in animal 2), only five (3.6%) contained units preferring the auditory and units preferring the visual modality at the same time (one highlighted in Aii). This demonstrates that neurons preferring different modalities only rarely occur along the same penetration but occur only spatially separated. To quantify this observation, Figure 2D displays the frequency of different combinations of modality preferences encountered along individual penetrations: most penetrations yielded (1) only units preferring the same modality, (2) modality-preferring and bimodal units together, or (3) only bimodal units. To determine which combinations of modality preferences could arise by chance, we used a randomization procedure to shuffle the assignment of units to individual penetrations and computed confidence intervals from 1000 such randomizations. This confirmed that the likelihood of finding two units with distinct modality preference in the actual data was significantly lower than chance ($p < 0.01$), whereas the likelihood of finding penetrations with only bimodal units was higher than chance ($p < 0.01$) (see confidence intervals in Fig. 2D). Overall, this demonstrates that the modality preference in uSTS is indeed spatially organized, with unimodal and bimodal neurons co-occurring along the same penetrations but neurons preferring distinct modalities being spatially separated.

Spatial organization of modality preferences: scale of penetrations
Next we determined the modality preference at the scale of penetrations (750 μm). By pooling all responses recorded along a penetration in an ANOVA, we obtained one modality preference for each penetration (termed "voxel" in the following, because it represents the aggregate response of all units along a penetration; see also Materials and Methods). For pooling, we included the responses of all units in the ANOVA used to determine the modality selectivity (hence including the variability between neurons in the estimate of the reliability of modality preferences). Noteworthy, at this scale, a structured organization was evident in both animals (Fig. 2B): many neighboring voxels share the same modality preference (around the yellow asterisks in Fig. 2B), and continuous regions of the same preference are apparent in the figures.

To test whether this spatial organization could arise by chance, we constructed an index capturing the number of neighboring voxels with same modality preference. Then we compared the actual index with a distribution of indices obtained from spatially randomized data (Fig. 2E). If the spatial organization was indeed random, the distribution of neighboring voxels with same modality preference would not differ from confidence intervals obtained from randomized data. However, in both animals, this was not the case. The likelihood of encountering several (more than three) neighboring voxels with the same modality preference was significantly higher than expected ($p < 0.01$) (see confidence intervals in the figure). This result demonstrates that a spatial organization of modality preference also prevails at the spatial scale of neighboring penetrations and, hence, at a spatial scale of several hundreds of micrometers.

Spatial organization of modality preferences: scale of millimeters

Because previous evidence for a spatial organization of modality preferences in the STS was obtained using functional imaging at resolution of millimeters (Beauchamp et al., 2004b), we extended our analysis to this scale. Responses were pooled across all units recorded on 2 × 2 neighboring penetrations (again using an ANOVA), to yield one modality preference at the scale of 1.5 mm (Fig. 2C). At this large scale, coherent regions of the same modality preference emerged in both animals. As above, we used a randomization test to assess the significance of this finding: the test confirmed that this spatial organization differed from chance and that these 2 × 2 voxels were more likely to have neighbors with the same preference than expected (supplemental Fig. 1, available at www.jneurosci.org as supplemental material).

In addition, we also verified the robustness of these maps using a split dataset approach [test–retest approach (cf. Beauchamp et al., 2004b)]. For each unit, the stimulus set was randomly split into two halves, and the responses and modality preference maps were computed for each half. A correlation analysis was then used to determine the similarity of the two resulting preference maps. The correlations were significant for both animals (animal 1, $r = 0.38$, $p < 0.001$; animal 2, $r = 0.33$, $p < 0.01$), demonstrating that modality maps are robust to the selection of a subset of the data.

Because these modality preferences, and the resulting spatial layouts, were derived only from responses to auditory and visual stimuli, it is important to confirm that the responses within these voxels behave as expected: visual voxels responded more strongly to visual than to auditory stimuli (27.1 ± 1.7 vs 6.8 ± 0.8 Imp/s, mean $\pm$ SEM; two-sided paired t test, $p < 10^{-10}$) and vice versa for auditory voxels (6.5 ± 0.6 vs 25.6 ± 1.1 Imp/s; $p < 10^{-10}$) (Fig. 3A). Importantly, for both kinds of voxels, audiovisual responses were comparable with responses to the preferred modality ($p = 0.85$ and $p = 0.44$, respectively), confirming the unimodal character of these voxels. Bimodal voxels, in contrast, responded similarly to both unimodal conditions (mean $\pm$ SEM, 13.0 ± 1.0 and 10.6 ± 1.1 Imp/s; $p = 0.37$) and significantly stronger to audiovisual stimuli (16.5 ± 1.3 Imp/s; $p = 0.01$ vs visual and $p = 0.0013$ vs auditory responses). Noteworthy, this audiovisual response in the bimodal voxels represented a subadditive superposition of response to unimodal visual and auditory stimuli (audiovisual, 16.5 vs auditory + visual, 23.6). As a result, cross-modal interaction, a typical sign of sensory integration, was only present in bimodal but not in unimodal regions.

Figure 3B displays how these voxel-based (large-scale) response properties relate to the distribution of individual units within these voxels. Most units in visual voxels preferred visual
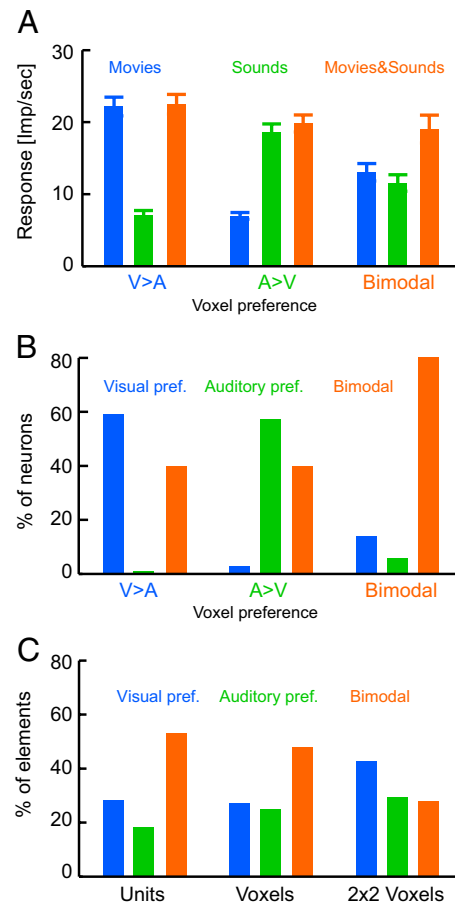

Figure 3. Properties of units within different penetrations (voxels). *A*, Average responses to the three conditions within voxels of different preference. Bars denote the mean and SEM, pooled over units and animals. *B*, Distribution of units within voxels. Bars indicate the frequency of units with different preferences in each kind of voxel (pooled over animals). *C*, Frequency of different modality preferences across spatial scales (units, penetrations, and 2 × 2 penetrations; data pooled over animals).

stimuli or were bimodal, whereas auditory-preferring units did not occur in visual voxels. The converse was true for auditory voxels, and bimodal voxels were mostly occupied by bimodal neurons. This further strengthens the notion that regions preferring one sensory modality arise from a mixture of modality-preferring and bimodal neurons, whereas regions without modality preference are mostly occupied by bimodal neurons.

Noteworthy, the fraction of bimodal units/voxels decreases with increasing spatial scale (Fig. 3C). At the smallest scale, bimodal units prevail over unimodal units, whereas at the scale of millimeters, modality-preferring voxels are more frequent. This results in a significant difference between spatial scales (χ^2 test comparing frequency of 2 × 2 voxels to frequency expected based on distribution of units, $\chi^2 = 12$; $p < 0.01$) and lets us conclude that modality-selective patterns at larger scales emerge from weakly modality-selective but topographically arranged neurons at the small spatial scale.

Stimulus selectivity

Neurons in the uSTS respond to complex objects, and preferences to motion- and action-related stimuli (Anderson and Siegel, 1999; Barraclough et al., 2005), as well as to faces and body parts, have been reported (Allison et al., 2000; Barraclough et al., 2005; Ghazanfar et al., 2008; Tsao and Livingstone, 2008). To account
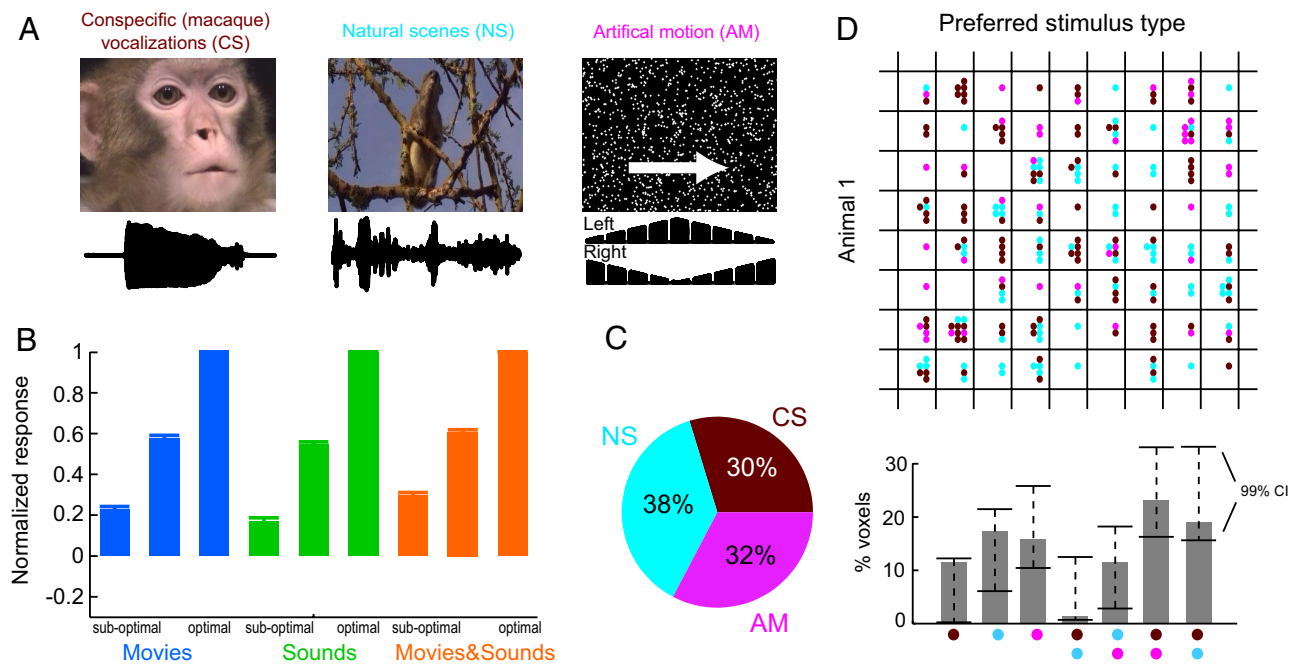
Figure 4. Responses to different types of stimuli. **A**, Example frames and sound waves of the three types of stimuli used. NS, Natural settings; CS, conspecific communication sounds; AM, artificial motion patterns. **B**, Many units responded selectively to only some stimuli. The graph displays the responses of individual units sorted according to increasing efficacy of the stimulus type (from left to right). Responses were normalized by the most efficient stimulus type to facilitate comparison across units. Bars denote mean and SEM across units. **C**, Frequency of preferred stimulus types across units. **D**, Spatial distribution of the preferred stimulus type of individual units, superimposed on the recording grid of one animal. Neighboring units along the same penetration often have heterogeneous preferences. The bottom displays the frequency of different combinations of stimulus preferences encountered along the same penetration for this animal. Colored dots denote the different combinations of preferences. Error bars indicate the 99% confidence intervals obtained from a randomization test. The actual and randomized data do not differ, as visible by the overlap of bars and confidence intervals.

for this range of preferences, our stimuli comprised scenes of conspecific vocalizing animals (CS), scenes of other animals in natural settings (NS), and artificial motion patterns (AM) (Fig. 4A). We investigated whether individual units show selectivity to these stimulus types and whether the spatial organization of modality preference depends on the stimulus type used.

Individual units typically responded stronger to one than to other stimuli, as demonstrated in Figure 4B. This graph reveals how responses differ between "optimal" and "suboptimal" stimulus types, by displaying them normalized by the strongest (optimal) response and sorted by increasing efficacy. For 38% of the units, the response to the intermediate type was <50% of the response to the optimal category (averaged across modalities of presentation), indicating considerable selectivity to individual stimuli. Across the population, however, the different stimulus categories were rather balanced, with a small majority of units preferring natural scenes over the other stimuli (Fig. 4C). This confirms that uSTS units respond to diverse stimuli ranging from artificial motion patterns to behaviorally relevant communications sounds, with different neurons preferring different stimuli.

These findings raise the question whether the spatial layout of modality preferences depends on the stimulus type. To directly address this, we followed the strategy used by a previous study on modality selectivity in the STS (Beauchamp et al., 2004b): we computed the modality preference at the scale of penetrations (750 μm voxels) for each stimulus type and used an ANOVA to reveal any interaction between modality preference and stimulus type (with modality preference and stimulus type as factors and voxels as elements). In visual ($F_{(2,1871)} = 4.9$, $p = 0.007$) and bimodal ($F_{(2,1565)} = 3.0$, $p = 0.045$) but not in auditory ($F_{(2,1349)} = 0.7$, $p = 0.46$) voxels, there was an overall effect of stimulus type, in agreement with a slight bias toward the natural settings (cf. Fig.

4C). However, there was no interaction between stimulus type and modality preference in any of the voxels (visual voxels, $F_{(4,1871)} = 2.3$, $p = 0.055$; auditory, $F_{(4,1349)} = 0.61$, $p = 0.65$; bimodal, $F_{(4,1565)} = 1.3$, $p = 0.23$). All in all, this demonstrates that the spatial topography of preferred modality does not depend on the stimulus type.

The observation that stimulus selectivity was independent of modality preference does by itself not allow definite conclusion about any spatial organization of preferences to particular kinds of stimuli. To determine whether neurons preferring the same kind of stimulus might cluster (in a manner unrelated to the modality preference), we displayed the preferred stimulus type for each unit. Again we used a permutation procedure to determine whether the distribution of preferred stimuli differs from a random pattern, using the frequency of different combinations of preferences along individual penetrations as index. In both animals (one shown in Fig. 4D), there was a large overlap of the actual and randomized ($p > 0.05$) distributions, leading us to conclude that, although modality preferences follow a spatial patterning in the uSTS, the preference for those types of stimuli as analyzed here does not.

Indices of multisensory interactions

To probe individual uSTS units for multisensory interactions, we used a frequently used index, the so-called linearity index (Stanford et al., 2005; Avillac et al., 2007). This probes whether the response in the bimodal condition deviates from a linear superposition of the two unimodal responses and hence reveals whether the different modalities interact in eliciting the bimodal response. For most units, the response in the bimodal condition was comparable with the sum of the two unimodal responses, as revealed by the scatter plot in Figure 5A. Less than one-quarter of

units (19%) exhibited significant (at $p < 0.05$ corrected for false discovery rates) multisensory interactions, with some units showing supralinear (AV > A + V) and other units showing sublinear responses (AV < A + V). This heterogeneity of response summation is also confirmed by the distribution of interaction indices (Fig. 5*B*). Across all units, the interaction was biased toward sublinear summation, but, of the units with significant effects (red), some were supralinear and others sublinear.

Noteworthy, units with significant multisensory effects often occurred along the same or neighboring penetrations (Fig. 5*C*). The top displays the recording location of units with significant interactions for one of the animals. A randomization test revealed that the number of units with significant multisensory interactions in neighboring voxels deviates significantly from chance ($p < 0.01$), and this was the case for the other animal as well. Although a random distribution of units would lead to few incidences of multisensory effects in the same or neighboring voxels, the actual numbers exceeded those expected by chance. These results demonstrate that units with multisensory response interactions also occur spatially clustered, hence forming "hotspots" of multisensory interactions.

Discussion

Our results show that modality preferences of neurons in the upper bank STS are spatially organized: units preferring the same modality often co-occur in close spatial proximity, or occur in proximity with bimodal neurons, whereas units preferring different modalities are spatially separated. This topographical organization at the small scale leads to extended patches of same modality preference at the scale of millimeters, showing that spatial organizations of response preferences are not limited to unisensory regions but also occur in multisensory association cortices.

Spatial organization of neuronal preferences in temporal association cortex

The spatial organization of modality preferences reported here fits with the general notion that neuronal feature selectivity in sensory cortices follows a columnar-like organization. Such an organization is widespread in early sensory regions but has also been reported in higher visual areas, such as in the temporal lobe (Tsunoda et al., 2001; Op de Beeck et al., 2008). Indeed, studies using electrophysiological or optical methods revealed similar feature preferences of neighboring neurons in inferotemporal cortex (Perrett et al., 1984; Wang et al., 1996; Tamura et al., 2005; Zangenehpour and Chaudhuri, 2005; Kreiman et al., 2006). The working hypothesis emerging from this is that neurons are organized according to some "critical features," which are shared across neurons within a couple of hundred micrometers (Sato et al., 2009).

The spatial organization of modality preferences found here could reflect a generalization of this principle from unisensory to multisensory cortices. With the preferred sensory modality being the critical feature, the spatial modality layout reflects the same
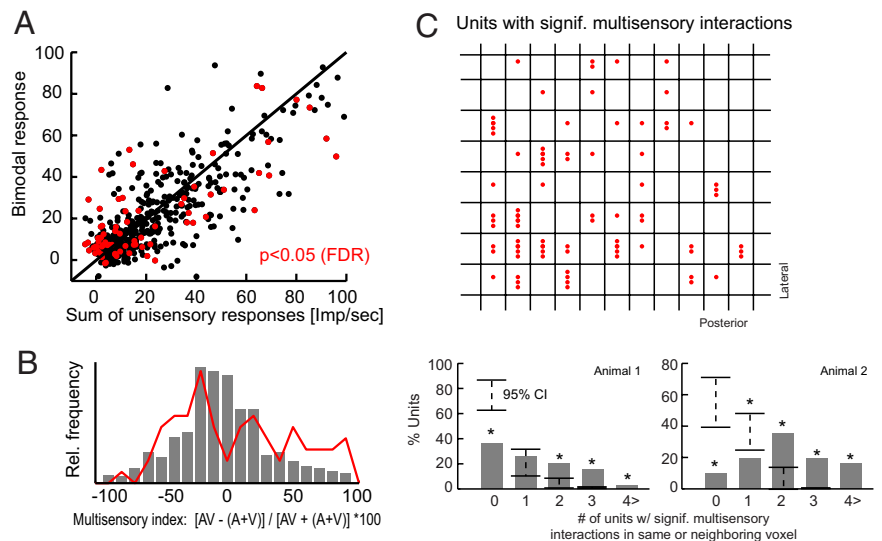

Figure 5. Significant multisensory response interactions. ***A***, Comparison of responses in the bimodal condition (AV, movie and sounds) to the sum of the responses in the two unimodal conditions (A + V). Red dots denote units that significantly deviate from a linear superposition [$p < 0.05$, corrected for false discovery rate (FDR)] and hence show significant multisensory interactions. ***B***, Distribution of the linearity index across units. Gray bars denote the values for all units; the red line denotes values for those units with significant multisensory interactions. Both distributions have been normalized to a total probability of one. ***C***, Top, Spatial distribution of units with significant multisensory interactions (red dots) for one animal. Bottom, Frequency of units with significant effects in the same or neighboring voxels. Bars denote the actual frequency, and error bars denote the 99% confidence interval obtained from a random distribution of units. Bars falling outside the confidence intervals (*) indicate distributions that deviate from chance and hence indicate a spatial pattern. Overall, the chance of finding units with significant multimodal effects in neighboring penetrations was significantly higher than expected by chance.

organizational principle underlying the columnar-like organization in unisensory cortices. Although both unisensory and multisensory organizations and their relation to the formation of sensory representations require additional analysis, the general notion that topographical organizations govern feature integration within and across sensory modalities is very appealing. Future studies, for example, could test this notion in other association cortices such as the prefrontal cortex, in which auditory and visual preferring neurons similarly co-occur with neurons integration information from both modalities (Romanski, 2004).

Response properties of neurons in the STS

The multisensory nature of the upper bank STS region has been ascertained in many functional imaging studies (for review, see Calvert, 2001; Amedi et al., 2005; Ghazanfar and Schroeder, 2006). However, given the indirect coupling of functional magnetic resonance imaging (fMRI)–blood oxygenation level-dependent (BOLD) and neural response, imaging studies cannot reveal the response properties of individual neurons or their spatial organization (Logothetis, 2008), and hence methods with higher spatial resolution and markers directly coupling to neuronal activity are required to address these questions.

Previous studies have characterized neurons in the uSTS region as responding preferentially to stimuli containing biological or artificial motion or containing face and object stimuli (Bruce et al., 1981; Oram and Perrett, 1996; Anderson and Siegel, 1999; Allison et al., 2000; Barraclough et al., 2005; Ghazanfar et al., 2008; Tsao and Livingstone, 2008). Our findings agree well with this characterization and revealed similar fractions of unimodal and bimodal units and units with significant multisensory interactions as reported previously (Benevento et al., 1977; Bruce et al., 1981; Baylis et al., 1987; Hikosaka et al., 1988; Barraclough et

al., 2005). However, the present is (to our knowledge) the first study to systematically investigate the spatial layout of auditory and visual response preferences at the same time. Although neuroanatomy has revealed afferent projections that could mediate a spatial organization of modality preferences (Seltzer and Pandya, 1994; Seltzer et al., 1996), experimental evidence at the neural level was missing so far (but see Hikosaka et al., 1988 for some anecdotal evidence).

In fact, our findings might provide a direct neural "explanation" for the observation of a "patchy" organization in the human STS by a previous human imaging study by Beauchamp et al. (2004b). These authors found a spatial patterning of voxels in which the fMRI–BOLD signal was preferentially activated by visual or auditory stimuli or that was activated by both modalities. These patches were of the scale of several millimeters and independent of the stimulus category used. The spatial organization of neuronal preferences and the resulting large-scale modality preferences found here might well be the neural substrate that elicited the respective BOLD activations seen in the imaging study.

Sensory integration and the organization of modality preferences

Our results provide evidence that a topographical organization might underlie the merging of sensory information in association cortex. Those neurons typically implicated in the process of sensory integration, such as neurons responding to several modalities or neurons with multisensory response interactions, occurred in spatial clusters that are interspersed with unisensory and modality-preferring regions. Hence, the typical signs of sensory integration were not distributed uniformly but confined to specific regions.

The notion that a topographical organization of modality representations might serve as principle underlying the merging of sensory information has been suggested based on two observations: Beauchamp et al. (2004b) imaged a millimeter-scale patchy organization of auditory- and visual-preferring voxels in the human STS, and Wallace et al. (2004) found that multisensory neurons in the rat brain are most frequent at the borders or intersections of unisensory regions. Our results provide strong experimental evidence for the existence of such an organization.

However, what remains unclear is the function or role of a spatial organization in the process of sensory integration. On the one hand, it might be that the spatial organization simply reflects the pattern of anatomical afferents (Seltzer and Pandya, 1994; Seltzer et al., 1996) but does not have an immediate functional implication. On the other hand, it might also be that a topographical arrangement of modality preferences facilitates the process of sensory integration in some respect. Because sensory integration is especially helpful in conditions in which unisensory responses are impoverished, integration serves to stabilize perception against external noise (Stein and Meredith, 1993; Ernst and Bülthoff, 2004). One way to stabilize neuronal representations might be to introduce redundancy, which could for example be achieved by spatially distributing neurons with similar preferences. Clearly, this is only a vague idea, and much future thinking and work will be required to understand the use of topographical originations in unisensory and multisensory processing.

Predictions for future work

Our data reveals that the patchy organization of modality preferences arises from the particular arrangement of individual neurons: patches preferring one modality arise from a mixed

population of units preferring this modality and modality-unselective units (i.e., bimodal), whereas bimodal patches contain mostly bimodal neurons. Hence, a bias in the distribution of individual units leads to a modality preference at the large scale. Importantly, we found that the fraction of modality-unselective responses was much larger at the scale of units than at the scale of millimeters. This suggests that techniques that average responses over large spatial regions (e.g., low-resolution functional imaging) might well underestimate the contribution of multisensory responses and provide a rather conservative estimate of which brain regions contain neurons participating in sensory integration. One possibility to overcome this limitation might be to exploit adaptation paradigms, which in principle can reveal the neuronal composition of individual voxels (Bartels et al., 2008; Goebel and van Atteveldt, 2009). Using such adaptation paradigms, one might be able to elucidate unimodal and bimodal response properties of image voxels in the STS and confirm the prediction that many apparently unimodal voxels contain bimodal neurons.

References

Allison T, Puce A, McCarthy G (2000) Social perception from visual cues: role of the STS region. Trends Cogn Sci 4:267–278.

Amedi A, von Kriegstein K, van Atteveldt NM, Beauchamp MS, Naumer MJ (2005) Functional imaging of human crossmodal identification and object recognition. Exp Brain Res 166:559–571.

Anderson KC, Siegel RM (1999) Optic flow selectivity in the anterior superior temporal polysensory area, STPa, of the behaving monkey. J Neurosci 19:2681–2692.

Avillac M, Ben Hamed S, Duhamel JR (2007) Multisensory integration in the ventral intraparietal area of the macaque monkey. J Neurosci 27:1922–1932.

Barnes CL, Pandya DN (1992) Efferent cortical connections of multimodal cortex of the superior temporal sulcus in the rhesus monkey. J Comp Neurol 318:222–244.

Barraclough NE, Xiao D, Baker CI, Oram MW, Perrett DI (2005) Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. J Cogn Neurosci 17:377–391.

Bartels A, Logothetis NK, Moutoussis K (2008) fMRI and its interpretations: an illustration on directional selectivity in area V5/MT. Trends Neurosci 31:444–453.

Baylis GC, Rolls ET, Leonard CM (1987) Functional subdivisions of the temporal lobe neocortex. J Neurosci 7:330–342.

Beauchamp MS, Lee KE, Argall BD, Martin A (2004a) Integration of auditory and visual information about objects in superior temporal sulcus. Neuron 41:809–823.

Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A (2004b) Unraveling multisensory integration: patchy organization within human STS multisensory cortex. Nat Neurosci 7:1190–1192.

Beauchamp MS, Yasar NE, Frye RE, Ro T (2008) Touch, sound and vision in human superior temporal sulcus. Neuroimage 41:1011–1020.

Benevento LA, Fallon J, Davis BJ, Rezak M (1977) Auditory–visual interaction in single cells in the cortex of the superior temporal sulcus and the orbital frontal cortex of the macaque monkey. Exp Neurol 57:849–872.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 57:289–300.

Bruce C, Desimone R, Gross CG (1981) Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. J Neurophysiol 46:369–384.

Calvert GA (2001) Crossmodal processing in the human brain: insights from functional neuroimaging studies. Cereb Cortex 11:1110–1123.

Efron B, Tibshirani RJ (1994) An introduction to the bootstrap. New York: Chapman and Hall/CRC.

Ernst MO, Bülthoff HH (2004) Merging the senses into a robust percept. Trends Cogn Sci 8:162–169.

Ghazanfar AA, Schroeder CE (2006) Is neocortex essentially multisensory? Trends Cogn Sci 10:278–285.

Ghazanfar AA, Chandrasekaran C, Logothetis NK (2008) Interactions be-

tween the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. J Neurosci 28:4457–4469.

Goebel R, van Atteveldt N (2009) Multisensory functional magnetic resonance imaging: a future perspective. Exp Brain Res 198:153–164.

Hikosaka K, Iwai E, Saito H, Tanaka K (1988) Polysensory properties of neurons in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. J Neurophysiol 60:1615–1637.

Jones EG, Powell TP (1970) An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. Brain 93:793–820.

Kayser C, Petkov CI, Logothetis NK (2007) Tuning to sound frequency in auditory field potentials. J Neurophysiol 98:1806–1809.

Kayser C, Petkov CI, Logothetis NK (2008) Visual modulation of neurons in auditory cortex. Cereb Cortex 18:1560–1574.

Kreiman G, Hung CP, Kraskov A, Quiroga RQ, Poggio T, DiCarlo JJ (2006) Object selectivity of local field potentials and spikes in the macaque inferior temporal cortex. Neuron 49:433–445.

Lakatos P, Pincze Z, Fu KM, Javitt DC, Karmos G, Schroeder CE (2005) Timing of pure tone and noise-evoked responses in macaque auditory cortex. Neuroreport 16:933–937.

Laurienti PJ, Perrault TJ, Stanford TR, Wallace MT, Stein BE (2005) On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies. Exp Brain Res 166:289–297.

Logothetis NK (2008) What we can do and what we cannot do with fMRI. Nature 453:869–878.

Logothetis NK, Guggenberger H, Peled S, Pauls J (1999) Functional imaging of the monkey brain. Nat Neurosci 2:555–562.

Op de Beeck HP, Dicarlo JJ, Goense JB, Grill-Spector K, Papanastassiou A, Tanifuji M, Tsao DY (2008) Fine-scale spatial organization of face and object selectivity in the temporal lobe: do functional magnetic resonance imaging, optical imaging, and electrophysiology agree? J Neurosci 28:11796–11801.

Oram MW, Perrett DI (1996) Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. J Neurophysiol 76:109–129.

Padberg J, Seltzer B, Cusick CG (2003) Architectonics and cortical connections of the upper bank of the superior temporal sulcus in the rhesus monkey: an analysis in the tangential plane. J Comp Neurol 467:418–434.

Perrett DI, Smith PA, Potter DD, Mistlin AJ, Head AS, Milner AD, Jeeves MA (1984) Neurones responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception. Hum Neurobiol 3:197–208.

Recanzone GH, Guard DC, Phan ML (2000) Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. J Neurophysiol 83:2315–2331.

Romanski LM (2004) Domain specificity in the primate prefrontal cortex. Cogn Affect Behav Neurosci 4:421–429.

Sato T, Uchida G, Tanifuji M (2009) Cortical columnar organization is reconsidered in inferior temporal cortex. Cereb Cortex 19:1870–1888.

Seltzer B, Pandya DN (1989) Frontal lobe connections of the superior temporal sulcus in the rhesus monkey. J Comp Neurol 281:97–113.

Seltzer B, Pandya DN (1994) Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: a retrograde tracer study. J Comp Neurol 343:445–463.

Seltzer B, Cola MG, Gutierrez C, Massee M, Weldon C, Cusick CG (1996) Overlapping and nonoverlapping cortical projections to cortex of the superior temporal sulcus in the rhesus monkey: double anterograde tracer studies. J Comp Neurol 370:173–190.

Sokal RR, Rohlf FJ (1995) Biometry. New York: Freeman and Company.

Stanford TR, Quessy S, Stein BE (2005) Evaluating the operations underlying multisensory integration in the cat superior colliculus. J Neurosci 25:6499–6508.

Stein BE, Meredith MA (1993) Merging of the senses. Cambridge, MA: MIT.

Tamura H, Kaneko H, Fujita I (2005) Quantitative analysis of functional clustering of neurons in the macaque inferior temporal cortex. Neurosci Res 52:311–322.

Tanaka K (2003) Columns for complex visual object features in the inferotemporal cortex: clustering of cells with similar but slightly different stimulus selectivities. Cereb Cortex 13:90–99.

Tsao DY, Livingstone MS (2008) Mechanisms of face perception. Annu Rev Neurosci 31:411–437.

Tsunoda K, Yamane Y, Nishizaki M, Tanifuji M (2001) Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. Nat Neurosci 4:832–838.

Wallace MT, Ramachandran R, Stein BE (2004) A revised view of sensory cortical parcellation. Proc Natl Acad Sci U S A 101:2167–2172.

Wang G, Tanaka K, Tanifuji M (1996) Optical imaging of functional organization in the monkey inferotemporal cortex. Science 272:1665–1668.

Zangenehpour S, Chaudhuri A (2005) Patchy organization and asymmetric distribution of the neural correlates of face processing in monkey inferotemporal cortex. Curr Biol 15:993–1005.