



Published in final edited form as:

*Exp Brain Res.* 2009 September ; 198(0): 127–136. doi:10.1007/s00221-009-1772-y.

## Spatiotemporal architecture of cortical receptive fields and its impact on multisensory interactions

**David W. Royal,**

Kennedy Center for Research on Human Development, Vanderbilt University, 7110 MRB III, Nashville, TN 37232, USA

**Brian N. Carriere,** and

Kennedy Center for Research on Human Development, Vanderbilt University, 7110 MRB III, Nashville, TN 37232, USA

**Mark T. Wallace**

Kennedy Center for Research on Human Development, Vanderbilt University, 7110 MRB III, Nashville, TN 37232, USA. Department of Hearing and Speech Sciences, Vanderbilt University, 7110 MRB III, Nashville, TN 37232, USA

David W. Royal: david.royal@vanderbilt.edu

### Abstract

Recent electrophysiology studies have suggested that neuronal responses to multisensory stimuli may possess a unique temporal signature. To evaluate this temporal dynamism, unisensory and multisensory spatiotemporal receptive fields (STRFs) of neurons in the cortex of the cat anterior ectosylvian sulcus were constructed. Analyses revealed that the multisensory STRFs of these neurons differed significantly from the component unisensory STRFs and their linear summation. Most notably, multisensory responses were found to have higher peak firing rates, shorter response latencies, and longer discharge durations. More importantly, multisensory STRFs were characterized by two distinct temporal phases of enhanced integration that reflected the shorter response latencies and longer discharge durations. These findings further our understanding of the temporal architecture of cortical multisensory processing, and thus provide important insights into the possible functional role(s) played by multisensory cortex in spatially directed perceptual processes.

### Keywords

Multisensory integration; Spatiotemporal receptive fields; Anterior ectosylvian sulcus; Multimodal; Polysensory

### Introduction

Natural stimulus environments are composed of an over-whelming array of differing sensory cues that are constantly changing in a dynamic fashion in both space and time. Consequently, real-world settings pose a significant challenge for the nervous system in constructing veridical percepts from this rich tapestry of multisensory information. It is not surprising, therefore, that a specialized neural architecture has evolved whereby multisensory information is processed in brain areas responsible for not only resolving the

ambiguity inherent to natural environments, but also for making use of complementary information provided by each of the sensory channels. Although the foundation of work describing the neural consequences of multisensory convergence have come from the brainstem (Stein and Meredith 1993; Stein and Stanford 2008), studies are increasingly focusing on multisensory domains within the cerebral cortex because of its seminal role in the creation of the perceptual gestalt (Calvert et al. 2004). One of the preeminent models for these studies has been the cortex surrounding the anterior ectosylvian sulcus (AES) of the cat. In addition to being divided into core visual, auditory and somatosensory domains (Olson and Graybiel 1987; Clemo and Stein 1982, 1983; Clarey and Irvine 1990a, b), this region also contains a considerable population of multisensory neurons (Wallace et al. 2006; Jiang et al. 1994a, b), suggesting that AES may play an important role in the perceptual synthesis of multisensory information.

Like multisensory neurons in other brain structures (and species), those in AES demonstrate a remarkable capacity of integrating their different sensory inputs in dynamic and nonlinear ways (Wallace et al. 1992; Stein and Wallace 1996; Carriere et al. 2008). Typically, these integrated multisensory responses differ significantly from both of the constituent unisensory responses, as well as from that predicted by their simple summation. Prior work has shown that these integrated responses are dependent upon the nature of the stimuli combined and on their relationships to one another, such that stimuli that are in close spatial and temporal proximity, and that are weakly effective when presented alone, typically result in large response enhancements when combined (Wallace et al. 1992; Meredith and Stein 1986). Although the spatial, temporal, and effectiveness principles of multisensory integration have been derived through extensive manipulation of these stimulus parameters in isolation, recent work has revealed a striking interdependency between them (Carriere et al. 2008). For example, it has been found that the large receptive fields (RFs) that characterize AES neurons exhibit a strikingly heterogeneous spatial organization, and that this RF architecture plays an important deterministic role in the multisensory interactions seen in these neurons.

During the course of this work it became apparent that there were also intriguing changes in the temporal profile of the multisensory responses that were being masked when neural activity was collapsed into a singular response measure (i.e., mean spikes/trial). Hints of the complexity of the temporal dynamics of the multisensory response have been noted in prior work (Kadunce et al. 1997, 2001; Eordegh et al. 2005; Peck 1996). Indeed, recent work in the superior colliculus (SC) has illustrated that single unit responses can be speeded under multisensory conditions (Rowland et al. 2007a). Such a shortening of response latency is likely to form the basis for the speeding of saccadic eye movements observed under multisensory conditions (Bell et al. 2005; Amlot et al. 2003; Frens and Van Opstal 1988; Hughes et al. 1998). In contrast, few studies have directly examined the temporal response dynamics characteristic of cortical multisensory neurons, and how the temporal signature evolves during the course of the multisensory response (see Kayser et al. 2008 and Lakatos et al. 2007). In the current study, we sought to address these issues by constructing a first-order representation of the spatiotemporal receptive fields (STRFs) of multisensory neurons in cat AES cortex.

## Methods

### Subjects

Four cats (*Felis catus*, ~3.5 kg, female) served as subjects. All procedures were carried out in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and the guidelines of the Vanderbilt University Animal Care and Use Committee under an approved protocol.

## AES localization and surgeries

For the implantation of the recording chamber over AES, anesthesia was induced with a bolus of ketamine [20 mg/kg, administered intramuscularly (i.m.)] and acepromazine maleate (0.04 mg/kg, i.m.), and animals were intubated and artificially respired. A stable plane of surgical anesthesia was achieved using inhalation isoflurane (1.0–3.0%). Body temperature, expiratory CO<sub>2</sub>, blood pressure, and heart rate were monitored continuously (VSM7, Vetspecs, GA, USA) and maintained within a range consistent with a deep and stable plane of anesthesia. A craniotomy was made over the AES, and a recording chamber was then secured (see following text) to the skull in order to allow direct access to AES. A head post was positioned stereotaxically over the cranial midline. Both the recording chamber and head post were attached to the skull using stainless steel screws and orthopedic cement in such a way that allowed the animal to be maintained in a recumbent position during recordings without obstructing the face and ears. Preoperative analgesics and postoperative care (i.e., analgesics and antibiotics) were administered in close consultation with veterinary staff. Animals were allowed to recover at least 1 week before the first experimental recording.

## Recording and stimulus presentation

On the day of the experimental session, animals were anesthetized with an induction dose of ketamine [20 mg/kg, administered i.m.] and acepromazine maleate (0.04 mg/kg, i.m.). The animal was supported comfortably in a recumbent position without any wounds or pressure points by the head post implanted during surgery. Once anesthetized, the animal's vital signs were monitored continuously (VSM7, Vetspecs, GA, USA) in order to ensure a stable plane of anesthesia. Anesthesia was maintained with a constant-rate infusion of ketamine [5 mg/kg h, administered intravenously (i.v.)] delivered through a cannula placed in either the saphenous or cephalic veins. Animals were artificially respired, paralyzed with pancuronium bromide (0.1 mg/kg h, i.v.) to prevent ocular drift, and administered fluids [lactated Ringer solution (LRS), 4 ml/h, i.v.] for the duration of the recording session. The animal's body temperature was maintained at 37°C throughout the experiment via a water-circulating heating pad. On completion of the experiment animals were given an additional 60–100 ml LR S subcutaneously to facilitate recovery. While an opaque contact lens was placed over the eye ipsilateral to the AES being recorded from, the eye contralateral to the targeted AES was refracted and an appropriate contact lens was placed over this eye in order to bring into focus the tangent screen positioned in front of the animal.

Extracellular single unit activity was recorded with Parylene-coated tungsten microelectrodes 1–3 M $\Omega$  (HJ Winston, Clemmons, NC, USA). Responses were amplified, bandpass filtered and routed to a window discriminator (BAK Instruments, Mount Airy, MD, USA), an audio monitor, and an oscilloscope for monitoring. A custom-built PC-based real-time data acquisition system controlled the trial structure, recorded spike times (sampling rate = 10 kHz), and controlled the presentation of stimuli. Analysis of the data was performed using customized scripts in the MATLAB (Math Works Inc., Natick, MA, USA) programming environment. Visual stimuli consisted of the illumination of stationary light emitting diodes (LEDs; 50–300 ms duration) or movement of a bar of light (0.11–13.0 cd/m<sup>2</sup> on a background luminance of 0.10 cd/m<sup>2</sup>) projected onto a translucent tangent screen (positioned 50 cm from the animal's eyes). Auditory stimuli were delivered through positionable speakers that were either clipped to the corresponding LED or pinned to the tangent screen adjacent to the corresponding visual stimulus location. Auditory stimuli consisted of 50–100 ms duration broadband (20 Hz–20 kHz) noise bursts ranging in intensity from 50.6 to 70.0 dB sound pressure level (SPL) on a background of 45.0 dB SPL (A-weighted).

The focus of the present study was on visual-auditory multisensory neurons found at the border between AEV and FAES. Recording penetrations targeted this zone, and multisensory single units were isolated and the borders of their unisensory RFs were mapped qualitatively using methods identical to those used in the past (Wallace et al. 1992; Meredith and Stein 1986). Multisensory interactions are most robust with combinations of weakly effective stimuli (Carriere et al. 2008; Peck 1996; Perrault et al. 2005). As such, weakly effective visual and auditory stimuli were presented in a randomized, interleaved manner at multiple azimuthal locations along a single line of elevation and distributed equivalently across the widest aspect of the mapped RFs. Typically, five discrete stimulus positions spanned the RF interior, and these were flanked by a single stimulus position on either side of the RF (Fig. 1). Multisensory combinations consisted of visual and auditory stimuli presented at the same spatial location (i.e., spatial coincidence). The order in which stimulus locations were tested was determined randomly. Unisensory and multisensory stimulus conditions were interleaved randomly until a minimum of 15 trials were collected for a given stimulus location. Consecutive stimulus presentations were separated by no less than 7-s in order to avoid response habituation.

Each animal was used in ~20 recording sessions that lasted on average 9 h. At the conclusion of each recording session, the animal was allowed to recover from the anesthetizing and neuromuscular blocking agents before being returned to its home cage.

### Data analysis

Peristimulus time histograms and collapsed spike density functions characterized the neuron's evoked response for each condition (i.e., visual alone, auditory alone, visual-auditory) and stimulus location. Spike density functions were created by convolving the spike train from each trial for a given condition and location with a function resembling a post-synaptic potential specified by  $\tau_g$ , the time constant for the growth phase, and  $\tau_d$ , the time constant for the decay, according to the following formula:

$$R(t) = (1 - \exp(-t/\tau_g)) * \exp(t/\tau_d)$$

Based on physiological data from excitatory synapses,  $\tau_g$  was set to 1 ms and  $\tau_d$  to 20 ms (Kim and Connors 1993; Mason et al. 1991; Sayer et al. 1990). Data recorded under a given stimulus condition were then arrayed in a matrix such that the single unit's temporal response dynamics were highlighted as a function of azimuthal location. Arranging the data in this manner effectively generated STRF for a single plane of azimuth (Fig. 1). The mean spontaneous firing rate recorded during the 500 ms immediately preceding stimulus onset was considered the cell's baseline activity. Collapsed spike density functions were thresholded at two standard deviations above their respective baselines in order to delimit the evoked responses (i.e., a suprathreshold response lasting at least 30 consecutive milliseconds with a latency shorter than 50 ms for the auditory condition and 250 ms for the visual condition). Single units that failed to demonstrate an evoked response for either unisensory condition were removed from further consideration. A one-way between subjects ANOVA and Tukey's HSD post-hoc test were used to evaluate significant differences in evoked response latencies, response durations, mean firing rates, and peak firing rates across unisensory and multisensory stimulus conditions.

**Spatiotemporal receptive field analyses**—For each neuron, the mean stimulus evoked firing rates were normalized with the highest stimulus evoked response recorded from all tested conditions and locations, producing a response continuum ranging from 0.0 to 1.0, with 0.0 being no measurable response and 1.0 being the maximum evoked sensory response. When arrayed in a matrix such that the relative positions of these discharge trains

within the matrix reflected stimulus location, these values revealed the STRF organization for the visual, auditory, predicted multisensory (linear addition of the visual and auditory STRFs), and actual multisensory conditions for the region of RF overlap (Figs. 2a, 3a). To aid visualization, STRFs were then interpolated in two dimensions using the method of cubic splines.

**Integrative capacity**—Single units were evaluated for integrative capacity along a sliding temporal scale (bin size = 10 ms) using the following formula:

$$I_t = (M_t - U_t) / U_t * 100$$

where  $I_t$  equals the multisensory interactive product during a given time bin ( $t$ ),  $M_t$  equals the mean multisensory stimulus evoked firing rate during time bin  $t$ , and  $U_t$  equals the mean firing rate of the predicted multisensory stimulus evoked response during time bin  $t$  (Figs. 2b, 3b). The multisensory and unisensory stimulus evoked response dynamics were characterized further by examining the response latencies and response durations calculated from the thresholded spike density functions (Figs. 2c, 3c). The temporal dynamics characterizing the multisensory stimulus evoked responses were compared with those of the linear model as a function of the change in firing rate (Fig. 4a). There was wide variability in the response latencies and response durations within and across stimulus conditions. To more accurately evaluate integrative capacity in a manner that was independent of this variability, multisensory response discharges and their respective linear model analogs were aligned to the onset of the linear model's response (Fig. 4b, bottom panel). The integrative capacity of the entire population of neurons was calculated using this new alignment scheme using the equation given above (Fig. 4b, top panel).

## Results

### Uniqueness of multisensory STRFs

A total of 111 sensory responsive neurons were isolated in electrode penetrations that targeted the multisensory domain between AEV and FAES of AES cortex in four adult cats. Of these, 38% ( $n = 42$ ) were categorized as visual-auditory multisensory neurons, the focus of the current study. The remaining neurons were categorized as unisensory visual ( $n = 39$ ) or unisensory auditory ( $n = 30$ ). Consistent with previous reports, the RFs of multisensory AES neurons typically exhibited a high degree of spatial overlap of their constituent unisensory RF (Wallace et al. 2006; Jiang et al. 1994a; Wallace et al. 1992), and showed marked response heterogeneities within these RFs as a function of stimulus location (Carriere et al. 2008).

In order to begin to gain a view into the temporal dynamics of the evoked responses of these multisensory cortical neurons, identical visual and auditory stimuli differing only in their spatial location along a single plane of azimuth were presented (Fig. 1a). Note that when visual-auditory stimulus combinations were presented, these stimuli were always presented at the same location (i.e., spatial coincidence). With this stimulus structure, a first-order approximation of the STRFs of these neurons in the visual, auditory and multisensory (i.e., visual-auditory) realms was constructed (Fig. 1b). In each of the 42 visual-auditory neurons, unisensory and multisensory responses exhibited striking differences in their temporal response profiles as a function of stimulus location. In every neuron in which a complete set of STRFs could be derived (i.e., a sampling of the full azimuthal extent of the RF at one elevation for the visual, auditory and multisensory conditions;  $n = 24$ ), the STRF architecture consisted of multiple regions of elevated response surrounded by regions of lesser response (Fig. 2a). When these STRFs were compared across conditions, there was

marked variance between the visual and auditory STRFs, and between these unisensory STRFs and the multisensory STRF (Figs. 2a, 3a). The most dramatic finding was that for every cell examined, the multisensory STRF exhibited significant response enhancements in multiple spatial locations and within multiple temporal epochs compared to either of the unisensory STRFs, as well as to the predicted summation of the unisensory STRFs (i.e., examine the  $[M - (V + A)]$  subtraction plots shown in Figs. 2b, 3b).

### Multisensory spatiotemporal response dynamics

In an effort to better characterize the spatiotemporal structure outlined above, the temporal dynamics of the unisensory and multisensory stimulus evoked responses were investigated by measuring response latencies, response durations, and peak firing rates. Although there was wide variation between individual neurons and across spatial locations for the same neuron (compare Fig. 2c with Fig. 3c), on average, multisensory stimulus evoked responses were characterized by the shortest latencies ( $M$ :  $40 \pm 9$  ms;  $A$ :  $42 \pm 14$  ms;  $V$ :  $128 \pm 52$  ms) and the longest discharge durations ( $M$ :  $244 \pm 100$  ms;  $A$ :  $128 \pm 111$  ms;  $V$ :  $216 \pm 114$  ms) when compared to the constituent unisensory stimulus evoked responses. When the response latencies were compared across conditions and stimulus locations for each individual neuron, only two neurons (8%) demonstrated significantly shorter multisensory response latencies compared to the respective auditory response latencies ( $P < 0.05$ ). While multisensory stimulus evoked response latencies were on average faster than the predicted multisensory stimulus evoked response latencies (see “Methods” for how these predicted values were generated), these differences failed to reach the level of significance (Fig. 4a, left panel). In contrast, multisensory stimulus evoked discharges were significantly longer than predicted based on the unisensory responses ( $F(2, 209) = 5.83$ ,  $P < 0.05$ ) (Fig. 4a, right panel). Interestingly, for 62% ( $n = 14$ ) of our multisensory neurons, shorter response latencies and longer response discharges co-occurred in at least one of the stimulus locations tested. Comparing the peak and mean stimulus evoked firing rates across conditions, multisensory peak and mean firing rates ( $67 \pm 47$  Hz and  $32 \pm 25$  Hz, respectively) were higher than those elicited under both the visual ( $49 \pm 46$  Hz and  $23 \pm 22$  Hz, respectively) and auditory ( $49 \pm 34$  Hz and  $24 \pm 20$  Hz, respectively) conditions (Fig. 4). Each of these activity measures was significantly different for the evoked multisensory response when compared to the predicted multisensory response (peak:  $F(2, 209) = 11.37$ ,  $P < 0.001$ ; mean:  $F(2, 209) = 11.59$ ,  $P < 0.001$ ).

### STRF heterogeneity and multisensory interactions

As mentioned earlier, unisensory and multisensory STRFs exhibit significant response heterogeneities that are characterized by unique temporal response dynamics (see panel c of Figs. 2, 3). To facilitate comparisons of integrative capacity across conditions for individual neurons, the corresponding actual and predicted multisensory evoked responses were aligned to a single time point—the response latency of the predictive multisensory evoked responses. When viewed in this way, two phases of multisensory integration were readily evident. The magnitude of the multisensory interaction (i.e., response enhancement) was very large within the first 10 ms of the stimulus evoked response and peaked by 30 ms (panel b of Figs. 2, 3), likely reflecting in part the shortened response latencies of the true multisensory stimulus evoked responses. Following this early peak, the integrative capacity of multisensory neurons decayed rapidly. The second phase of integration was seen late in the response profile, and in large measure appears to reflect the extended duration of the multisensory response. This pattern is most evident in the population analyses shown in Fig. 4, and which relates the multisensory interaction to both the predicted onset and offset of the multisensory response.



## Discussion

In the current study, we have provided the first description of the STRF structure of cortical multisensory neurons. Extending our prior work examining the spatial receptive field (SRF) organization of these neurons (Carriere et al. 2008), we have shown that the STRF architecture of AES neurons is quite complex, and reveals a dynamism to the multisensory interactions exhibited by these neurons on a temporal scale not previously appreciated. Most important in the current study was the observation that many of the temporal aspects of the evoked multisensory response (e.g., response latency, response duration, etc.) differed substantially from the component unisensory responses and from simple combinatorial predictions based on these unisensory responses.

Although multisensory responses tended to be speeded relative to the predicted latency based on the unisensory responses, the effect failed to reach significance. Given the short response latencies of the contributing auditory channel, this is likely explained by a floor effect, in which the minimum latency is determined by the fastest arriving input. In contrast, multisensory responses were significantly longer in duration than predicted, suggesting one tangible temporal benefit to multisensory stimulus combinations is in their ability to produce longer duration discharge trains. These temporal differences in the evoked multisensory responses translated into two critical intervals of multisensory integration; one near response onset and the other near response offset. Future work will further explore this temporal architecture by manipulating both the timing and the efficacy of the unisensory responses.

Although no specific functional roles have yet to be attributed to cortical area AES, the data provided herein support a number of intriguing possibilities. First, based upon what is known about unisensory and multisensory processing performed by its constituent multisensory neurons, it is possible that area AES plays a key role in coordinate transformations, as has been proposed for other multisensory cortical domains (Salinas and Abbott 2001; Pouget and Snyder 2000; Deneve et al. 2007; Porter and Groh 2006). Areas responsible for coordinate transformations are challenged with the role of synthesizing sensory information initially encoded in different peripherally constructed representations (i.e., retinotopic, craniotopic, somatotopic) into a common framework able to support directed action and singular veridical percepts. A multisensory area like AES seems uniquely suited for this task given the strong convergence of different sensory inputs onto single neurons, and the spatial aspects of the stimulus encoding exhibited by these neurons. Second, multisensory AES neurons may play a role in the coding and binding of multisensory motion. Evidence for this possibility includes the tight register between the different unisensory RF of AES multisensory neurons (Jiang et al. 1994a, b; Wallace et al. 1992; Carriere et al. 2008), and the high degree of motion selectivity of visual neurons in AEV (Mucke et al. 1982; Benedek et al. 1988; Olson and Graybiel 1987; Scannell et al. 1996). Indeed, the spatiotemporal architecture of AES neurons may represent a mechanism for the encoding of multisensory motion, as has been demonstrated for the STRFs of visual neurons in the lateral geniculate nucleus (Ghazanfar and Nicolelis 2001; DeAngelis et al. 1995). Third, these findings suggest that the temporal profiles of multisensory responses are unique from the composite unisensory responses and, as such, multisensory interactions should be evaluated at a level beyond a simple rate code. The temporal dynamics of these responses may provide important insights into the functions subserved by AES. Although it must be noted that the results obtained in the current study were gathered in the anesthetized preparation, and acknowledged that a complete detailing of STRF organization and its functional utility await investigations in awake and behaving animals, we believe that the results are of general significance that transcends the preparation. In support of this, prior work examining multisensory interactions in the SC (a subcortical multisensory structure) of the awake and behaving cat (Wallace et al. 1998) showed there to be little difference in the

general integrative features of multisensory neurons. In addition, these multisensory neurons were noted to have marked response differences within their SRF; differences that closely resemble those seen in anesthetized recordings in both subcortical and cortical structures.

As alluded to in the introduction, the nature of multisensory interactions at the level of the single neuron and beyond depend critically on the spatial, temporal and inverse effectiveness principles of multisensory integration (Stein and Meredith 1993; Calvert et al. 2004). Nonetheless, despite their utility, these principles represent but a first-order set of guidelines, and often fail to capture the complete integrative capacity of individual multisensory neurons. This is even more germane in the current context where we have expanded our view of multisensory responses into the temporal domain, and have extended our analyses beyond the singular measures (e.g., mean spikes per trial, mean firing rate, etc.) typically used to evaluate sensory responses in multisensory studies. Here we can see for the first time how the domains of space and time interact in the generation of a multisensory neuron's integrated responses. Perhaps most important is that despite exhibiting a striking and complex spatiotemporal heterogeneity, the multisensory interactions of AES neurons showed a universal temporal "signature," being largest (and often non-linear) within two distinct temporal epochs—one very early in the response and the other at a time when each of the unisensory responses had abated. These findings are provocative in that they may offer new mechanistic insights into multisensory processes; insights that will likely be of great utility for the recent surge of interest in modeling multisensory interactions (Colonius and Diederich 2004; Avillac et al. 2005; Rowland et al. 2007b; Anastasio and Patton 2003; Xing and Andersen 2000).

As previously alluded to, the current study represents but a first-order characterization of the spatiotemporal response dynamics of multisensory neurons in cortical area AES. In an effort to make these complex experiments tractable, they were reductionist in nature in that the visual and auditory stimuli were very simple and were placed under very tight and stereotyped spatiotemporal constraints. However, as we learn more about the functional character of unisensory and multisensory processing in AES, it will become increasingly important to transition to experiments that utilize more dynamic stimuli in order to reveal AES activity under more naturalistic conditions. Among other considerations, this endeavor should involve the use of a richer stimulus library (one that addresses all three stimulus modalities represented in AES), moving stimuli, spatially disparate stimulus combinations and multiple stimulus onset asynchronies. Additionally, the methods of analysis of these more complex datasets will need to be refined and reimagined. Instead of the relatively crude (and time consuming) technique used here to generate STRFs, the transition to naturalistic stimuli will necessitate the use of more powerful and sophisticated approaches for the generation and analysis of STRFs (e.g., reverse-correlations (de Boer and de Jongh 1978; DeAngelis et al. 1995; Jones and Palmer 1987; Ringach et al. 1997) and response-plane techniques (Gerstein et al. 1968; Felleman and Kaas 1984; Mullikin et al. 1984). It is our expectation that the complexity of the spatiotemporal response patterns of these neurons under these circumstances will increase dramatically to reflect the richer and more dynamic stimulus relations that characterize more biologically meaningful stimuli.

## Acknowledgments

This work was supported by the National Institute of Mental Health Grant MH-63861 and by the Vanderbilt Kennedy Center for Research on Human Development. We would like to acknowledge the expert technical assistance of Matthew Fister and Juliane Krueger.

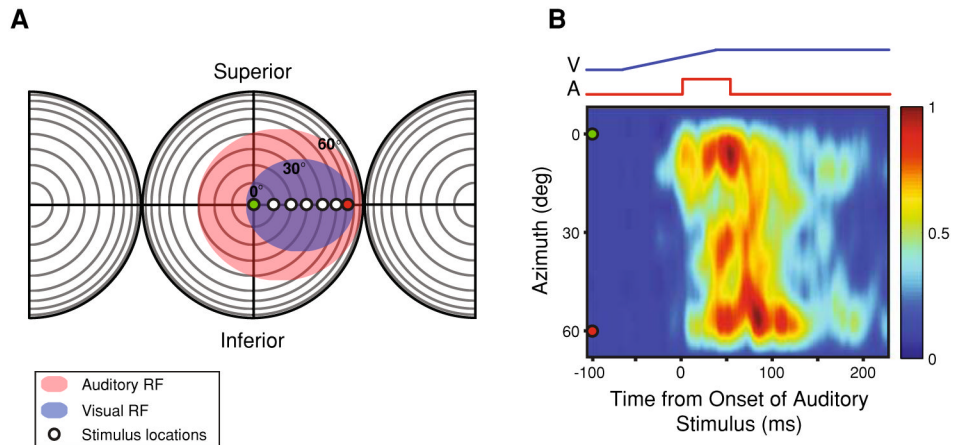


## References

- Amlot R, Walker R, Driver J, et al. Multimodal visual-somatosensory integration in saccade generation. *Neuropsychologia*. 2003; 41:1–15. [PubMed: 12427561]
- Anastasio TJ, Patton PE. A two-stage unsupervised learning algorithm reproduces multisensory enhancement in a neural network model of the corticotectal system. *J Neurosci*. 2003; 23:6713–6727. [PubMed: 12890764]
- Avillac M, Deneve S, Olivier E, et al. Reference frames for representing visual and tactile locations in parietal cortex. *Nat Neurosci*. 2005; 8:941–949. [PubMed: 15951810]
- Bell AH, Meredith MA, Van Opstal AJ, et al. Crossmodal integration in the primate superior colliculus underlying the preparation and initiation of saccadic eye movements. *J Neurophysiol*. 2005; 93:3659–3673. [PubMed: 15703222]
- Benedek G, Mucke L, Norita M, et al. Anterior ectosylvian visual area (AEV) of the cat: physiological properties. *Prog Brain Res*. 1988; 75:245–255. [PubMed: 3187054]
- Calvert, G.A.; Spence, C.; Stein, B.E. *The handbook of multisensory processes*. The MIT Press; USA: 2004.
- Carriere BN, Royal DW, Wallace MT. Spatial heterogeneity of cortical receptive fields and its impact on multisensory interactions. *J Neurophysiol*. 2008; 99(5):2357–2368. [PubMed: 18287544]
- Clarey JC, Irvine DR. The anterior ectosylvian sulcal auditory field in the cat: I. An electrophysiological study of its relationship to surrounding auditory cortical fields. *J Comp Neurol*. 1990a; 301:289–303. [PubMed: 2262593]
- Clarey JC, Irvine DR. The anterior ectosylvian sulcal auditory field in the cat: II. A horseradish peroxidase study of its thalamic and cortical connections. *J Comp Neurol*. 1990b; 301:304–324. [PubMed: 1702108]
- Clemons HR, Stein BE. Somatosensory cortex: a ‘new’ somatotopic representation. *Brain Res*. 1982; 235:162–168. [PubMed: 7188320]
- Clemons HR, Stein BE. Organization of a fourth somatosensory area of cortex in cat. *J Neurophysiol*. 1983; 50:910–925. [PubMed: 6631469]
- Colonius H, Diederich A. Multisensory interaction in saccadic reaction time: a time-window-of-integration model. *J Cogn Neurosci*. 2004; 16:1000–1009. [PubMed: 15298787]
- de Boer E, de Jongh HR. On cochlear encoding: potentialities and limitations of the reverse-correlation technique. *J Acoust Soc Am*. 1978; 63:115–135. [PubMed: 632404]
- DeAngelis GC, Ohzawa I, Freeman RD. Receptive-field dynamics in the central visual pathways. *Trends Neurosci*. 1995; 18:451–458. [PubMed: 8545912]
- Deneve S, Duhamel JR, Pouget A. Optimal sensorimotor integration in recurrent cortical networks: a neural implementation of Kalman filters. *J Neurosci*. 2007; 27:5744–5756. [PubMed: 17522318]
- Eordeghe G, Nagy A, Berenyi A, et al. Processing of spatial visual information along the pathway between the supragenulate nucleus and the anterior ectosylvian cortex. *Brain Res Bull*. 2005; 67:281–289. [PubMed: 16182935]
- Felleman DJ, Kaas JH. Receptive-field properties of neurons in middle temporal visual area (MT) of owl monkeys. *J Neurophysiol*. 1984; 52:488–513. [PubMed: 6481441]
- Frens MA, Van Opstal AJ. Visual-auditory interactions modulate saccade-related activity in monkey superior colliculus. *Brain Res Bull*. 1988; 46:211–224. [PubMed: 9667814]
- Gerstein GL, Butler RA, Erulkar SD. Excitation and inhibition in cochlear nucleus. I. Tone-burst stimulation. *J Neurophysiol*. 1968; 31:526–536. [PubMed: 5709869]
- Ghazanfar AA, Nicolelis MA. L. The structure and function of dynamic cortical and thalamic receptive fields. *Cereb Cortex*. 2001; 11:183–193. [PubMed: 11230091]
- Hughes HC, Nelson MD, Aronchick DM. Spatial characteristics of visual-auditory summation in human saccades. *Vision Res*. 1998; 38:3955–3963. [PubMed: 10211387]
- Jiang H, Lepore F, Ptito M, et al. Sensory modality distribution in the anterior ectosylvian cortex (AEC) of cats. *Exp Brain Res*. 1994a; 97:404–414. [PubMed: 8187853]
- Jiang H, Lepore F, Ptito M, et al. Sensory interactions in the anterior ectosylvian cortex of cats. *Exp Brain Res*. 1994b; 101:385–396. [PubMed: 7851509]

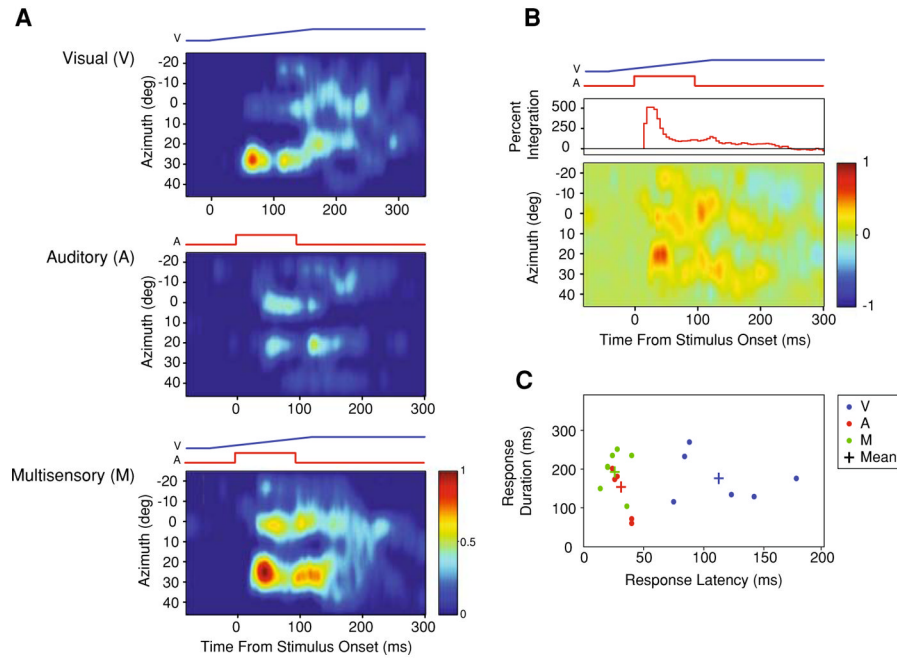
- Jones JP, Palmer LA. The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *J Neurophysiol.* 1987; 58:1187–1211. [PubMed: 3437330]
- Kadunce DC, Vaughan JW, Wallace MT, et al. Mechanisms of within- and cross-modality suppression in the superior colliculus. *J Neurophysiol.* 1997; 78:2834–2847. [PubMed: 9405504]
- Kadunce DC, Vaughan JW, Wallace MT, et al. The influence of visual and auditory receptive field organization on multisensory integration in the superior colliculus. *Exp Brain Res.* 2001; 139:303–310. [PubMed: 11545469]
- Kayser C, Petkov CI, Logothetis NK. Visual modulation of neurons in auditory cortex. *Cereb Cortex.* 2008; 18(7):1560–1574. [PubMed: 18180245]
- Kim HG, Connors BW. Apical dendrites of the neocortex: correlation between sodium- and calcium-dependent spiking and pyramidal cell morphology. *J Neurosci.* 1993; 13:5301–5311. [PubMed: 8254376]
- Lakatos P, Chen CM, O’Connell MN, et al. Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron.* 2007; 53(2):279–292. [PubMed: 17224408]
- Mason A, Nicoll A, Stratford K. Synaptic transmission between individual pyramidal neurons of the rat visual cortex in vitro. *J Neurosci.* 1991; 11:72–84. [PubMed: 1846012]
- Meredith MA, Stein BE. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J Neurophysiol.* 1986; 56:640–662. [PubMed: 3537225]
- Mucke L, Norita M, Benedek G, et al. Physiologic and anatomic investigation of a visual cortical area situated in the ventral bank of the anterior ectosylvian sulcus of the cat. *Exp Brain Res.* 1982; 46:1–11. [PubMed: 7067781]
- Mullikin WH, Jones JP, Palmer LA. Receptive-field properties and laminar distribution of X-like and Y-like simple cells in cat area 17. *J Neurophysiol.* 1984; 52:350–371. [PubMed: 6481437]
- Olson CR, Graybiel AM. Ectosylvian visual area of the cat: location, retinotopic organization, and connections. *J Comp Neurol.* 1987; 261:277–294. [PubMed: 3624544]
- Peck CK. Visual-auditory integration in cat superior colliculus: implications for neuronal control of the orienting response. *Prog Brain Res.* 1996; 112:167–177. [PubMed: 8979828]
- Perrault TJ Jr, Vaughan JW, Stein BE, et al. Superior colliculus neurons use distinct operational modes in the integration of multisensory stimuli. *J Neurophysiol.* 2005; 93:2575–2586. [PubMed: 15634709]
- Porter KK, Groh JM. Visual perception—fundamentals of awareness: multi-sensory integration and high-order perception. *Prog Brain Res.* 2006; 155(2):313–323. [PubMed: 17027396]
- Pouget A, Snyder LH. Computational approaches to sensorimotor transformations. *Nature Neurosci.* 2000; 3:1192–1198. [PubMed: 11127837]
- Ringach DL, Hawken MJ, Shapley R. Dynamics of orientation tuning in macaque primary visual cortex. *Nature.* 1997; 387:281–284. [PubMed: 9153392]
- Rowland B, Quessy S, Stanford TR, et al. Multisensory integration shortens physiological response latencies. *J Neurosci.* 2007a; 27:5879–5884. [PubMed: 17537958]
- Rowland B, Stanford T, Stein BA. Bayesian model unifies multisensory spatial localization with the physiological properties of the superior colliculus. *Exp Brain Res.* 2007b; 180:153–161. [PubMed: 17546470]
- Salinas E, Abbott LF. Advances in neural population coding. *Prog Brain Res.* 2001; 130:175–190. [PubMed: 11480274]
- Sayer RJ, Friedlander MJ, Redman SJ. The time course and amplitude of EPSPs evoked at synapses between pairs of CA3/CA1 neurons in the hippocampal slice. *J Neurosci.* 1990; 10:826–836. [PubMed: 2319304]
- Scannell JW, Sengpiel F, Tovee MJ, et al. Visual motion processing in the anterior ectosylvian sulcus of the cat. *J Neurophysiol.* 1996; 76:895–907. [PubMed: 8871207]
- Stein, BE.; Meredith, MA. *The merging of the senses.* MIT Press; Cambridge: 1993.
- Stein BE, Stanford TR. Multisensory integration: current issues from the perspective of the single neuron. *Nat Rev Neurosci.* 2008; 9:255–266. [PubMed: 18354398]

- Stein BE, Wallace MT. Comparisons of cross-modality integration in midbrain and cortex. *Prog Brain Res.* 1996; 112:289–299. [PubMed: 8979836]
- Wallace MT, Meredith MA, Stein BE. Integration of multiple sensory modalities in cat cortex. *Exp Brain Res.* 1992; 91:484–488. [PubMed: 1483520]
- Wallace MT, Meredith MA, Stein BE. Multisensory integration in the superior colliculus of the alert cat. *J Neurophysiol.* 1998; 80:1006–1010. [PubMed: 9705489]
- Wallace MT, Carriere BN, Perrault TJ Jr, et al. The development of cortical multisensory integration. *J Neurosci.* 2006; 26:11844–11849. [PubMed: 17108157]
- Xing J, Andersen RA. Models of the posterior parietal cortex which perform multimodal integration and represent space in several coordinate frames. *J Cogn Neurosci.* 2000; 12:601–614. [PubMed: 10936913]



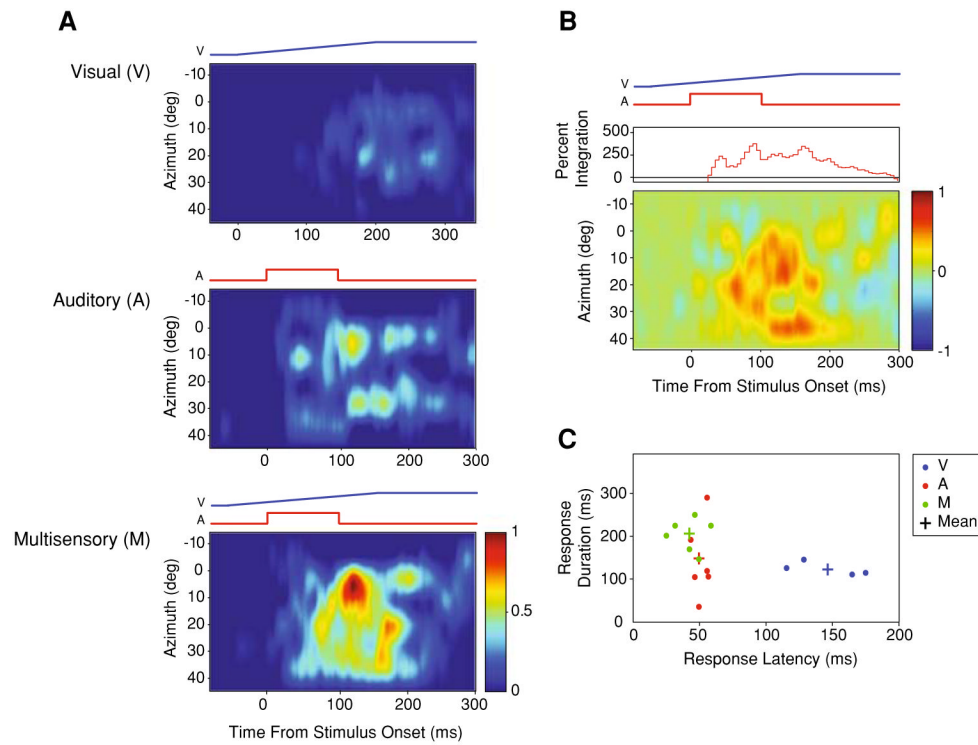
**Fig. 1.**

Construction of spatiotemporal receptive fields (STRFs). **a** The borders of the RFs of multisensory AES neurons were mapped using established criteria (visual RF: *blue* oval; auditory RF: *red* oval) (see “Methods”). *Circles* positioned at different azimuthal locations at 0° elevation represent positions at which unisensory and multisensory stimuli were presented in isolation and combination. In this representation of visual and auditory space, each concentric *circle* represents 10°, and the intersection of the *horizontal* and *vertical* lines represents the point directly in front of the animal’s eyes. **b** Example of a multisensory STRF constructed from the evoked responses generated from the experimental design diagrammed in **a**. The *traces* at the top represent the stimulus ramps for the visual (*blue*) and auditory (*red*) stimulus conditions. The pseudocolor plot represents the mean multisensory stimulus evoked responses that were recorded at each of the seven spatial locations. Note that the responses were normalized using the highest stimulus evoked response recorded from all tested conditions and locations, producing a response continuum ranging from 0.0 to 1.0, with 0.0 being no measurable response and 1.0 being the maximum evoked sensory response. The *green* (0°) and *red* (60°) *circles* are markers to represent the spatial transformation connecting panels **a** and **b**

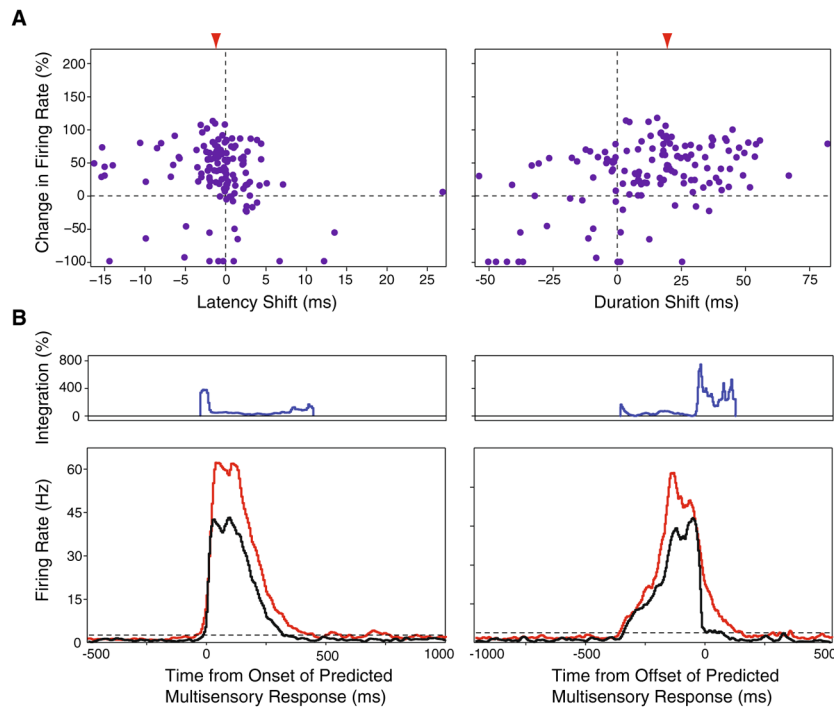


**Fig. 2.** STRFs produced from a visual-auditory multisensory neuron recorded from cortical area AES. **a** Visual (*top*), auditory (*middle*), and multisensory (*bottom*) STRFs aligned such that the relative timing of the stimuli depicted in the multisensory condition is preserved across panels. **b** The difference STRF generated by subtracting the predictive multisensory STRF (linear sum of the visual and auditory STRFs) from the true multisensory STRF. Warmer colors reflect areas where the actual multisensory response exceeds the predicted multisensory response. The *curve* shown in the *top panel* represents the magnitude of multisensory integration (%) as the response evolves over time. **c** Scatterplot highlights the relationship between response latency and response discharge duration plotted as a function of the stimulus condition. *Plus signs* represent the mean values for each stimulus condition. Note the *leftward* and *upward* shift in the multisensory response relative to the auditory and visual responses, reflecting the speeded and longer lasting responses, respectively





**Fig. 3.** STRFs produced from a visual-auditory multisensory neuron recorded from cortical area AES. Whereas, the neuron shown in the previous example (Fig. 2) was driven more strongly by the auditory condition, this neuron is driven more strongly by the visual condition (**a**). In panel **b**, note the temporal profile of the integrated multisensory response differs substantially from the previous example. Nonetheless, once again the multisensory response is faster and of longer duration than the unisensory responses (**c**)



**Fig. 4.** Spatiotemporal response dynamics of multisensory AES neurons. **a** Relationship between the relative change in firing rate between the multisensory and predicted multisensory conditions as a function of the concomitant shifts in evoked response latency (*left panel*) and duration (*right panel*). *Left panel* each point represents a response from one neuron at one tested stimulus location. Points to the left of the origin represent conditions where the multisensory response was faster than the predicted multisensory response. *Right panel* points to the right of the origin represent conditions where the multisensory response discharge duration was longer than the predicted multisensory response. *Red arrowheads* indicate the mean shift in latency (*left panel*) and duration (*right panel*). **b** Temporal patterns of multisensory response and multisensory integration in relation to the onset and offset of the predicted multisensory response. *Top* temporal pattern of multisensory integration. *Bottom* collapsed evoked response profiles for the predicted multisensory (*black*), and the multisensory (*red*) conditions aligned to the predicted multisensory response's onset (*left panel*) and offset (*right panel*). As shown by the individual examples in Figs. 2 and 3, the greatest amount of integration tends to occur towards to onset and offset of the stimulus evoked response, corresponding to the shorter response latencies and longer discharge durations recorded under the multisensory conditions