Properties of Correlated Neural Activity Clusters in Cat Auditory Cortex Resemble Those of Neural Assemblies
Jos J. Eggermont

You might find this additional information useful...

This article cites 62 articles, 29 of which you can access free at:
http://jn.physiology.org/cgi/content/full/96/2/746#BIBL

Medline items on this article's topics can be found at http://highwire.stanford.edu/lists/artbytopic.dtl on the following topics:
Veterinary Science .. Auditory Cortex
Veterinary Science .. Cats

Updated information and services including high-resolution figures, can be found at:
http://jn.physiology.org/cgi/content/full/96/2/746

Additional material and information about Journal of Neurophysiology can be found at:
http://www.the-aps.org/publications/jn

This information is current as of July 11, 2006.
Properties of Correlated Neural Activity Clusters in Cat Auditory Cortex Resemble Those of Neural Assemblies

Jos J. Eggermont

Departments of Physiology and Biophysics and Psychology, University of Calgary, Calgary, Alberta, Canada

Submitted 18 January 2006; accepted in final form 3 May 2006

INTRODUCTION

The frequency of sound is represented in the form of frequency-place maps in many auditory cortical areas. However, other representations in the auditory cortex may take the form of clusters rather than maps. A potential example is the spatial representation of sound (Cohen and Knudsen 1999). More complex perceptual properties have been linked to distributed neural representations called neural assemblies (Harris 2005). Do the properties of neural clusters have similarities to those attributed to neural assemblies? Comparison of the properties of clustered neural activity with those attributed to neural assemblies will be facilitated if the same criteria are used to define neural clusters and neural assemblies.

Neural assemblies have been defined as “a group of neurons [that are] at least transiently working together as indicated by correlation of unit activity” (Gerstein and Kirkland 2001). In visual cortex, cells with ~0.5-mm separation showed the highest correlation among cells with similar receptive fields and similar connectivity from the lateral geniculate nucleus (Lampl et al. 1999). This suggests that connectivity is a dominant factor in neural assembly formation. Correlation and connectivity may ensure efficient propagation of neural activity through downstream stages in the neural pathway (Kimpo et al. 2003; Kistler and Gerstner 2002; Reyes 2003).

It is common to think about a neural assembly as widely distributed in cortical space, potentially extending over various subdivisions of cortex (Singer and Gray 1995). Connections over large spatial divisions of auditory cortex are provided by the thalamic cell axonal divergence and convergence, often estimated to be between 2 and 5 mm at the cortical level (Lee et al. 2004a) and intracortically through horizontal fibers (Clarke et al. 1993; Wallace et al. 1991). In visual cortex, the spatially periodic effects of the patchy connections of the horizontal fibers have been shown by cross-correlation (Ts’o et al. 1986). These cortico-cortical connections are for a sizable part heterotopic, i.e., do connect cell groups with characteristic frequencies (CFs) differing by more than one octave (Lee et al. 2004b).

Neural assembly membership is expected to be stimulus dependent and context specific and may reflect the number and functional strength of its common inputs under different conditions (Usrey and Reid 1999). It is, however, likely that at any point in time several, spatially overlapping, neural assemblies are active. In response to external events, a group of neurons forming a dynamical cell assembly may spontaneously organize itself temporarily by correlated firing of their spiking activity.

Neural assemblies, thus defined, may potentially be probed using micro-electrode arrays that allow recording from a set of relatively widely spaced neurons. These neurons could participate in one or more neural assemblies. The quantification of the correlation in spiking activity occurring between pairs of such widely spaced neurons thus becomes crucial in defining membership of neural assemblies. The stimulus will be one of the dominant sources of neural correlation because of its common input character. Although it is common to correct for stimulus-induced correlations by using shift predictors or joint PST histogram techniques (Eggermont 1994), the brain does not have that luxury but may exploit this stimulus-dependent correlation to change the extent and structure of the neural assemblies.
To be confident that the observed changes in the cross-correlation matrix are functional, one has to eliminate other ways in which neural correlations can change. It is known that significant correlations can arise from covariation in neuronal excitability, i.e., co-modulated firing rates, and from covariation in response latency (Brody 1999a,b) as well as from trial-to-trial variability in the stimulus or the effect thereof (Ben-Shaul et al. 2001). The latter may also be a reflection of changes in the underlying assembly. It is therefore important to use experimental conditions that will not allow stimulus variability, e.g., resulting from movement of the animal, which can be accomplished by using anesthesia. It is, however, also important to minimize state changes in the brain that can be caused by fluctuating or slowly changing anesthesia levels (Erchova et al. 2002) and their effect on cross-correlation strength. This may be harder to accomplish because a change from a state showing periodic cross-correlograms in the auditory thalamus, likely caused by spindle activity due to the use of ketamine anesthesia, to one without these periodicities could be induced by stimulation with continuous dynamic ripple noise (Miller and Schreiner 2000). So, even in stable anesthetized animals, brain states are not invariant, and transitions may be induced by appropriate auditory stimulation. One can also posit that appropriate auditory stimulation breaks down those assemblies that during silence synchronize to produce spindle activity. The result could be changing pair correlations between the neurons recorded from. These changing correlation strengths may in turn cause neurons to associate in different assemblies.

To make putative neural assemblies last longer, and easier to detect, compared with their fleeting nature during wakefulness and behavior, we used long-lasting steady-state stimuli. We analyzed neural activity recorded simultaneously by two eight-electrode arrays, each having a size of 1.5 × 0.5 mm, from middle layers (III/IV) of three auditory cortical areas under several 15-min duration conditions. These included silence (spontaneous activity) and stimulation with Poisson-distributed click trains (Eggermont and Smith 1995), low-pass noise amplitude-modulated wide-band noise, and multi-frequency tone-pip ensembles with random frequency and presentation time and presentation rates of 20, 120, and 240 pips/s (Tomita and Eggermont 2005; Valentine and Eggermont 2004). The electrode arrays were either both in primary auditory cortex (AI) or one was in AI and the other in the anterior auditory field (AAF) or posterior auditory field (PAF). The matrix of pairwise peak correlation values was subjected to a hierarchical clustering procedure. Coincident-firing spectrotemporal receptive fields (STRF) were subsequently constructed for the clusters obtained. This was an extension of the coincident-firing pair STRFs reported previously (Tomita and Eggermont 2005).

Our findings suggest that cortical area boundaries nearly always act as cluster boundaries. Clusters were in large majority contiguous in space. Different clusters were formed under different stimulus conditions. Several operations on the coincident output of the clusters, as reflected in the common-spike and all-spike STRFs, may serve different modes of stimulus coding.

**METHODS**

**Animal preparation**

All animals were deeply anesthetized with the administration of 25 mg/kg of ketamine hydrochloride and 20 mg/kg of sodium pentobarbital, injected intramuscularly. A mixture of 0.2 ml of acepromazine (0.25 mg/ml) and 0.8 ml of atropine methyl nitrate (25 mg/ml) was administered subcutaneously at ~0.25 ml/kg body wt. Lidocaine (20 mg/ml) was injected subcutaneously prior to incision. The tissue overlaying the right temporal lobe was removed, and the dura was resected to expose the area bounded by anterior and posterior ectosylvian sulci. The cat was then secured with one screw cemented on the head without any other restraint. The wound margins were infused every 2 h with lidocain, and additional acepromazine/atropine mixture was administered every 2 h. The ketamine dose to maintain a state of areflexive anesthesia was on average 7.4 mg/kg·h (range: 3–12 mg/kg·h).

The care and the use of animals reported in this study was approved (BI 2001–021) and reviewed on a yearly basis by the Life and Environmental Sciences Animal Care Committee of the University of Calgary. All animals were maintained and handled according to the guidelines set by the Canadian Council of Animal Care.

**Acoustic stimulus presentation**

Stimuli were generated in MATLAB and transferred to the DSP boards of a TDT-2 (Tucker Davis Technologies) sound-delivery system. Acoustic stimuli were presented in an anechoic room from a speaker system [Fostex RM765 in combination with a Realistic super-tweeter that produced a flat spectrum (±5 dB) ≤40 kHz measured at the cat’s head] placed ~30° from the midline into the contralateral field, ~50 cm from the cat’s left ear. Calibration and monitoring of the sound field was accomplished with a condenser microphone (Bruel and Kjaer 4134) placed above the animal’s head, facing the speaker and a measuring amplifier (Bruel and Kjaer 2636). Prior to acute recordings, peripheral hearing sensitivity was determined using auditory brain stem response (ABR) thresholds (details in Noreña et al. 2003).

**Cortical frequency-tuning properties**

Frequency-tuning curves were measured by randomly presenting 27 gamma-tone pips with frequencies covering five octaves (e.g., 1.25–40 kHz) in equal logarithmic steps and presented at eight different stimulus levels in 10-dB steps (e.g., 5–75 dB SPL) at a rate of 4/s such that each intensity-frequency combination was repeated five times. The envelope of the gamma tones is given by

\[
g(\tau) = (\tau/4)^2 \exp(-\tau/4) \quad (1)
\]

with \(\tau\) in ms. The duration of the gamma tones at half-peak amplitude was 15 ms, and the envelope was truncated at 50 ms, where the amplitude is down by 64 dB compared with its peak value.

STRFs were obtained by presenting multi-frequency stimuli consisting of randomly presented gamma-tone pips. Here, tone pips for each of 81 frequencies in five octaves were randomly presented according to a Poisson process (Blake and Merzenich 2002), with similar average rate but different realization for each frequency. Three stimulus ensembles were used for which each tone pip frequency was presented at a rate of 0.25, 1.5, or 3 Hz so that the aggregate tone-pip rates were 20/s, 120/s, or 240/s.

Poisson-distributed click trains, with mean click rate of 8/s and dead time of 20 ms, and lasting 15 min were also presented with peak click levels equal to those for the gamma-tone pips. Wide-band noise (bandwidth: 40 kHz) modulated with a 30-Hz low-pass filtered noise and lasting 15 min was also presented.

**Recording and spike separation procedure**

Two arrays of eight electrodes (Frederic Haer) each with impedances between 1 and 2 MΩ were used. The electrodes were arranged in a 4 × 2 configuration with inter-electrode distance within rows and columns equal to 0.5 mm. Each electrode array was oriented such that
all electrodes were touching the cortical surface and then were manually and independently advanced using a Narishige M101 hydraulic microdrive (1 drive for each array). The depth of recording was between 700 and 1,200 μm and thus the electrodes were likely in deep layer III or layer IV. The signals were amplified 10,000 times using a Frederic Haer HiZx8 set of amplifiers with filter cut-off frequencies set at 300 Hz and 5 kHz. The signals were processed by a TDT-Pentusa multi-channel data-acquisition system (filter bandwidth: 300 Hz to 10 kHz). Spike sorting was done off-line using a semi-automated procedure based on principal component analysis and K-means clustering implemented in MATLAB. The spike times and waveforms were stored. The multiple single-unit (MSU) data presented in this paper represent only well-separated single units that, because of their regular spike wave form, likely are dominantly from pyramidal cells. For statistical purposes, the separated single-unit spike trains were added again to form a multi-unit spike train, thereby eliminating potential contributions from thalamocortical afferents or fast spikes from interneurons.

**Data analysis**

**FREQUENCY-TUNING.** To assess frequency-tuning properties, the peak number of action potentials in the post stimulus time histogram (PSTH, 5-ms bins) calculated over the first 100 ms after gamma-tone presentation was estimated. The results were calculated per stimulus intensity and were combined into an intensity-frequency-rate profile which from tuning curves, rate-intensity functions, and iso-intensity rate-frequency contours could be derived (Eggermont 1996) using routines implemented in MATLAB. The frequency-tuning curve was defined for a firing rate at 25% of the maximum peak-firing rate (FRmax). This was ~10–20% above the background firing-rate, but as the latter was dependent on the level of stimulus-induced suppression, the tuning curve criterion based on a percentage of peak firing rate was preferred over that based on increase over background activity. The values of firing rate given in this paper are in spikes/s.

For that purpose, the number of spikes in a 5-ms bin is divided by the number of repetitions (5) times the bin size in seconds (0.005 s), i.e., by 0.025. This gives very high values of FRmax when the MSU spikes are well synchronized to the onset of stimulus. The threshold was determined from the tuning curve.

The STRF was determined by constructing PSTHs for each of the preceding gamma tones in a 100-ms window. For that purpose, each spike elicited was plotted several times in the appropriate frequency bins and in the 100-ms time window after the onset of each of the preceding gamma tones (because spike latency is a priori unknown). If the gamma tones had no effect on a spontaneously firing neuron, the entire matrix of 81 frequency bins by 50 (2-ms duration) time bins would be filled uniformly. If certain frequencies consistently produce excitation in a certain latency window, then this part of the frequency-time plane would receive more hits. In case certain frequencies produce consistently lower activity than average, this would be interpreted as inhibition (Tomita and Eggermont 2005; Valentine and Eggermont 2004).

For STRFs, typically obtained at 65 dB SPL, contour lines or color-coded images are plotted. STRF overlap was measured by taking the contour line representing 30% of the difference between the maximum response and the mean value of the STRF for each of the neurons, calculating the number of pixels for each neuron (each pixel defined as: 1 frequency bin by 1 2-ms time bin) as well as the number of overlapping pixels. The weighted STRF overlap was defined as overlap size/sqrt(STRF1 size * STRF2 size), which scaled between 0 (no overlap) and 1 (complete overlap). The minimum latency of the 30% excitation contour line, the temporal duration of the 30% excitation contour line, the frequency of the center of gravity, and the spectral extent of the 30% contour lines were determined for each stimulus.

Overlap of inhibitory areas was not calculated because the regions are typically shallow and their presence correlates with the peak strength of the excitatory area (Valentine and Eggermont 2004).

**CORTICAL AREA BOUNDARIES.** The following properties were used in the assessment of cortical area boundaries: reversal of the CF gradient in the tonotopic map and along the electrode array, minimum latency values, the shape of the STRF, and the peak value of the cross-correlation coefficient for recordings straddling boundaries. For delineating the border between AI and AAF, we first of all used the sign and/or reversal of the gradient of CF along the electrode array with distance in the anterior direction (Norena and Eggermont 2005). The general shorter minimum latency in AAF compared with AI, and particularly the much higher frequency-tuning curve bandwidth at 20 dB above threshold in AAF (Eggermont 1998) were important as well. For the distinction between AI and PAF or potentially EPI (intermediate part of the posterior ectosylvian gyrus), we used mainly latency, which was ≥20 ms larger in non-AI areas. In addition, the sudden drop in peak cross-correlation coefficient across area boundaries under spontaneous firing conditions (Eggermont 2000) was a highly consistent indicator. It was, however, not possible to differentiate between PAF and EPI on any of those parameters, but on anatomical grounds, it is likely that the recordings were likely all from PAF, excluding the banks of the PES.

**CROSS-CORRELATION.** Cross-correlograms were calculated using custommade programs in Matlab (Eggermont 1992; Tomita and Eggermont 2005). Quantification of neural correlation in those studies was done on the basis of the cross-correlation coefficient

\[
R(\tau) = (R_{AB}(\tau) - E_{AB})(N_{A} - N_{A}^2/N)(N_{B} - N_{B}^2/N)^{0.5} \tag{2A}
\]

which for relatively low firing rates reduces to

\[
R(\tau) = (R_{AB}(\tau) - E_{AB})(N_{A}N_{B})^{0.5} \tag{2B}
\]

where \(R_{AB}(\tau)\) is the number of coincidences in the bin corresponding to lag time \(\tau\), \(E\) is the expected value for coincidences under the assumption of independent spike trains, \(E = (N_{A}N_{B})/N\), with \(n = T\Delta\), where \(N_{A}\) and \(N_{B}\) are the number of spikes in the recording. \(\Delta\) is the bin size and \(T\) the duration of the recording. \(|R(\tau)| \leq 1\). Stationarity estimates of the recordings were based on firing rate (mean and variance) in 100-s-long segments of the 15-min recordings. Functional correlation strength is not independent of the firing properties of the neurons in the pair. Specifically, periodicities in the neuronal firing imposed by cortical oscillations, e.g., in the spindle frequency range, may affect the peak cross-correlation coefficient (Eggermont and Smith 1995). Thus a deconvolution of the cross-covariance by the geometric mean of the auto-covariance functions of the two spike trains was implemented here. To correct for the overall firing rate, burst firing and periodicities in the firing of the neurons, the cross covariance, \([R_{AB}(\tau) - E_{AB}]\), was deconvolved with the square root of the product of the autocovariance functions \([R_{AA}(\tau) - E_{A}]\) and \([R_{BB}(\tau) - E_{B}]\). Here \(E_{AB}\) and \(E_{A}\) and \(E_{B}\) are the expected values for the cross- respectively auto-covariance functions under the assumption of independence and Poisson processes. This deconvolution was done in the frequency domain, where it becomes a simple division, Fourier transformation back to the time domain resulted in the corrected cross-correlation coefficient function \(R_{AB}(\tau)\). This correction procedure is an extension of the method used in Eggermont and Smith (1996) where the deconvolution was done with the auto-covariance of the trigger signal only. However, because most correlations in cortex result from shared input, the more symmetric form indicated above was used. The cross-correlations were all significantly different from zero at a level of 3 SD \((P < 0.01)\).

Because none of the stimuli used in this study repeated in their 15-min duration, a standard shift predictor could not be used to estimate the effect of the stimulus locking on the value of \(R_c\). For the Poisson-distributed clicks, the PSTH predictor was used (Perkel et al.
HIERRARCHICAL CLUSTERING. The method of hierarchical clustering (Everitt 1978) applies when some clusters are nested within other clusters and the technique operates on a matrix of individual similarities or distances. In our case, we use pair-wise log(RC) values, RC being the peak value of the corrected cross-correlogram, or pair-wise STRF weighted overlap values, which serve as similarity measures. The log(RC) was used because the distribution of log(RC) values was approximately normal, and thus equal weight was given to high as well as low values. The implementation in MATLAB uses the nearest-neighbor single-linkage method to form clusters, which are visualized in a dendrogram. As inconsistency level a value of 1.0 was used. The goodness of fit of the dendrogram to the original similarities in the matrix is assessed by the cophenetic correlation coefficient, which is the Spearman rank correlation between the heights of the link in the dendrogram at which the two observations are first joined into a cluster and the original similarities of these two observations. The result from hierarchical clustering is generally unacceptable if the cophenetic correlation coefficient is <0.75.

For comparison, we also used hierarchical clustering based on the “ward” linkage method, which uses a minimum variance algorithm. Results were generally the same for multi-electrode arrays, but most of the nonclustered electrodes were clustered using this method. Inspection of those clusters showed that all poorly correlated electrodes were grouped using this method. As a consequence the nearest neighbor clustering was preferred.

STABILITY OF CLUSTERING UNDER DIFFERENT STIMULATION CONDITIONS. This was measured by calculating per electrode the stability of clustering and averaging this over the number of electrodes and the number of stimulus conditions. The cluster stability was calculated by assigning a value of one for each stimulus that resulted in the same largest cluster and then subtracting the number of conditions that resulted in different clusters. Thus if for a six-stimulus case, four stimuli result in the same cluster then the cluster value is 4 – 2 = 2. If the cluster size is equal to half the number of conditions, then the cluster value is 0. If the largest cluster was only obtained for two of six stimuli, then the cluster value is 2 – 4 = -2, etc. This is calculated for each electrode and averaged over the number of electrodes. The same analysis was done for six time segments (equal to 150 s) of the 900-s silence condition, and those results were used in a pair-wise comparison with the one for different stimulus conditions.

COMPARISON OF CLUSTERING BY STRF-OVERLAP AND PEAK R PER PP-STIMULUS. Here we calculated a simple correlation per stimulus condition, assigning a one to each electrode when the clustering was the same based on pair-wise peak correlation and on pair-wise STRF overlap. The value divided by the number of electrodes was tested against the expected value of getting that number of clusters under random assignments.

Locally weighted scatter plot smoothing (Lowess) curves were displayed in several figures. All statistical analyses were performed using Statview 5 (SAS Institute).

RESULTS

The results shown here were obtained from 30 sets of stationary recordings obtained from AI, PAF, and AAF in 16 adult cats. We selected recording sites with ≥12 functional electrodes and where at least four stimulus conditions each lasting 15 min gave stationary recordings. This comprised the calculation of 15,308 MSU pair correlations, and 5,712 pair-wise STRF overlaps.

Pair-wise cross-correlation based clustering

AN ILLUSTRATIVE EXAMPLE. We illustrate the basic method with data from a 15-min recording of spontaneous activity using two eight-electrode arrays in AI. The neighbor-electrode distances in an array are 0.5 mm and 15 of the 16 electrodes recorded spiking activity. Figure 1A shows a subset of electrode MSU-pair correlograms between electrode 1 and electrodes 3–5, electrodes 2 and 3–5, and electrode 3 with 4 and 5. A three-point smoothing is applied to eliminate spurious peaks here and in all correlograms to follow. This resulted in the autocorrelation (3–3) to be ~0.6 in value rather than 1.0 because it is essentially only one bin wide. Each panel shows the correlograms scaled on their own extremes. The time base is from ~0.3 to 0.3 s. The red traces show the standard correlograms calculated according to Eq. 2B; large oscillations with a frequency of ~7 Hz, reflecting anesthesia-induced spindling activity are evident. The blue traces represent the corrected correlograms (RC, see METHODS). Most of the oscillatory activity has disappeared and the peak values have been reduced. The peak RC values are indicated in the right-hand corner of each panel. Figure 1B shows the full matrix of the logarithm of RC, the subset shown in Fig. 1A corresponds to a region in the top-left corner (columns 1–3 with rows 3–5). Column numbers and row numbers correspond to electrode numbers as indicated. The dark row and column for electrode 9 indicate that this electrode was not recording spontaneous activity; the value of the peak correlation for this electrode was arbitrarily set at 0.01 (only for graphical purposes) to allow use of the full color scale for representing the matrix entries. In cluster calculations, we used a value of RC = 0.001 for pairs with nonresponding electrodes. The color bar to the right indicates the range of log(RC) values. Based on this pair-wise RC matrix, a hierarchical clustering procedure (Fig. 1C) resulted in four clusters, one comprising activity on five electrodes (1–5), one comprising three electrodes (6–8), and two comprising two electrodes (13, 15 and 14, 16), and there were three electrodes that did not cluster with other electrodes (10, 11, 12). Electrode 9, which did not show spontaneous activity, also formed a separate cluster. Clusters 13–15 and 14–16 are very close to being clustered in one group. The multi-electrode clusters are indicated on the dendrogram with colored dots at their final branching points. In Fig. 1D, the electrode positions are indicated on the surface of the auditory cortex by the centers of the colored circles, and the clusters are indicated by the same colors as used in the dendrogram. Black circles reflect nonclustered electrodes. Clustered electrodes are grouped in close proximity and do not cross electrode-array boundaries. The size of the individual clusters does not extend much beyond 1 mm (inter-electrode distance is 0.5 mm). Both electrode arrays were in AI, based on the criteria outlined in the Methods section.

GROUP DATA. For a total of 132 multi-electrode-array recordings subjected to the clustering procedure, we found that 6 had a cophenet correlation coefficient <0.75 (see METHODS), whereas the average value was 0.92 ± 0.08 (SD). Thus in the vast majority (95.5%) the hierarchical clustering procedure was justified. The six cases with low (<0.75) cophenet correlation were not considered in the group data.

The average pair-wise RC within a cluster is significantly larger than that between clusters, not surprising because this
difference is the basis of cluster formation. Figure 2 shows the distribution of the logarithm of the average $R_C$ values within clusters (Fig. 2A), the average $R_C$ value between clusters (Fig. 2B), and the geometric mean of the within-cluster $R_C$ values that participate in the between-cluster calculations (Fig. 2C).

Best-fitting normal distributions are shown; the distributions appear to be approximately normally distributed. Figure 2D shows the scatter plot of between-cluster $R_C$ values and the geometric mean of the corresponding within-cluster $R_C$ values. All between-cluster $R_C$ values are well below those for within cluster values (see Table 1).

**STABILITY OF CLUSTERING.** Here we show an example of the result of the stability calculation (see METHODS) for the cluster procedure applied to six consecutive 150-s sections of a spontaneous activity recording. We illustrate the time dependence in Fig. 3 with a “flower-plot” where “petal” colors indicate the assigned cluster and petal position indicates the time section. Electrode 14 did not record neural activity. Activity on electrode 10 did not form clusters and neither did the activity on electrodes 1 and 6 most of the time. One observes that occasionally there is a flip over to another cluster. The time segment in which this occurs is different for different electrodes and the observed changes also differ per electrode. So these changes either reflect small changes in $R_C$ values caused by noise or reflect the fleeting nature of neural connection strengths. This procedure was carried out for the silence condition on all recordings and quantified in the same way as for stimulus effects on clustering (see METHODS) and was compared therewith (next section).

**STIMULUS EFFECTS ON CLUSTERING.** For the recording condition shown in Fig. 1, six different stimulus conditions, i.e., silence, three sets of multi-tone stimuli with aggregate gamma-tone-pip rates of 20/s, 120/s, and 240/s, Poisson-distributed click trains and low-pass noise amplitude-modulated wide-band noise (lpamn), each lasting for 15 min were presented. We illustrate the stimulus dependence of clustering in Fig. 4 again with a flower-plot where petal colors again indicate the assigned cluster and petal position now indicates the presented stimulus (labeled for electrode 10 in the top right corner, electrode 9 was not functional in this recording and the flower petals were left white). The base color (same as the center of each flower) again indicates those electrodes that form a cluster of their own. Electrodes 1–8 and 9–16 belong to the more posterior and anterior array in Fig. 1, respectively. Electrodes 6–8 were assigned to the same cluster regardless stimulus type and so were, albeit to a different cluster, electrodes 13 and 15. Electrodes 1–5 showed the same clustering for all stimuli except for the 240/s multi-tone stimulus. This was the only stimulus for which clustering crossed electrode array boundaries and pro-
duced a 12-electrode cluster. Low-pass amplitude modulated noise created a unique cluster for electrodes 12 and 13. Potential effects of changes in the mean and variance of the pairwise correlation matrix with time after the first recording was evaluated for all stimulus conditions and silence. Neither the mean $R_C$ nor its variance changed as a function of time (regression analysis, slope not significantly different from 0, $P > 0.06$).

Cluster stability was compared between stimulus effects and those of sectioning the 900 s of recording during silence into six subsets. A paired comparison showed that clusters were significantly more stable for different 150-s sections within a 900-s period than under different stimulus conditions ($P < 0.01$).

Table 1 shows the group means for within- and between-cluster $R_C$ values per stimulus condition and also combined across all conditions (bottom row). Whereas the within-cluster values are all significantly larger ($P < 0.0001$) than the between-cluster values for any given stimulus, there were no significant differences in within- and between-cluster $R_C$ values for different stimuli.

The fraction of electrodes that participated in the same cluster for all pairs of different stimulus conditions is shown in a similarity matrix form (Table 2). The highest fraction of common-clustered electrodes (0.747) is between two multi-tone stimuli with aggregate rates of 120/s and 240/s, whereas the lowest fraction (0.433) is found between the 20/s multi-tone stimulus and Poisson-distributed clicks. Multi-dimensional scaling on the dissimilarity matrix (unity matrix minus

![Figure 2](image-url)

**FIG. 2.** Within- and between-cluster pair-wise correlation strengths. A: distribution of the logarithm of the average within-cluster $R_C$ values. B: for average between-cluster $R_C$ values. C: distribution of the logarithm of the geometric mean of the average within-cluster $R_C$ values that participate in the between-cluster pairs. D: scatterplot of between-cluster $R_C$ values against the geometric mean of the corresponding within-cluster $R_C$ values.

![Figure 3](image-url)

**FIG. 3.** Flower plot of the stability of clustering for consecutive 150-s segments. Petal color indicates clusters. The numbers in the center of each flower indicate electrode number. White petals indicate that the electrode was not recording spike activity. Green petal color indicates nonclustered activity. Color changes for individual electrodes indicate signs of clustering instability.

![Figure 4](image-url)

**FIG. 4.** Flower plot of the stability of clustering under various stimulus conditions (indicated on the petals of the electrode 10 representation). Details same as in Fig. 3. The 240/s multi-tone stimulus induces large changes in the clustering.

<table>
<thead>
<tr>
<th>Stimulus Type</th>
<th>Within Clusters</th>
<th>Between Clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>$R_C$ values</td>
</tr>
<tr>
<td>Multi-tone 20/s</td>
<td>97</td>
<td>0.072 ± 0.035</td>
</tr>
<tr>
<td>Multi-tone 120/s</td>
<td>97</td>
<td>0.060 ± 0.029</td>
</tr>
<tr>
<td>Multi-tone 240/s</td>
<td>62</td>
<td>0.061 ± 0.025</td>
</tr>
<tr>
<td>Silence</td>
<td>96</td>
<td>0.069 ± 0.033</td>
</tr>
<tr>
<td>Poisson clicks</td>
<td>46</td>
<td>0.073 ± 0.037</td>
</tr>
<tr>
<td>Low-pass AM noise</td>
<td>34</td>
<td>0.070 ± 0.041</td>
</tr>
<tr>
<td>All stimuli</td>
<td>432</td>
<td>0.067 ± 0.033</td>
</tr>
</tbody>
</table>
the similarity matrix) shows (Fig. 5A) that the clusters for silence, and the multi-tone stimuli with rates of 120/s and 240/s are relatively close, whereas those for the broadband, i.e., Poisson and low-pass amplitude-modulated noise, stimuli are very different from the other stimuli but also differ considerably among themselves (fraction of electrodes in the same clusters: 0.491). This is also reflected in the changes in cluster size. From silence to multi-tone stimuli, 70% of the cluster sizes did not change significantly and only 13% resulted in larger clusters and 17% is smaller clusters. In contrast, for the comparison of the effect of silence and broadband stimuli, one finds that in 40%, the size did not change, in 40%, became larger, and in 20%, became smaller. Nine simultaneous recordings from PAF and AI showed no significant differences in the fraction of electrodes that participated in the same cluster in PAF or AI. Clusters on the same electrode array were generally contiguous, only in 18/432 clusters did we observe otherwise. Of these 10 were found for the 20/s multi-tone stimulus and 4 for the 120/s multi-tone stimulus. The others were for Poisson-distributed clicks (2), low-pass amplitude modulated noise (1), and silence (1).

This all suggests that stimulation changes cluster size. Clusters most often became larger under broadband stimulation compared with silence, whereas under multi-tone stimulation they were more likely to stay the same in size as under spontaneous conditions.

For the Poisson clicks, the stimulus-predictor based on the cross-correlation of the average PSTHs (akin to a time-inverted 1st-order Poisson kernel) was calculated. The mean (SD) peak value of the kernels was 0.039 ± 0.041 spikes/click, with a range of −0.25–0.15 spikes/click. With this predictor, 1,231 cross-correlograms were calculated. The peak correlation coefficient $R_C$ was strongly correlated with the one calculated with the PSTH predictor ($r^2 = 0.965$) and the regression line was: $\log_{10}(R_{PSTH}) = -0.062 + 0.97 \times \log_{10}(R_C)$. The relationship is shown Fig. 5B. In general, the PSTH-corrected $R$ was slightly smaller than the one corrected on basis of the average firing rates, i.e., $R_C$. In incidental cases, the PSTH-corrected $R$ was larger than $R_C$. This happened among others when one kernel was initially negative and the other initially positive. The cluster assignments based on $R_{PSTH}$ and $R_C$ were very similar: In 12 of the 13 recordings, the cluster assignment was exactly the same, in 1 recording the clustering based on $R_{PSTH}$ split one cluster into 2. As a result, there was only a minute effect in the multivariate scaling plot (Fig. 5A).

CORTICAL AREAS AND CLUSTERING. Table 3 shows the mean $R_C$ values for within- and between-cortical area clusters (across all stimulus conditions because according to ANOVA there was no interaction between area and stimulus type). Not all electrodes could be unambiguously assigned to an auditory cortical area, i.e., in case of nonresponsiveness to tones, so the total number is slightly less than the number of clusters in Table 1. One observes that the within-area, within-cluster $R_C$ values are much higher than the within-area, between-cluster values, and those in turn are higher than the between-area, between-cluster values. The within-cluster $R_C$ values are significantly larger ($P < 0.0001$) within AI than those within PAF and AAF, which are not significantly different from each other.

Clusters were mostly confined within an electrode array; only 11 clusters of 243 multi-electrode clusters were crossing electrode array boundaries. We also observed that 3 of the 243 $R_C$-based clusters crossed area boundaries and comprised activity from AI as well as PAF. Given that there were 64 simultaneous recordings of 27 clusters in PAF, 34 in AI, and 3 across PAF and AI, this suggests that only ∼5% (3/61) common inputs are found between these areas. Table 4 shows the

![FIG. 5. A: multi-dimensional scaling of the average difference in the fraction of electrodes that participate in the same cluster for 6 stimulus conditions. For silence and multi-tone stimuli, the differences are relatively small. Large changes occur for Poisson-distributed click stimuli and low-pass amplitude modulated noise (lpamn). B: comparison of the stimulus-corrected peak cross-correlation coefficient (Rpsth) and the peak $R_C$ value for Poisson-distributed clicks.](jneurophysiol.org)
cluster size both in number of participating electrodes and in largest distance (in mm) between participating electrodes split according to cortical area and stimulus type. There were no significant differences in the cluster size as expressed by number of electrodes between cortical areas or stimulus type. In contrast, the spatial extent of the clusters (across all stimulus conditions) was significantly larger in PAF compared with AI (P < 0.001). Because there was no difference in cluster size between the three tonal stimulus ensembles (comprising the 3 multi-tone ensembles with aggregate presentation rates of 20/s, 120/s, and 240/s) or between the two broadband stimuli (Poisson clicks and AM noise), we grouped the stimuli into tonal stimuli, broadband stimuli, and silence and found cluster size significantly larger (across all areas) for broadband stimuli compared with tonal stimuli (P < 0.05). There was no interaction between cortical area and stimulus type.

**TABLE 3. Within and between cluster R_c values with SD and number of clusters (n) involved for within and between cortical areas**

<table>
<thead>
<tr>
<th>Area</th>
<th>Within Cluster</th>
<th>Between Cluster</th>
<th>Between Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>R_c value</td>
<td>n</td>
</tr>
<tr>
<td>AI</td>
<td>318</td>
<td>0.073 ± 0.031</td>
<td>362</td>
</tr>
<tr>
<td>PAF</td>
<td>74</td>
<td>0.047 ± 0.023</td>
<td>32</td>
</tr>
<tr>
<td>AAF</td>
<td>16</td>
<td>0.037 ± 0.028</td>
<td>9</td>
</tr>
<tr>
<td>PAFxAI</td>
<td></td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>AAFxAI</td>
<td></td>
<td></td>
<td>9</td>
</tr>
</tbody>
</table>

AI, primary auditory cortex; PAF, posterior auditory field; AAF, anterior auditory field.

**Dependence of R_c on inter-electrode distance.** The different values within and between clusters as well as within and between cortical areas are potentially related to the decrease in peak R_c value with increasing inter-electrode distance. Figure 6 shows this distance function for randomly selected 50% (to avoid cluttering the graph) of the pair-wise recordings during silence. Different symbols indicate different cortical areas. Whereas there can be almost a factor 100 (2 log units) difference in R_c values at any particular distance, there is a significant decrease with distance d (P < 0.0001). For unit pairs in AI only, the nonlinear regression line (through all the points) is R_c = 0.05*exp(-d/4.2), r^2 = 0.24, (P < 0.0001) and for unit pairs restricted to PAF, the decrease with distance is given by R_c = 0.05*exp(-d/1.1), r^2 = 0.4, (P < 0.0001). This suggests that the pair-wise R_c decreases in AI by a factor 2.72 every 4.2 mm but every 1.1 mm in PAF. This could be related to a different CF gradient in the two fields. Pair correlations between units in AI and PAF can be as high as for those within AI for the same distance range, but on average the mean value is significantly lower between AI and PAF (P < 0.0001). Similarly in the distance range >3 mm, the within AI correlation was significantly (P < 0.0001) larger than between AI and AAF. The regression between R_c and d for pairs between AI and PAF reads: R = 0.014*exp(-d/8.7), r^2 = 0.02 (P < 0.01) and thus decreases even slower (a factor 2.72 every 8.7 mm) than within AI. Thus the sparseness of clusters crossing area borders is likely not related to the decrease of R_c with distance.

**Pair-wise STRF-overlap based clustering**

STRFs. An example set of STRFs obtained with the 120/s multi-tone stimulus is shown in Fig. 7. The electrode numbers read from left to right and top to bottom as indicated. The positions of the electrodes on the cortical surface are shown in Fig. 8D. Frequency is shown in kilohertz along the vertical axes using a logarithmic scale, and time since tone-pip onset is shown on the horizontal axes from 0 to 100 ms. The frequency-time plots basically show the frequency-dependent PSTH with 81 frequency bins and 50 time bins of 2-ms duration. Dark red colors indicate maximum response; dark blue colors indicate minimal response. The most general pattern consists of a
strong excitation followed after 10–20 ms by post activation suppression, sometimes extending for a wider frequency range than the excitation. Best response frequencies range from just below 5 kHz (electrode 1) to just below 20 kHz (electrode 11). Some electrodes show broad sensitivity to frequency (electrodes 3 and 13). Electrodes 8 and 9 have weak responses and their STRFs are ill defined.

STRF-BASED CLUSTERING. The weighted overlap of two STRFs (see methods) was used as an alternative measure for hierarchical clustering and compared with that based on $R_C$. The weighted overlap is based on a percentage from peak value and therefore is normalized for firing rate. Figure 8A shows pairwise overlap of STRFs based on the 30% contour lines (between peak and mean value) of STRFs for electrodes 10, 12, 14, and 16 from Fig. 7. Neighboring electrode STRFs (e.g., 10 and 12) tend to have larger overlap than more remote ones (e.g., 10 and 16). The full matrix of weighted pair-wise overlap (see methods) is shown in Fig. 8B and values range from 1 (on the diagonal, complete overlap) to close to 0 (hardly any overlap). The pair-wise weighted STRF-overlap values correlate with the pair-wise correlation strength as shown in Fig. 8C with a Pearson correlation coefficient of 0.66. Following a hierarchical clustering procedure based on STRF overlap, three clusters were assigned (Fig. 8D, left half of the circles). Electrodes 1–8, comprising the more caudal array, feature as one cluster, whereas electrodes 10–12, 14, and 16 form one cluster and electrodes 9 and 13 form another cluster on the more anterior array. Both arrays were in AI. Note that black again indicates single-electrode clusters. This cluster assignment can be compared with the clustering based on the $R_C$ values that resulted in only two multi-electrode clusters as shown in Fig. 8D (right half of the circles). Both clusters comprised six electrodes and were each confined to one array. One observes that there is a good correspondence in the assigned clusters for the majority of the electrodes, suggesting that at least in this case similarities in STRF are important in determining neural synchrony.

GROUP DATA. Figure 9A shows the scatter gram of the average within-cluster pair-wise $R_C$ and the average pair-wise weighted STRF overlap, and $B$ shows the same for the average between-cluster values. The average weighted STRF-overlap per cluster and the average pair-wise log($R_C$) are significantly correlated with $r^2 = 0.2$, those between clusters are correlated with $r^2 = 0.36$. Thus the pair-wise $R_C$ between clusters is more dependent on STRF overlap than that within clusters, regardless of the fact that the range of overlap values and log($R_C$) values is the same.

Table 5 shows the statistics for the Pearson correlation between the pair-wise STRF overlap and peak $R_C$ value per multi-tone stimulus file regardless of clustering. There was no significant difference between the three multi-tone stimulus ensembles, and the average Pearson correlation was 0.615,
suggesting that, on average, 38% of the variance in $R_C$ is explained by STRF overlap. A likely source for the unexplained variance is cortico-cortical connections such as those provided by the horizontal fibers, but it is also likely that less specific cortical network properties play a role.

We also compared the clustering based on pair-wise overlap with that based on pair-wise $R_C$ for the multi-tone stimulus conditions and found that the results of the two clustering methods were significantly more in agreement than what chance assignments predicted ($P < 0.0001$).

Whereas the average number of clusters based on $R_C$ values (3.4 ± 1.0) was the same as that based on weighted overlap (3.4 ± 1.2), the type of clusters differed. We considered the following types: spatially contiguous clusters, noncontiguous clusters and inter area clusters. The average number of contiguous clusters per recording was significantly larger ($P < 0.0001$) for $R_C$ based clustering (3.0 ± 1.0) compared with weighted overlap based clustering (2.2 ± 1.4). The average number of noncontiguous clusters was significantly lower ($P < 0.0001$) for $R_C$ based clustering (0.4 ± 0.6) compared with weighted overlap based clustering (0.9 ± 0.9). This was also the case ($P = 0.0001$) for the average number of inter-area $R_C$ based clusters (0.1 ± 0.3) compared with weighted overlap based clusters (0.4 ± 0.7).

Coincident-firing cluster STRFs versus all-spike cluster STRFs

Figure 10 shows an example of the STRFs in an $R_C$-based cluster for stimulation with the 20/s multi-tone stimulus ensemble. The cluster contains the activity of 12 electrodes across both electrode arrays positioned within AI (left 3 columns). The percentage of common spikes, defined as those that occur within a 10-ms window across all these 12 electrodes, is only 0.7 but still results in a clearly defined common-spike STRF (top panel in right column) with a CF around 7 kHz. The merged activity of all spikes from the 12 electrodes in the cluster results in an STRF that is shown at the second row of right-most column and the 30% contour line (between peak and mean value) of the common-spike based STRF is superimposed. The common-spike STRF corresponds to the lower half of the all-spike STRF and consequently is not just a diluted version of the all-spike STRF. The signal-to-noise ratio (SNR), defined as $20\log_{10}(\text{peak/mean})$, of the common-spike STRF is 15.3 dB, whereas that for the merged-activity STRF is 4.3 dB. This suggests that synchronized neural activity can enhance signal-to-noise ratios.

Figure 11 shows four comparisons, each from a different recording, of common-spike STRFs with merged all-spike
correlation structure as well as the fraction of common spikes in the cluster.

The common-spike STRFs show an SNR that is significantly larger than for the all-spike STRFs (average 3.7 ± 3.3 dB, minimum: -2.9 dB, maximum: 15 dB; Fig. 12B). Split among multi-tone stimulus type the mean improvement of the SNR was 5.0 dB (20/s), 3.0 dB (120/s), and 1.9 dB (240/s). The values for the 20/s stimulus were significantly higher (P < 0.005) than those for the 120/s and 240/s multi-tone stimuli. Those between the 120/s and 240/s were not significantly different (P = 0.6). The fraction of common spikes in the clusters was significantly higher (P < 0.05) in AI clusters (0.33 ± 0.23, n = 251) compared with PAF clusters (0.20 ± 0.19, n = 29) and AAF clusters (0.14 ± 0.11, n = 5). In case of AAF, the small number cannot exclude a sample bias, and no firm conclusions can be drawn from this statistical difference.

Functional cluster properties

If individual clusters from the same multi-electrode recording show quite different cluster STRFs, one is inclined to consider the clusters functionally different, as they will respond preferentially to different stimuli. This is not always that clear. For example, Fig. 13 shows a set of common- and all-spike STRFs for four clusters from the same recording that feature only slight differences across clusters and those are largely in CF and frequency bandwidth. This tonotopic gradient with very similar excitatory STRFs shapes was the most common finding when the electrode arrays were both in AI. The common-spike STRFs show an SNR that is significantly larger than for the all-spike STRFs (average 3.7 dB). The percentage of common spikes shown above the common-spike STRFs, and the average within-cluster STRF overlap (Fig. 12B). The same is shown for the average between-cluster values in B.

STRFs for multi-tone stimuli with rate of 120/s (A and B) and 240/s (C and D). The percentage of common spikes is shown above the common-spike STRFs, and the maximum and mean values of the STRFs are also indicated. In Fig. 11A, the common-spike STRF (top) for a cluster of six electrodes contained 6.4% of the spikes and was similar in shape to the all-spike STRF albeit reduced in both frequency and time domain. The common-spike STRF had an SNR (3.7 dB) that was higher than for the all-spike STRF (1.8 dB). In contrast, in Fig. 11B, the common-spike STRF (7 electrodes; 7.1% of the spikes) represents only the center of the all-spike STRF and again shows a larger SNR (3.1 dB) compared with the all-spike STRF (1.7 dB). For Fig. 11C, the common-spike STRF (4 electrodes; 33.2% of the spikes) has an SNR of 2.6 dB and the all-spike SNR = 2.0 dB. In this case, both STRFs are nearly identical in shape. In Fig. 11D, (7 electrodes; 10.5% of the spikes) the SNRs are, respectively, 1.6 and 1.4 dB, and consequently the STRFs are both somewhat fuzzy.

The fraction of common-spikes in the cluster is strongly correlated (r² = 0.88) with the ratio of the peak STRF values in common- and all-spike STRFs (Fig. 12A), suggesting that the fraction of common spikes is strongly related to the stimulus-driven firing rate. This again indicates that clustering may reflect the communality in the receptive fields of the thalamic neurons and that this largely determines the cross-

**TABLE 5.** Pearson's correlation between pair-wise STRF overlap and pair-wise Rs for the three multi-tone stimulus ensembles

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>n</th>
<th>Pearson’s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi tone 20/s</td>
<td>30</td>
<td>0.59 ± 0.14</td>
</tr>
<tr>
<td>Multi tone 120/s</td>
<td>29</td>
<td>0.62 ± 0.16</td>
</tr>
<tr>
<td>Multi tone 240/s</td>
<td>17</td>
<td>0.65 ± 0.12</td>
</tr>
<tr>
<td>Total</td>
<td>76</td>
<td>0.62 ± 0.14</td>
</tr>
</tbody>
</table>
all cases, the common-spike STRF overlaps substantially with the all-spike STRF.

Quite often, and especially for clusters comprising a large number of electrodes, the $R_C$-based clustering resulted in a noise-like common-spike STRF. Figure 15 shows an example where the cluster comprised 12 electrodes; the individual STRFs for the 20/s multi-tone stimulus are clear (columns 1–3) but widely scattered in best frequency and with moderate pair-wise correlations (average: 0.032) and pair-wise STRF overlap (average: 0.26). In this case, the similarity in $R_C$ values leading to the clustering may have resulted from a common drive other than that related to frequency tuning. The spike files did not show bursting or spindle-like firing behavior, and the firing patterns were stationary throughout. The all-spike STRF shows three areas centered around 2.5, 7, and 20 kHz; the common STRF does not show a distinct activity pattern. For

FIG. 10. STRFs belonging to 1 12-electrode cluster for 20/s multi-tone stimulation (left 3 columns), the common-spike STRF (top row, right column), and the all-spike STRF (second row, right column). The common-spike STRF represents 0.7% of the spikes recorded on the 12 electrodes and shows a clear activity region between 5 and 10 kHz and 15–25 ms after stimulus onset. The contour of this common-spike STRF is shown superimposed on the all-spike STRF and corresponds with the lower half of it.

FIG. 11. Comparison of common-spike STRF and all-spike STRFs for clusters of several sizes. The 4 recordings represent different animals. In 3 of the 4 cases, the common- and all-spike STRFs overlap nearly completely. In 1 case (B), the common-spike STRF is again much smaller than the all-spike one. The signal-to-noise ratios, defined as peak value divided by the mean (indicated on top of each panel) are larger for the common-spike STRFs compared with the all-spike STRFs.
the other stimuli, the clusters were much smaller in size, in particular for the 120/s multi-tone stimulus the large 12-electrode cluster broke up in two clusters comprising electrodes 5–8 and 9–13 and 16, respectively. The R\(_C\)-based clustering for the 120/s stimulus largely reflected the similarity in STRFs.

Generally, the fraction of common spikes in a cluster decreased with increased cluster size, either defined by number of participating electrodes or as spatial size (Fig. 16). The decrease showed the same trend regardless of the presence or absence of a common-spike STRF, albeit that the fraction of common spikes was larger when a common-spike STRF was present. In our sample, in 141 of 243 cases (58%), a clear common-spike STRF was present. The average fraction of common spikes was 0.33 ± 0.23 (range: 0.005–0.86) when a clear common-spike STRF was present and 0.13 ± 0.16 (range: 0.001–0.68) when it was not. The difference was highly significant (\(P < 0.0001\)), but it is clear that even without a common-spike STRF present the fraction of common spikes can be as large as 68%. There was no effect of multi-tone stimulus type on these values. It is noted that even for clusters as small as two electrodes, and correspondingly spatially close, a common-spike STRF is missing in a substantial number of cases.

It must be emphasized that STRFs are average results and that although high pair-wise correlations may exist, the probability of all units in the cluster firing synchronously to frequencies that make up the STRF decreases with the number of participating units. Under the assumption that the units would generally fire independent of each other in response to a stimulus, but when they fire together show high temporal precision, one could estimate the probability of \(N\) units firing together from the probability of common-spikes for pairs. In our data for clear common-spike STRFs, the mean value of the fraction of common spikes for all electrode pairs was 0.535. This can be interpreted as a single-electrode probability of firing to each stimulus contributing to the STRF of \(\sqrt{0.535} = 0.731\). Thus the fraction of common-spikes in a cluster of size \(N\) would change according to \((0.731)^N\). This results in a slight overestimation of the fraction of common spikes to what was observed. If we estimate the average “probability of common spikes” from the linear regression of log(fraction common) on \(N\) (from Fig. 16A), then one is slightly underestimating the observed mean values of the fraction of common spikes as a function of the number of electrodes.

This was further analyzed for five randomly selected multi-electrode recordings. For each set of recordings, the fraction of common spikes was calculated for all possible pairs of electrodes and then the mean value thereof was used to predict the expected values for all possible three, four, five, six, etc electrode combinations. In all cases, the prediction on basis of that mean values was slightly larger than the actual mean values for the various electrode combinations.

Based on these calculations, one cannot reject the assumption that the decrease in the fraction of common spikes with the number of electrodes in the cluster is due to some independence in firing to the stimulus. This also suggests that most pair-wise correlations contain significant contributions of non-stimulus-related spikes (but likely not driven by brain rhythms in the >1 Hz range because those effects are corrected for in the calculation of \(R_C\)).

**DISCUSSION**

We have shown that neural activity recorded from the auditory cortex can be clustered on basis of pair-wise cross-correlation. The resulting neuron clusters are local and typically do not cross cortical-area boundaries and are even confined to a particular electrode array in the majority of cases. As a consequence, the spatial size of the clusters rarely exceeds a few mm and on average is ~1 mm. Cluster size was significantly larger in PAF compared with AI, whereas the fraction of common spikes (within a 10-ms window) across electrodes participating in a cluster was significantly higher in AI compared with PAF. The difference in the fraction of common spikes could result from the longer latencies and hence larger SD in first spike firing of PAF neurons (Phillips and Orman...
The average cluster size may reflect the decrease in $R_C$ with distance (Fig. 6). However, within AI this decrease is much slower with distance than in PAF, yet cluster size is on average larger in PAF. Thus other factors play a role in the local changes in $R_C$. For instance the difference between the average within-cluster $R_C$ values is much larger in AI than in PAF (Table 2) so that the cluster boundaries may be less sharp in PAF.

Large $R_C$-based clusters comprising activity from both multi-electrode arrays were observed only 11 times (4.5%). These large clusters were never found for silence or low-pass amplitude-modulated noise and were most frequently encountered for the Poisson-distributed click trains and the 20/s multi-tone stimulus. These two stimuli have distinct onsets with sufficiently long silent periods between the stimuli. In general, cluster boundaries changed under different stimulus conditions.
conditions. Taking the cluster sizes found for spontaneous firing as a reference, Poisson-distributed clicks and low-pass amplitude-modulated noise produce the largest changes in cluster size, whereas multi-tone stimuli have a more modest effect. This suggests that the clusters found under spontaneous firing conditions are largely determined by the same common inputs that govern the STRF properties of the neurons, i.e., the specific thalamic afferents. Topographic mapping of activity in visual cortex using intrinsic optical signals also showed a strong similarity in the locations of patches of synchronous spontaneous firing and evoked firing of the order of 1 mm in size (Tsodyks et al. 1999). This is furthermore reflected in the fact that 70% of the cluster sizes did not change significantly between the silence and multi-tone stimuli condition, whereas only 40% stayed the same in comparison between silence and broadband stimulation. Thus broadband stimulation with low-pass amplitude-modulated noise has very different effects than stimulation with a multi-tone 240/s stimulus, which has a similar amplitude-modulation spectrum and the same long-term carrier spectrum but of course sounds completely different.

Coarse-grained recording and homogeneity in response

Responses for units recorded with the same electrode in AI show strong homogeneity for CF, threshold at CF, 20-dB bandwidth, and minimum latency (Eggermont 1996). The spatial representation in AI also changes slowly for those parameters (Schreiner et al. 2000). The information about a stimulus in AI is most likely contained in this slowly changing spatial representation of neural activity, i.e., in a population code, thus the size and shape of the clustered activity should be important. It is therefore assumed that for mapping those parameters one can use a relatively coarse sampling as done in the present study. There is one caveat, the continuity in mapping is generally observed for near threshold properties (CF, threshold at CF, 20-dB bandwidth), whereas the stimuli used in the present study were all presented at 65 dB p.e. SPL, and thus well above the level of the 20 dB bandwidth of the neurons. For those intensity levels, tone pips activated a large (∼4 mm along an isofrequency contour as well as perpendicular to it) and patchy area in AI (Phillips et al. 1994), so there is no a priori reason to expect spatially contiguous clusters. Furthermore, our two electrode arrays were mostly positioned in a caudal to rostral alignment (as in Fig. 8) and only infrequently within the direction of the isofrequency sheets. Yet, 89% of the $R_C$-based clusters on the same electrode array were contiguous. In contrast, and probably relating more to the Phillips et al. (1994) findings, 26% of clusters based on STRF overlap were noncontiguous. This suggests that despite the potentially significant STRF overlap for neurons in different patches, there is probably a weaker correlation between different patches than within a single patch.

Should one correct for stimulus-correlation?

Calculating a shift predictor or joint-PSTH predictor is one way to extract the effects of stimulation on the effective neural
connectivity (Espinoza and Gerstein 1988). However, as we have previously argued (Tomita and Eggermont 2005), it is unlikely that the nervous system performs a correction for stimulus-induced correlation as estimated by the various predictors. It is the actual spike coincidences that are affecting the potential for firing in a target neuron not the stimulus-corrected ones. Thus our raw correlations may effectively estimate those coincident firings between neurons that could play a role in neural population coding of sound. The effect of removing common periodicities in spike firing and common bursting activity form the cross-correlation can be justified because these contributions typically do not occur in awake animals. Nevertheless, our data show that in the case of the Poisson-distributed clicks, the effects of a stimulus correction based on the first-order Poisson kernels on clustering are minimal and do not affect our conclusions. In case of the multi-frequency stimuli, one expects a larger reduction of the peak cross-correlation from corrections for stimulus locked contributions. This is in first approximation shown by the clustering on basis of STRF overlap. The STRFs are a linear predictor of stimulus effects on the firing rate (or vice versa), and the overlap of STRFs of two recordings is a linear estimator of the fraction of spikes they have in common. Thus the STRF-overlap would correspond to the area under the PSTH-product predictor. Given that the STRF overlap explains 38% of the variance in the pair-wise $R_C$, and that $R_C$-based clusters for multi-frequency stimuli are similar, but not identical, to those based on spontaneous activity, one has to conclude that clustering based on stimulus-corrected correlation coefficients would also be similar.

**Basis for clustering**

Cross-correlation measures allow the derivation of the functional ordering of neurons exemplified in the pair-wise peak-correlation matrix. Functional organizations that show topographic order, such as the tonotopic organization of the auditory cortex, are in fact only observable by the experimenter, whereas the animal has likely to use correlation measures to access that order (Koenderink 1984a,b). From the similarity of $R_C$-based and STRF-overlap-based clustering, one could, however, claim that the correlation is largely caused by the topographic order of the common inputs. However, <40% of the variance in $R_C$ was explained by STRF overlap, suggesting dominant other contributions.

The correlation measure used is calculated for multiple single-unit activity based on spike-sorting procedures. How many neurons are recorded from with a single microelectrode? Spike separation usually results in two to three well-resolved units and a remaining spike cluster that cannot be resolved. Typically the firing rates of the sorted groups are of the same order of magnitude and can each reach 10/s (Eggermont et al. 1993; Legendy and Salcman 1985). However, optical recordings with single-neuron resolution in vivo from motor cortex in urethan-anesthetized rats and follow-up patch-clamp recordings suggest that individual cortical neurons have a low probability of firing (Kerr et al. 2005). Spontaneous activity in layer II/III of motor cortex during up states was found to be <0.1/s, and only 10% of neurons were simultaneously active during an up state. DeWeese et al. (2003) found that patch-clamped neurons in auditory cortex (ketamine anesthesia) fired either 0 or 1 spike. These findings may suggest a different selection bias for standard extra cellular recordings and patch-clamp recordings (see also Margrie et al. 2002), they may indicate a profound cortical layer difference in firing rate or suggest that the number of units contributing to a well-sorted cluster in extra cellular recordings is grossly underestimated. In awake paralyzed cats, 40% of neurons in AI had spontaneous firing rates <1 spike/s, but no significant layer difference in response to tones was found (Abeles and Goldstein 1970). If the number of contributing units is underestimated, even well-separated single-unit recordings must then be comprised of activity from ≤10 not-well-synchronized units (otherwise the spikes would be superimposed and likely sorted as 1 unit).

Clustering is based on the distribution of $R_C$ values across electrode pairs. A potential cause for changing correlation strengths under different stimulus conditions is the use of unresolved multi-unit activity on pairs of electrodes to probe neural correlations (Bedenbaugh and Gerstein 1997; Gerstein 2000). As these authors note, the peak cross-correlation coefficient between small groups of units is generally larger than the cross-correlation between a pair of single units from each group. However, the measured inter-electrode correlation is reduced in turn by a nonlinear function of the cross-correlations between the single units recorded on each electrode. This becomes important if the number of units recorded from under different stimulus conditions would change, and this would in turn change the measured cross-correlation strength between...
the two unit clusters. From our spike-separation procedure, which resulted in reasonable well-isolated spike-waveforms, we could not detect new units activated by different stimuli. Thus the changes in $R_C$ clustering are likely not the result of sampling different sets of units for different stimuli. Although if one accepts that single units fire $<1$ spike/s, then even close similarity in spike waveform cannot rule out activation of different units by different stimuli.

What causes changes in clustering?

Following the definition of cluster membership, changes in $R_C$-based clustering are the result of changes in the distribution of pair-wise cross-correlation. In its most simple form, changes in clusters could result from recording different units for different stimulus conditions and hence obtaining different correlation strengths. As discussed in the preceding text, I consider this not very likely.

Changing anesthesia levels can also cause changes in the peak cross-correlation coefficient (Erchova et al. 2002). With deeper anesthesia levels, the spontaneous firing activity recorded with microelectrode arrays in rat barrel cortex took the shape of periodic bursting with periods of $\sim 2$–$2.5\,s$, and the firings became more synchronized between barrels. The peak correlation coefficient for this spontaneous activity increased by about a factor 4 between light and deep urethan anesthesia (Erchova et al. 2002). It is possible that in the course of our 1- to 1.5-h recording that comprised the four to six different stimulus presentations of 15 min each, the anesthesia level changed sufficiently to affect the cross-correlation strength. In contrast to Erchova et al. (2002), we correct for the effect of firing rate, common periodicities and bursting behavior in the neurons on the cross-correlation, so we expect that slight changes in anesthesia levels will have very little effect on the peak $R_C$ values. This is supported by the fact that both the mean value and the variance of the pair correlation matrix did not change significantly with time.

It is also possible that the difference in the clustering of neurons for a long-lasting period of silence and for the other stimuli is related to a state change in the brain caused by appropriate stimulation (Miller and Schreiner 2000). These authors showed that during stimulation with dynamic-ripple noise the spindle-wave-like periodicity in the autocorrelograms of individual units recorded from the auditory thalamus disappeared. It was argued that the random frequency and temporal changes in such a stimulus, clamped the thalamocortical system in an “aroused” state. The multi-tone stimuli that we use also have this property (we found that during presentation of these stimuli the 7- to 14-Hz spindle rhythm in cortical local field potentials and in spike autocorrelograms can be completely suppressed; T. Britvina and J. J. Eggermont, unpublished observations).

One of the properties attributed to a neural assembly is its ability to change neuron membership as a result of a change in brain state (Abeles et al. 1995). The multi-tone stimuli that we employ have a small but distinct effect on the brain state (Abeles et al. 1995). The multi-tone stimuli that we use are the result of competitive interactions resulting from the demand to map several stimulus parameters on the same cortical surface (Schreiner 1995). Neural assemblies are generally proposed to explain particular forms and properties of population coding. The point is then whether neural clusters and neural assemblies have a common base. The brain-state dependence in cluster membership mentioned in the preceding text is one property that the neuron clusters have in common with putative neural assemblies. Our clusters appear to be largely confined by cortical area boundaries and are affected by stimulus type. Only three $R_C$-based clusters (4.6% of the number of simultaneous PAF and AI recordings) were found that crossed area boundaries potentially because of a drop in pair-wise correlation strength across area boundaries as noticed before (Eggermont 2000). This drop in cross-correlation strength may reflect the paucity of bifurcating axons that innervate cortical cells in different cortical areas (shown for AAF and AI) (Lee et al. 2004a), thereby only allowing a modest common input from thalamic neurons to cortical neurons. This forms another argument for our assumption that the clustering reflects anatomical connectivity. In contrast, STRF overlap-based clustering resulted in $\sim 15\%$ of the number of simultaneous PAF and AI recordings as inter-area clusters, which was significantly higher ($P < 0.005$) than for $R_C$-based clusters. Assuming that overlap in STRF implies overlap in thalamocortical afferent input, but not necessarily provided by bifurcating axons, this suggests that purely stimulus-based correlations do not play a large role in the $R_C$-based clustering compared with connectivity-based correlations.

Anatomical connectivity may also underlie the synchrony for sensitivity to visual stimulus orientation in spatially segregated patches. For example, gamma-oscillation synchrony in the visual cortex crosses area boundaries and is driven by common sensitivity to stimulus aspects such as orientation sensitivity has also been related to neural assembly activity. This inter-area synchrony was frequently observed (König et al. 1995) and could reflect a more widespread common input between different cortical visual areas compared with the auditory ones. It should be mentioned that gamma oscillations are conspicuously absent in spike firing from auditory cortex (Horikawa et al. 1994). Comparison of the topological organization of the auditory and visual systems (Scannell et al. 1995) shows that striate cortex is interconnected with eight other visual cortical areas and AI with five other auditory cortical areas, albeit that Winer (1992) lists interconnections of AI with seven areas. This suggests no a priori differences in connectivity. Whereas in the visual system intercortical connections largely reflect similar orientation columns and ocular dominance columns (Schmidt et al. 1997), in the auditory system,
this connectivity may reflect besides frequency tuning only binaural interaction type (Winer 1992). The fact that our stimuli were presented in the contralateral field may have restricted the probability to observe inter-cortical area clusters. Another reason might be that PAF and AI are not directly connected (Rouiller et al. 1991) and any clustering would thus be based solely on shared thalamo-cortical afferents.

\[ R_C \] and weighted STRF overlaps in AI were larger than for PAF and AAF, suggesting that the divergent thalamocortical fiber activity is either more synchronized in AI or that the afferent fibers make stronger synaptic connections with the pyramidal cells than is the case in PAF and AAF. The longer response latencies and the temporally less precise STRFs in PAF indicate that the lower input synchrony in PAF plays a large role in its lower \( R_C \) values. In cat AAF, latency is typically shorter than in AI (Eggermont 1998) and STRFs are well defined, but within-cluster \( R_C \) values are even smaller than in PAF. Therefore the different origin of the thalamo- and cortico-cortical afferents and their connection strengths in the three areas may be at the base of the differing \( R_C \) values.

The STRF determines part of the functional operation of a cluster in the auditory cortex. The STRFs can be considered as spectrotemporal filters that take partial derivatives of the sound, i.e., in time (modulation envelope) and frequency (Fishbach et al. 2001; Koenderink and van Doorn 1990) and thus act as modulation selective filters (Chi et al. 2005). One can ask what is more important for describing the functional properties of a cluster, its common-spike STRF or the one calculated from all spikes from the units taking part in the cluster. The common-spike STRF is a subset of the all-spike STRF. In some cases, however, properties from the common-spike STRF largely mimic those from the all-spike STRF (Fig. 11, A, C, and D).

One of the things a cluster accomplishes is that it extracts common-spike STRFs with higher signal-to-noise ratio than that of population STRFs that result from merging all spikes emitted by the cells participating in the cluster. This suggests that an all-spike STRF is preferable over a common-spike STRF as far as stimulus encoding is concerned. This, however, has to be tested.

Limitations of this clustering approach

The clustering was deliberately studied for steady-state stimuli that each lasted 15 min and allowed stationary recordings and good statistics for the neuronal firing and synchrony. In reality, the brain has to deal with activity that changes meaning and context over time scales of the order of seconds (Aertsen et al. 1994; Ahissar et al. 1992; Eggermont 1994). Thus time-dependent evaluation of correlations and dynamical clustering need to be considered. The statistically less robust results of using short-duration spike trains will likely result in less stable clusters, which either reflect the noisiness of the data or the actual transient nature of neural assemblies. However, the gravitational clustering method in its newest adaptation (Lindsey and Gerstein 2006) has potential for this purpose.

Is pair correlation sufficient or even meaningful as a basis for clustering? It is for linear systems. It is also sufficient for a nonlinear system with only feed-forward connections, but it is not in the case of nonlinear systems with feedback connections (Johannesma et al. 1986). Because the thalamo-cortical system is obviously nonlinear and has strong feedback connections, one potentially needs to extend this analysis by incorporating nonlinear relations, e.g., by using mutual information as a basis for clustering.

Summary

Neuron clusters in auditory cortex formed on basis of the neural correlation matrix, reflecting the firing synchrony across cortical areas of several mm², expand and contract in response to different stimuli. The most striking result was that the cluster positions and size were very similar for spontaneous activity and multi-tone stimuli evoked activity. Clusters rarely comprised neurons from more than one cortical area, as follows from the decrease in average correlations strength across area boundaries. Clusters are therefore not synonymous with the view of neural assemblies, advanced for visual cortex, that assumes synchrony across several areas. Several operations on the coincident output of the clusters, as reflected in the common-spike and all-spike STRFs, may serve different modes of stimulus coding.

Acknowledgments

G. Shaw provided programming assistance. A. Noreha, M. Tomita, and N. Aizawa helped with the data collection. A. Noreha and B. Gourevitch commented on an earlier version of this manuscript.

Grants

This work was supported by the Alberta Heritage Foundation for Medical Research, by the Natural Sciences and Engineering Research Council, and by the Campbell McLauren Chair of Hearing Deficiencies.

References


