Behavioral/Systems/Cognitive

Auditory Cortical Activity after Intracortical Microstimulation and Its Role for Sensory Processing and Learning

Matthias Deliano,[1] **Henning Scheich,**[1] **and Frank W. Ohl**[1,2]

[1]Leibniz-Institute for Neurobiology, D-39118 Magdeburg, Germany, and [2]Institute of Biology, Otto-von-Guericke-University, D-39118 Magdeburg, Germany

Several studies have shown that animals can learn to make specific use of intracortical microstimulation (ICMS) of sensory cortex within behavioral tasks. Here, we investigate how the focal, artificial activation by ICMS leads to a meaningful, behaviorally interpretable signal. In natural learning, this involves large-scale activity patterns in widespread brain-networks. We therefore trained gerbils to discriminate closely neighboring ICMS sites within primary auditory cortex producing evoked responses largely overlapping in space. In parallel, during training, we recorded electrocorticograms (ECoGs) at high spatial resolution. Applying a multivariate classification procedure, we identified late spatial patterns that emerged with discrimination learning from the ongoing poststimulus ECoG. These patterns contained information about the preceding conditioned stimulus, and were associated with a subsequent correct behavioral response by the animal. Thereby, relevant pattern information was mainly carried by neuron populations outside the range of the lateral spatial spread of ICMS-evoked cortical activation (∼1.2 mm). This demonstrates that the stimulated cortical area not only encoded information about the stimulation sites by its focal, stimulus-driven activation, but also provided meaningful signals in its ongoing activity related to the interpretation of ICMS learned by the animal. This involved the stimulated area as a whole, and apparently required large-scale integration in the brain. However, ICMS locally interfered with the ongoing cortical dynamics by suppressing pattern formation near the stimulation sites. The interaction between ICMS and ongoing cortical activity has several implications for the design of ICMS protocols and cortical neuroprostheses, since the meaningful interpretation of ICMS depends on this interaction.

Introduction

Electrical stimulation of sensory cortex has recently regained interest as a tool for establishing causal links between cortical activity and perception (Cohen and Newsome, 2004). Furthermore, in clinical research, sensory cortical neuroprostheses are developed with the aim of replacing lost functions of afferent pathways based on electrical stimulation (Brindley and Lewin, 1968; Dobelle et al., 1973; Normann et al., 1999; Donoghue, 2002; Kipke et al., 2008).

This renewed interest in sensory cortex stimulation originates in part from experiments using intracortical microstimulation (ICMS) techniques, which yield defined perceptual effects based on a focal cortical activation. In animal models it has been shown that electrically evoked percepts vary in accordance with the location of the stimulation site within retinotopic (Bradley et al., 2005), tonotopic (Scheich and Breindl, 2002; Otto et al., 2005), and somatotopic (Leal-Campanario et al., 2006; Fitzsimmons et al., 2007) map representations. Several studies have specifically

demonstrated that psychophysical properties of electrically evoked percepts can be linked to the representational topographies of cortical maps at the stimulation site (Romo et al., 2000; Bartlett et al., 2005), in some cases even to tuning properties of small sets of directly excited neurons located within one or a few cortical columns (Tehovnik et al., 2006).

Although psychophysical properties of electrically evoked percepts can often be explained by the local activation of cortical representations, the puzzling question remains how the focal, artificial activity imposed onto an extended, complex network like the cortex (Butovas and Schwarz, 2003) can yield a meaningful signal that can be interpreted in the context of a behavioral task. Thereby, beyond the local encoding of stimulus features, learning plays an important role, as by learning, meaning is attributed to the electrical stimuli (Scheich and Breindl, 2002; Bradley et al., 2005; Fernández et al., 2005; Leal-Campanario et al., 2006). When animals learn to discriminate or categorize natural, e.g., visual, auditory, or somatosensory stimuli, large-scale stimulus-specific patterns emerge from the ongoing activity of sensory cortex as has been demonstrated by recording electrocorticograms (ECoGs) at high spatial resolution (Barrie et al., 1996; Freeman, 2000; Ohl et al., 2001). Whereas evoked cortical activity mainly carries information about physical stimulus properties (Ohl et al., 2000), these distributed patterns carry learning-dependent information about the behaviorally relevant stimulus classes, and thus rather reflect the meaning attributed to stimuli
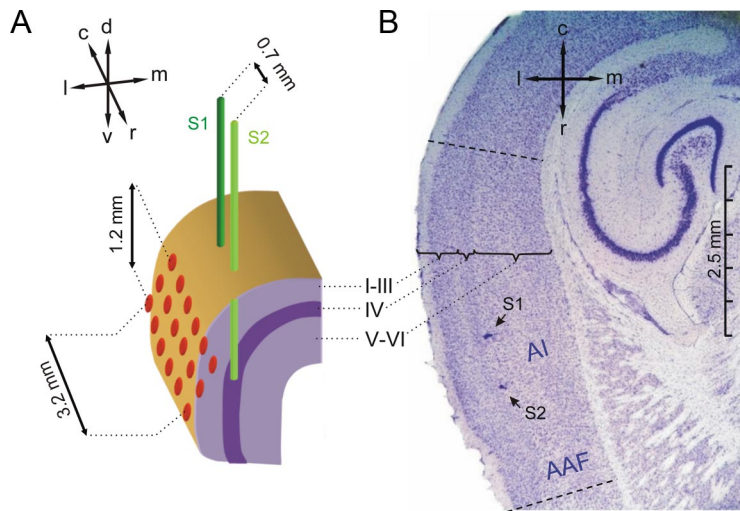
Figure 1. Positioning of electrodes. **A**, A pair of stimulation electrodes S1 (dark green) and S2 (light green) was implanted into the depth of the right AI close to its input layer IV. Electrode tips were positioned along the rostrocaudal axis of AI (caudal electrode S1, rostral electrode S2) with an interelectrode distance of ~0.7 mm. A 3 × 6 ECoG recording array (red) with 0.6 mm interelectrode distances was centered epidurally over the right AI. Anatomical directions are indicated by arrows (d, dorsal; c, caudal; l, lateral; m, medial; r, rostral; v, ventral). **B**, Nissl-stained horizontal section showing two small lesions (arrows) at the tips of the two implanted stimulation electrodes indicating their location within temporal cortex. The cortical layering is marked by the curly brackets. The rostrocaudal extent of the auditory core fields [AI and anterior auditory field (AAF)] is indicated by dashed lines, and can be approximately determined from the pronounced layer IV in these fields identifiable by the band of densely packed cells at a cortical depth of ~0.5 mm.

within a behavioral context by learning (Ohl et al., 2001; Ohl and Scheich, 2005). Here, we investigate how the focal, artificial activation of sensory cortex by ICMS gives rise to similar spatial activity patterns, when animals learn to perceptually interpret electrical stimuli. For this, we trained gerbils to discriminate closely neighboring ICMS sites within primary auditory cortex evoking responses largely overlapping in space. In parallel, we recorded ECoGs at high spatial resolution from the stimulated area. A combined analysis of evoked potentials and of the ongoing ECoG using a multivariate pattern classification procedure allowed us to shed light on the interaction between ICMS and the ongoing activity, which generates patterns that constitute meaningful signals with respect to the behavioral interpretation of ICMS.

Materials and Methods
In our study we trained 6 Mongolian gerbils (*Meriones unguiculatus*) to discriminate between two different sites of ICMS at a distance of ~0.7 mm across the tonotopic gradient of right primary auditory cortical field AI (Figs. 1, 2A). During training, we recorded ECoGs at high spatial resolution (~0.6 mm interelectrode-distance) from the 18 channels of a 3 × 6 electrode array positioned epidurally onto AI, above the simulation electrodes (Fig. 1A). The recording array covered the entire tonotopic gradient along the rostrocaudal extent of AI, about two thirds of its dorso-ventral extent, and parts of the neighboring auditory fields. The male, adult Mongolian gerbils had an age of ~6 months, weighed 85–100 g, and were housed under a 12 h light-dark cycle with free access to water and food between experimental sessions. All animal experiments were surveyed and approved by the animal care committee of the Land Sachsen-Anhalt (No. 43.2-42502/2-325; IfN MD).

Multielectrode arrays. For the application of ICMS, a 2 × 1 array of depth electrodes was built from two Teflon-insulated platinum-iridium microwires (51 μm diameter) aligned in parallel at distance of ~0.7 mm. For recording of high resolution multichannel ECoGs, rectangular 18-electrode arrays were built from Teflon-insulated stainless steel microwire (256 μm diameter) arranged in a 3 × 6 matrix and embedded in dental acrylic at a distance of ~0.6 mm (Fig.

1) (cf. Ohl et al., 2000). The interelectrode distances of ~0.6 mm had been optimized by spatial spectral analysis of recorded potentials (cf. Freeman and Baird, 1987).

Preparation and electrode implantation. Electrode arrays were chronically implanted under deep anesthesia, induced by halothane (4 vol %), and maintained by ketamine (20 mg/100 g body wt) and xylazine (2 mg/100 g body wt). Cranial skin was removed from the interoccipital, parietal, and frontal bones. A Teflon-insulated stainless steel microwire (256 μm diameter), and a Teflon-insulated platinum-iridium microwire (128 μm diameter), both with bared and looped endings, were placed onto the median blood sinus through holes drilled between the parietal bones. The stainless steel microwire served as common ground/reference for recording, and the platinum-iridium microwire as return electrode for ICMS.

Craniotomy over the right temporal cortex and positioning of the electrode arrays was guided by anatomical landmarks, by stereotactic coordinates, and by the visual inspection of the vascularization pattern (Ohl et al., 2000). Stimulation electrodes were implanted through the incised dura into the depth of AI in a sagittal plane parallel to the lateral surface of the lissencephalic cortex. Implantation depth was approximated by a medial offset of ~0.5 mm of this plane to the lateral surface, close to the thalamic input layer IV of AI. Electrode tips were oriented along the rostrocaudal axis and thus positioned along the tonotopic gradient of AI (Figs. 1, 2).

The recording array was positioned epidurally onto AI, right above the tips of the stimulation electrodes. Its length axes was rostrally inclined into the ventral direction, 10° from the horizontal plane (Fig. 1A) to achieve orientation parallel to the tonotopic gradient in AI (Ohl et al., 2000). All electrodes and connectors were fixed with dental acrylic. After implantation, animals were allowed to recover for 3–5 d.

Histology. To verify the positioning of the stimulation electrodes and to analyze the effects of chronic implantation and ICMS, animals were deeply anesthetized using halothane (4 vol %) and killed by an intrapulmonary injection of T61 (Intervet) ~1 week after finishing the training. Nissl-stained sections (40 μm) were prepared from the removed, frozen brain cut in a cryostat microtome (Fig. 1B).

Training, stimulation, and recording. Training was performed in an acoustically and electrically shielded shuttle-box, in 7 consecutive sessions with sessions succeeding in intervals of 1–3 d (Ohl et al., 1999). Before the start of each session, animals were allowed to explore the shuttle-box for 3 min, to adapt to the environment. At the beginning of each trial ICMS delivered through one of the two stimulation electrodes served as conditioned go-stimulus (CS$_{go}$), and stimulation at the other electrode as conditioned no-go-stimulus (CS$_{no-go}$), respectively (Fig. 2A,B). The use of caudal and rostral electrodes for go- and no-go stimulation was balanced across animals (Fig. 1). ICMS was generated by an isolated pulse stimulator (model 2100, A-M Systems), and consisted of 610 ms long, 100 pps trains of current pulses. Pulses were biphasic (charge-balanced), cathodic first, with 200 μs phase duration, and 50–80 μA current amplitude.

After go-trials, animals received a footshock as unconditioned stimulus (US) after a randomly selected delay of 3.5 s, 4.5 s, or 5.5 s after the onset of the CS$_{go}$, unless it changed its compartment within the delay (Fig. 2A). The US lasted for up to 6 s and could be terminated by the animal in changing the compartment. In no-go-trials, an unconditioned error-footshock (US$_e$) was given for 1 s, if the animal did not stay in its momentary compartment for at least 3.5 s (Fig. 2B). Current amplitudes for US and US$_e$ were adjusted individually for each animal between 200 and 600 μA (Ohl et al., 1999). By introducing a delay of at least 3.5 s
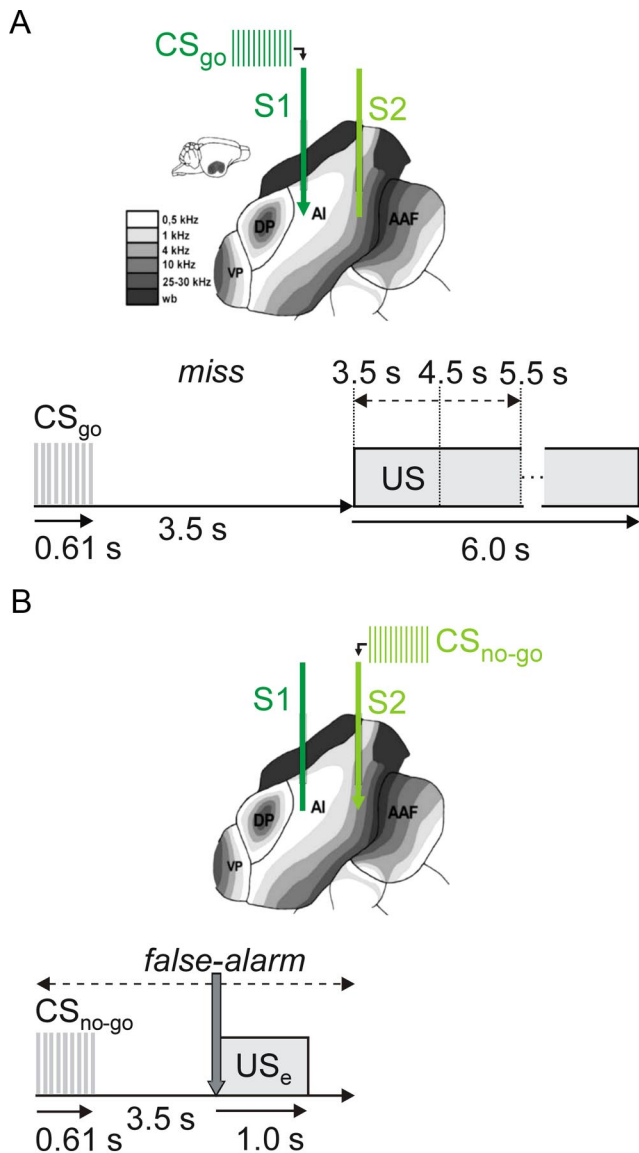
A

B

Figure 2. ICMS and behavioral paradigm. **A**, **B**, Gerbil AI has a tonotopic organization with a caudal-to-rostral gradient of increasing frequencies. Accordingly, the caudal stimulation electrode S1 was positioned in a lower-frequency part of AI than the rostral electrode S2. The schematic views show the case of a CS_{go} delivered at S1 (**A**) and a CS_{no-go} at S2 (**B**). Using a go/(no-go) active avoidance paradigm, animals were trained in a two-compartment shuttle-box to discriminate between two sites (S1 and S2) of ICMS (0.61 s train length). At the beginning of go-trials (**A**), ICMS was delivered through one of the two stimulation electrodes (CS_{go}), and in no-go-trials (**B**) through the other electrode (CS_{no-go}), respectively. Within a random delay after CS_{go} onset of 3.5, 4.5, or 5.5 s (indicated by the dashed arrow in **A**), animals had to change the shuttle-box compartment (go response, CR_{go}), to avoid a mild electrical foot shock (US, 6.0 s maximum duration, terminated by compartment change). When the CS_{no-go} was presented, animals had to stay in their compartment (no-go response, CR_{no-go}) to avoid a mild electrical error-footshock (US_e) with 1 s duration, given immediately after a compartment change (gray arrow in **B**) within the first 3.5 s after CS_{no-go} onset. Behavioral responses were sorted in one of four categories: hit (CR_{go} to the CS_{go}), miss (CR_{no-go} to the CS_{go}), false-alarm (CR_{go} to the CS_{no-go}), and correct-rejection (CR_{no-go} to the CS_{no-go}).

between conditioned stimulus (CS) onset and US onset, we performed trace conditioning to put sufficient demand on cortical processing (Leal-Campanario et al., 2006).

In each session, 60 CS_{go} and 60 CS_{no-go} trials were presented in a pseudo-random sequence with an intertrial interval varying randomly between 15 and 19 s. If the animal changed compartment within 3.5 s after CS_{go} onset, its response was counted as a hit, otherwise as a miss. If the animal changed compartment within 3.5 s after CS_{no-go} onset, its

response was counted as a false-alarm, otherwise as a correct-rejection. For each trial, response type and reaction times were stored to disk by the shuttle box system.

Concurrently with the training, ECoGs were recorded from all electrodes of the 6×3 epidural surface array monopolarly against a common reference/ground-electrode. The signals recorded from the 18 electrodes were fed into an amplifier by a harness of thin, flexible cables permitting the animal to move freely and to turn around in the shuttle-box. The signals were amplified either 5000 or 10, 000 times and bandpass filtered between cutoff frequencies of 0.1 Hz and 100 Hz (Xcell-3x4; FHC). Data from 2 s prestimulus and 3 s poststimulus were sampled at 1 kHz and stored to disk.

Analysis of behavioral data. Rates of hit and false-alarm responses calculated session-wise for each animal. To quantify the behavioral discrimination performance (d') independent of experimental conditions biasing the response of the animal (Green and Swets, 1966), d' values derived from signal detection theory were calculated for each session and animal, as well. For this, z-scores of hit and false-alarm rates were derived from the inverses of a standardized normal distribution function, respectively. d' values were then calculated by subtracting these z-scores. For a statistical analysis of hit- and false-alarm rates and of d' values, repeated-measure ANOVAs were calculated using SPSS 8.0. If necessary, we applied a Huynh-Feldt correction of sphericity (Huynh and Feldt, 1976).

Preprocessing of electrophysiological data. Digitized ECoG-recordings were analyzed session-wise and trial by trial using Matlab (The Math-Works). Amplifier clippings, as they occurred with movements of the animal, and large, fast oscillations in the recording associated with muscle artifacts were identified by visual inspection and regarded as artifacts. Trials containing these artifacts at more than two channels were excluded from analysis. In trials with no more than two channels affected by the artifact, the signals from bad channels were replaced by averages across the signals at their neighboring channels. The first session of training was excluded from analysis, because of the high incidence of movement artifacts in this session with sometimes <50% of the trial remaining after artifact removal.

In our ECoG-recordings, an electrical stimulus artifact occurred with each ICMS pulse which consisted of a short ($\sim$1 ms) negative peak followed by a positive peak decaying within <8 ms (see Fig. 4A). Hence, the pulse train (61 pulses at an interval of 10 ms) created an artifact component in a narrow frequency band around 100 Hz (see supplemental Fig. 1A, available at www.jneurosci.org as supplemental material). An interpolation procedure was chosen to remove these electrical stimulus artifacts, because we wanted to exclude all ECoG data points affected by this artifact from any further analysis. For this, in the time window from 1 ms before to 609 ms after stimulus onset, all data points were discarded except for those unaffected by the artifact 1 ms before and/or 9 ms after the onset of each pulse. The signal was then reconstructed by interpolating between these remaining points which expanded the down-sampled input vector (62 "artifact-free" data points) to its original length (609 data points). This was achieved by inserting zeros between the remaining data points, and by applying a specially designed symmetric FIR filter that minimized the mean-square errors between the interpolated points and their ideal values (Matlab's interp.m function, The MathWorks) (cf. Digital Signal Processing Committee, 1979). The procedure was applied to the raw signal, separately at each channel before digital filtering and further analysis. By this, we avoided filter-ringing due to the sharp artifact peaks (see supplemental Fig. 1C, available at www.jneurosci.org as supplemental material). Down-sampling and interpolation sufficiently suppressed the electrical stimulus artifact (see supplemental Fig. 1B, available at www.jneurosci.org as supplemental material), and did not introduce any changes in the raw signals outside the interval from 1 ms before to 609 ms after stimulus onset. Also, the distortion of the evoked signal by down-sampling to 100 Hz and subsequent interpolation in the stimulation interval (-1 to 609 ms) was small, as the spectrum of the evoked potential was dominated by frequencies below 50 Hz (see supplemental Fig. 1B, available at www.jneurosci.org as supplemental material).

Nevertheless, after artifact removal, our analyses included digital filtering in the β- and γ-band (15 Hz to 80 Hz), and time windowing (time

frames of classification), by which remaining minor distortions of the evoked signals could have still affected results outside the interval from 1 ms before to 609 ms after stimulus onset. To exclude any of these effects, the time interval affected by ICMS was defined from 290 before to 900 ms after stimulus onset. This interval included the 610 ms pulse train, the half lengths of time frames of classification (90 ms; see below), and the length of the filter kernel (200 ms; see below).

After the removal of the electrical stimulus artifact, 50 Hz power line interference was largely reduced by subtracting a 50 Hz sine wave fitted channel- and trial-wise to the data.

Analysis of electrically evoked potentials. Electrically evoked potentials (EEPs) were calculated separately for each channel, stimulus condition (CS_{go} or CS_{no-go}), session, and animal by averaging ECoGs across artifact-free trials in reference to stimulus onset. The mean of a baseline 1.0 s before stimulus onset was subtracted to remove offsets. To characterize the early input evoked by ICMS, we determined the amplitude and the latency of the first peak of the EEP in a time window from 0 to 50 ms after stimulus onset at its maximum across channels. Furthermore, the spatial distribution of the EEP amplitude across the recording array was determined at the maximum peak latency. The spatial maximum of this distribution of EEP amplitudes was used to precisely define the position of the CS_{go} and the CS_{no-go} stimulation sites relative to the coordinates of the recording array.

To determine the lateral spatial spread of the early cortical activation evoked by ICMS, we analyzed the variation of EEP amplitudes at the maximum peak latency as a function of distance to the stimulation site. For this, separately for the CS_{go} and the CS_{no-go}, each spatial distribution of EEP amplitudes was normalized to zero and one between the maximum peak amplitude at the stimulation site and the lower amplitude offset in each session (see Results). This offset was estimated by the mean of the three lowest EEP amplitude values at the maximum peak latency. We then calculated for each recording channel its distance to the stimulation site, respectively. Close to the stimulation sites, channels with equal distances were binned. At larger distances broader bins were used, to obtain amplitude values for the same bins in each stimulus condition and animal. This was due to the fact that the position of stimulation sites varied relative to the recording array across stimulus conditions and animals (see Results). For each bin, normalized EEP amplitudes were averaged across the channels of the bin. After averaging across sessions, the resulting distance functions of the normalized EEP could be fitted well with a Gaussian (Freeman, 1975) using nonlinear least-square fitting (Press et al., 2007). Adjusted goodness-of-fit statistics for grand mean distance functions weighted by the variances in the means of the data points across animals yielded $\chi^2 = 11.14$, $p = 0.29$ for the CS_{go}, and $\chi^2 = 4\,0.47$, $p = 0.68$ for the CS_{no-go}. This allowed us to quantify the range of the lateral spread of early ICMS-evoked activity by the half-width of the Gaussian fitted to the mean distance functions of the EEP amplitude in response to CS_{go} and the CS_{no-go}, respectively.

Analysis of spatial patterns in the ongoing ECoG using a multivariate classification procedure. As has been shown previously, spatial patterns containing information about the behaviorally relevant stimulus classes emerge from the ongoing cortical activity with learning. Such learning-dependent activity patterns are carried by aperiodic oscillations in the β- and γ-range of the ECoG, and cannot be retrieved by temporal ensemble averages of the ECoG across trials, because at each time a pattern emerges the waveform carrying the pattern is different (Barrie et al., 1996; Ohl et al., 2001). Therefore, to identify spatial patterns carrying information about the stimulus classes CS_{go} and CS_{no-go}, we obtained instantaneous spatial amplitude distributions of oscillations in the β- and γ-range (15–80 Hz), and analyzed them trial-by-trial as a function of time using a multivariate classification procedure. For this, we determined for each of the 18 channels the signal power in 180 ms time frames that were stepped through the 5.0 s single-trial recordings in 20 ms steps. Signal power was quantified by the root mean squared (RMS) amplitude within in a 180 ms time frame after digitally bandpass-filtering the ongoing ECoGs (FIR-filter of order 200) in the β- and γ-range (15–80 Hz). Notably, RMS- and FFT-based estimation of signal power in the β- and γ-range actually have yielded similar results in classification studies comparable to our study (cf. Barrie et al., 1996). However, the RMS method is

computationally much more effective than FFT-based methods. For removing baseline offsets, mean RMS values calculated from a 1.0 s prestimulus baseline were subtracted channel-wise, and trial by trial. Furthermore, a frame-wise z-standardization across the 18 recording channels (subtraction of mean, and division by the SD) removed activity common to all channels, and pronounced the spatial variation of power. Parameters of this procedure like window lengths and filter bands had been optimized for yielding best classification results in a previous study (Barrie et al., 1996). A detailed description of the classification procedure can be found in the studies by Barrie et al. (1996) and Ohl et al. (2003a). In short, sets of all artifact-free CS_{go} trials and all artifact-free CS_{no-go} trials of a session were each split in half by even and uneven set indices. This gave four disjunctive trial subsets: CS_{go-A}, CS_{go-B}, $CS_{no-go-A}$, and $CS_{no-go-B}$. For each subset, a centroid of the spatial distribution of normalized RMS values was calculated within each time frame across the trials of the subset. Based on Euclidean distances, the question was asked whether an instantaneous spatial RMS distribution in a trial of subset CS_{go-A} was closer to the centroid of CS_{go-B} or of $CS_{no-go-B}$. In the former case, the RMS distribution was said to classify correctly; in the latter case it was said to classify incorrectly. In the same manner, the RMS distributions from set $CS_{no-go-A}$ were classified against the centroids of CS_{go-B} and $CS_{no-go-B}$. Analogously, the RMS distributions in of CS_{go-B} and $CS_{no-go-B}$ were classified against the centroids of CS_{go-A} and $CS_{no-go-A}$, respectively (cross classification). The number of correctly classified trials was summed up over trials and subsets, separately for each time frame. Goodness of pattern classification was quantified in percentage of correctly classified trials in each time frame. Significance of pattern classification was tested against the null hypothesis of randomly sorting the instantaneous RMS distributions into the stimulus classes. The probability $p_n(i)$ for obtaining at least the observed number of correctly classified trials $N(i)$ from a total number of trials n_0 by chance ($p = 0.5$) was determined for each time frame i using the cumulative binomial distribution ($N(i)$, number of correctly classified trials at time step i; n_0, total number of classified trials; n_w: number of 20 ms time steps; n_s, number of classified trials in a subset), as follows:

$$p_n(i) = \sum_{n=N(i)}^{n_0} \binom{n_0}{n} p^n (1-p)^{n_0-n} \; \forall \; i \in \{1, \ldots, n_w\}; \; n_0 = \sum_{s=1}^{4} n_s. \quad (1)$$

In a second statistical analysis, we investigated how the information discriminating between the stimulus classes was spatially distributed over the recording array. For this discriminant analysis, we determined how recording sites contributed to pattern classification in the last session of training (session 7), when all animals had reached high levels of discrimination performance, and significant pattern classification (see Results). For each of the possible 262,142 combinations of the 18 recording channels the goodness of classification was determined in selected time frames, i.e., in an early time frame coinciding with the first peak in the EEP, and in late time frames yielding maximum goodness of classification in the poststimulus/preresponse interval (between 900 ms after stimulus onset and the behavioral response time quantified by the mean hit reaction time in each session). To obtain a measure for the relative contribution of a channel to pattern classification, the mean of goodness of classification across all possible sets of channels containing this channel ($n = 14,563$) was calculated, and normalized to 0 and 1 between its minimum and its maximum across recording channels. This measure was used to quantify the discriminant pattern information contributed to classification independently by each single recording channel. To compare the distribution of discriminant pattern information across the recording array with the lateral spread of ICMS-evoked activity (see Results, Electrically evoked potentials), we again calculated distance functions relative to the stimulation sites. For this we determined the minimal distance of each recording channel to the two stimulation sites. Minimal channel-distances were binned in a similar way as in the analysis of the lateral spread of the first peak in the EEP. The relative contribution of recording sites to classification was then averaged across the channels of each bin, separately for early and late time frames. By this we obtained for each animal, a distance function of discriminant pattern information for early and late time frames, respectively.
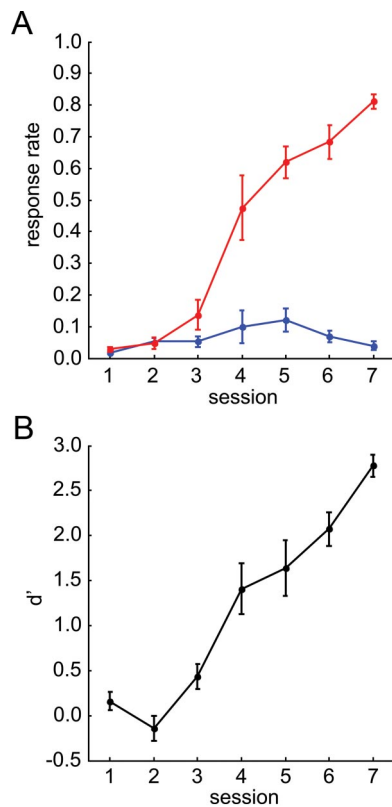
Figure 3. Discrimination learning. **A**, Learning curves. Mean and SEs of the hit (red) and the false-alarm (blue) rate across animals is plotted over training sessions. **B**, Discrimination performance. Mean and SE of the sensitivity measure d' across animals is plotted over training sessions.

Correlation analysis. Common correlation coefficients between goodness of classification and discrimination performance, d', over sessions were calculated using the Fisher z-transform corrected for small sample size (Hotelling, 1953). First, single correlation coefficients were calculated for each animal, and transformed to Fisher z-values. Using a χ^2-statistics with 5 degrees of freedom (Hotelling, 1953), Fisher's z-values were tested for homogeneity across animals. If the null hypothesis of homogeneity could not be rejected on a significance level of $p < 0.01$, Fisher's z-values were averaged across animals, and transformed back to yield a common correlation coefficients across animals. Limits of significant positive and negative correlation were calculated on a significance level of α. For this the SD for the Fisher's z-values was estimated (Hotelling, 1953) and the range including $1-\alpha\%$ of the variance around zero was calculated. The upper and lower limit of this range was transformed back into limits of significant correlation.

Results

Discrimination behavior

Using a go/(no-go) active avoidance paradigm (Fig. 2*B*), gerbils were trained for 7 daily sessions in a two-compartment shuttle box (see Materials and Methods) to discriminate two nearby ICMS sites within the right cortical field AI. From the learning curves (Fig. 3*A*) it can be seen that from session 3 on mean hit rate steeply increased over sessions finally reaching values of ~0.8 in the last session, whereas the mean false-alarm rate always stayed below 0.2. Thus, on average, animals started to show significant discrimination performance from session 3 on. A biphasic rise-fall time course of the false-alarm rate could be observed between session 3 and 6 (Fig. 3*A*) indicating a certain amount of generalization between CS_{go} and $CS_{no\text{-}go}$ during early training sessions. Huynh-Feldt-corrected two-way repeated-measures ANOVA of

the conditioned-response rates yielded significant main effects for the factors "stimulus condition" (CS_{go} and $CS_{no\text{-}go}$, $F_{(1,5)} = 156.274$, $p < 0.001$), and "session" ($F_{(3.344,16.719)} = 28.621$, $p < 0.001$), and a significant interaction between these 2 factors ($F_{(4.581,22.905)} = 47.579$, $p < 0.001$). The increase of the difference between hit and false-alarm rates over sessions was therefore highly significant. Discrimination performance quantified by the mean of the sensitivity measure d' across animals (Fig. 3*B*) showed a significant monotonic increase of values around zero (chance level) to values larger than 1.3 (~75% correct) was found over sessions (one-way repeated measure ANOVA with Huynh-Feldt correction: $F_{(6,30)} = 40.168$, $p < 0.001$). High discrimination performance ($d' > 1.3$) was reached within ~5 sessions.

Electrically evoked potentials

To characterize the electrically evoked input to AI, electrically evoked potentials (EEPs) were calculated from the ECoG-recordings during training in response to CS_{go} and $CS_{no\text{-}go}$, respectively (see Materials and Methods). Thereby, our two main objectives were to determine the stimulation sites within the coordinates of the recording array, and to quantify the lateral spatial spread of cortical activation evoked by ICMS. Before all analyses, electrical stimulus artifacts in the ECoG had been largely removed in each trial by an interpolation procedure (Fig. 4*A*) (see Materials and Methods). A first prominent early negative peak 20 ms after stimulus onset (N20) was reliably found in the EEPs of all sessions and stimulus conditions (Fig. 4*A*). Presumably, this peak reflects early, transsynaptic excitation of the auditory cortex by ICMS (see Discussion). Amplitudes and latencies of the N20 peak at its maximum were averaged across sessions in each animal (see Materials and Methods). No significant differences were found in the grand mean of this N20 amplitude and N20 latency between CS_{go} (amplitude mean ± SE: -335 ± 94 μV; latency mean ± SE: 22 ± 2 ms) and $CS_{no\text{-}go}$ (amplitude mean ± SE: -385 ± 95 μV; latency mean ± SE: 20 ± 1 ms) across animals (pairwise Wilcoxon sign-rank tests, $n = 6$, $p > 0.1$).

To characterize the lateral spatial spread within auditory cortex evoked by ICMS, spatial distributions of EEP amplitudes across the recording array were determined at the N20 latency for each stimulus condition, session and animal. Spatial amplitude distributions where focal, i.e., they displayed a single spatial peak (Fig. 4*B*). In 5 of the 6 animals these spatial N20 peaks were centered within the array. Only in one animal the spatial maxima, both for CS_{go} and $CS_{no\text{-}go}$, were located at the ventral edge of the array. Stimulation sites seemed to be located ventral to the recording array in this animal. Notably, in all animals, spatial distributions of the EEP amplitude at the N20 latency displayed a constant offset across channels (~40% of the N20 amplitude), which was estimated by the mean of the three lowest EEP amplitude values at the N20 latency within the recording array. In the session-mean, this offset did not differ between stimulus conditions across animals (pairwise Wilcoxon sign-rank test, $n = 6$, $p > 0.1$).

Although the cortical regions activated by caudal and rostral ICMS had a large overlap, the positions of the spatial maxima of the N20 peak were systematically located at more caudal and rostral positions in the tonotopic map, respectively. In accordance with a distance of ~0.7 mm between the implanted stimulation electrodes, the mean distance between N20 peak maxima averaged across sessions and then across animals was 0.74 ± 0.13 mm (mean ± SE), and did not differ significantly from 0.7 mm (pairwise Wilcoxon sign-rank test, $n = 6$, $p > 0.1$). The maxima of the spatial distributions of EEP amplitudes at the N20 latency
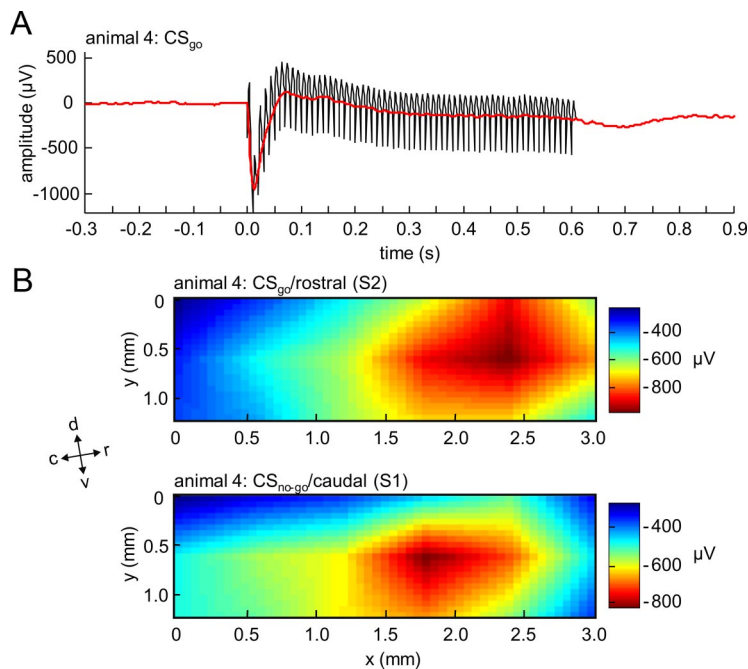
Figure 4. Time course and spatial distribution of early ICMS-evoked activation. **A**, Typical example of an EEP from a single animal averaged across CS_{go} trials in a single session of training. The EEP is shown before (black) and after (red) removal of single pulse stimulus artifacts in the ECoG-recording (see Materials and Methods). A first, prominent negative peak in the EEP can be seen at a latency of ~20 ms (N20, red arrow). **B**, Typical example of spatial distributions of the N20 amplitude in response to a CS_{go} at the rostral stimulation electrode (top) and to a CS_{no-go} at the caudal stimulation electrode (bottom). Anatomical directions relative to the recording array are indicated by arrows (d, dorsal; c, caudal; l, lateral; m, medial; r, rostral; v, ventral).

were therefore be used to precisely define the positions of the stimulation sites relative to the coordinates of the recording array (see Materials and Methods).

We then determined for each stimulus condition the lateral spread of ICMS-evoked cortical activation by analyzing the EEP amplitude at the N20 latency as a function of distance to the stimulation sites (Fig. 5A,B). Normalized and binned distance functions were averaged across sessions, and then across animals for CS_{go} (Fig. 5A) and CS_{no-go} (Fig. 5B), respectively. To quantify the lateral spatial spread, the fall-off of the grand mean of the normalized EEP amplitude with distance from the stimulation site was fitted to a Gaussian (see Materials and Methods). The range of the lateral spatial spread defined as the half-width of the Gaussians (Fig. 5A,B) was highly similar for the CS_{go} (1.25 mm, 99% confidence interval from 1.09 to 1.41 mm) and the CS_{no-go} (1.22 mm, 99% confidence interval from 0.98 to 1.47 mm). Thus, ICMS imposed transsynaptic activity onto the cortex up to a distance of at least 1.2 mm lateral to the stimulation sites, well beyond the range of significant effects by volume conduction (see Discussion).

Classification of cortical activity patterns

We then studied the interaction between ICMS and the learning-dependent ongoing cortical dynamics. After filtering the ECoGs in the β- and γ-band (15–80 Hz), spatial distributions of activity were obtained by determining the root mean squared (RMS) amplitude for each of the 18 channels in 180 ms time frames stepped through the 5.0 s single-trial recordings in 20 ms steps. To identify spatial patterns carrying information about the stimulus classes CS_{go} and CS_{no-go}, we applied a multivariate classification procedure to the spatial distributions of normalized RMS values, time frame by time frame (see Materials and Methods). From this pattern classification analysis we obtained for each

animal, session and time frame the goodness of pattern classification (percentage correct classification) and its significance quantified by the p value for obtaining the observed number of correctly classified trials in a session just by chance.

Figure 6 shows a representative example of this pattern analysis from a single animal. Results are displayed for a session before the steep increase of discrimination performance (Fig. 6A, session 3; see also Fig. 3), and for the last session of training, when all animals had obtained high levels of discrimination performance (Fig. 6B, session 7). The decadic logarithm of the p value of pattern classification is plotted as a function of time (from −2 to 3 s relative to stimulus onset), and each point includes filtered ECoG data of a 180 ms frame of classification (from −90 to 90 ms relative to that point in time; cf. Barrie et al., 1996). The stimulus interval was defined as the time interval affected by ICMS (from −290 to 900 ms relative to stimulus onset) (Fig. 6A,B) including the 610 ms pulse train, the half lengths of the classification frame (90 ms) and the length of the applied filter kernel (200 ms). Highly significant (defined as $\log_{10}(p_n) < -3$) early peaks of pattern classification were found around stimulus onset (Fig. 6A,B, blue arrows), and lasted for ~11 consecutive classification frames from about −90 ± 90 ms to 110 ± 90 ms. Thus, the early peak of pattern classification was temporally overlapping with the N20 peak of the EEP (Fig. 4A). Moreover, these early peaks were time-locked to stimulus onset, and occurred in almost all of the training sessions, independent of the learning state of the animals. This suggests, that they reflect the discernible focal activation patterns of the N20 peak in the EEP in response to CS_{go} and CS_{no-go}, respectively (see Electrically evoked potentials).

In the course of training, further late peaks of significant pattern classification were found in the poststimulus time interval (Fig. 6B, red arrow). Similar to what has been observed in gerbils learning to discriminate acoustic stimuli (Ohl et al., 2001), these late significant peaks of pattern classification were more likely to occur with higher levels of behavioral performance, in our case after the steep increase of discrimination performance between session 3 and 5. Thus, before session 4, late significant peaks rarely occurred, whereas such peaks were found in each of the animals in session 7 (Fig. 6B), i.e., when all animals had reached high levels of discrimination performance. Notably, significant pattern classification almost never occurred during prestimulus time.

To evaluate the behavioral relevance and learning dependence of the observed classifiability of cortical activity patterns, we studied the correlation between goodness of classification and behavioral discrimination performance over training sessions as a function of time within the trial (Fig. 7A,B). In consecutive 0.5 s time windows, maximum goodness of classification values were obtained for each session and animal. Common correlation coefficients (see Materials and Methods) were calculated for each time window between maximum goodness of classification values (z-standardized across sessions) and discrimination perfor-
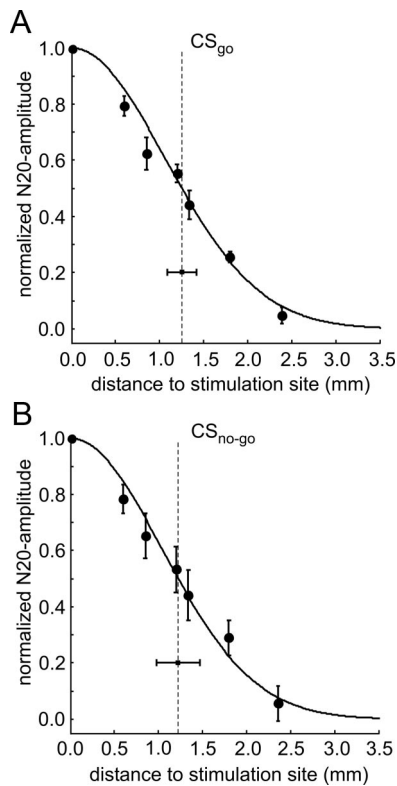
Figure 5. Lateral spatial spread of ICMS-evoked activation. ***A***, ***B***, The normalized and binned EEP amplitude at the N20 latency is displayed as a function of distance to the stimulation site in response to CS$_{go}$ (***A***) and CS$_{no\text{-}go}$ (***B***). Up to 1.3 mm from the stimulation site, normalized EEP amplitudes were binned and averaged across channels of equal distance (see Materials and Methods). As the position of stimulation sites varied relative to the recording array, broader bins had to be used at larger distances, to obtain amplitude values for all stimulus conditions and animals (the penultimate bin included N20 amplitudes at 1.7, 1.8, and 1.9 mm and the last bin N20 amplitudes between 2.1 and 3.2 mm). Grand mean and SE of the distance functions of EEP amplitudes across animals are displayed by circles and error bars. The lateral spread of early cortical activation evoked by CS$_{go}$ and CS$_{no\text{-}go}$ was estimated by the half-width of a Gaussian fitted to each of the two grand mean distance functions, respectively (indicated by the vertical dashed lines) (see also Materials and Methods). Confidence intervals of the half-widths are displayed by horizontal bars.
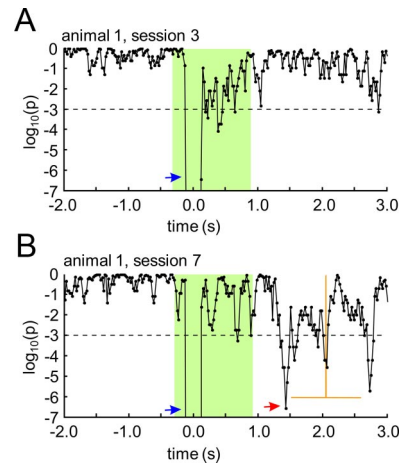

Figure 6. Spatial patterns in the β- and γ-bands of the ongoing ECoG. Significance of ECoG pattern classification (decadic logarithm of the *p* value) as a function of time relative to stimulus onset (time is given as the center of each 180 ms classification time frame stepped through the ECoG in 20 ms steps). ***A***, ***B***, A typical example is shown for sessions from the beginning (***A***) and the end (***B***) of training. Mean and SD of hit reaction times in the last session are displayed by vertical and horizontal orange lines, respectively. The time range affected by the electrical stimulus, including the half-length of the classification frame (90 ms) and the length of the filter kernel (200 ms), is marked by a green area. Arrows indicate highly significant peaks of pattern classification (*p* < 0.001, red dashed horizontal line) in the stimulus interval reflecting early, ICMS-evoked patterns (blue arrows), and late, learning-dependent patterns in the poststimulus/preresponse interval (red arrow).

mance quantified by *d*′ over sessions. A highly significant peak of positive correlation (Fig. 7*A*, *p* < 0.01, *r* = 0.68, asterisks) was found in the poststimulus time interval between 1.0 and 1.5 s after stimulus onset preceding the behavioral response characterized by the grand mean and the SD of the hit reaction times across animals and sessions (*n* = 23). Correlation was also significant (*p* < 0.01, *r* = 0.43) in the late stimulus interval between 0.5 and 1.0 s after stimulus onset. No significant correlation was found in the early stimulus interval between 0.0 and 0.5 s after stimulus onset, and in the prestimulus intervals. From this it can be inferred, that late after stimulus onset, in the poststimulus/preresponse interval, spatial patterns of cortical activity develop, when animals learned to discriminate the stimulation sites.

To further characterize the temporal occurrence of late patterns within the trial, the latency of the most significant time frame of classification in the poststimulus/preresponse interval (between 0.9 s after stimulus onset and the mean hit-reaction time in a session) was determined for each session with significant pattern classification. In the grand mean, most significant late peaks of pattern classification occurred at 1.36 ± 0.32 s, and significantly 0.69 ± 0.3 s before the reaction time at 2.1 ± 0.43 s after stimulus onset (mean ± SD, paired-wise t tests, *p* < 0.01,

n = 23, data pooled across sessions with significant patterns classification and animals, Fig. 8*A*,*B*). As indicated by the SDs, latencies of late peaks of pattern classification largely varied across sessions and animals, both, relative to stimulus onset and behavioral response. Thus, in contrast to the early peaks, late peaks were not precisely time-locked, neither to stimulus onset, nor to the behavioral response. Furthermore, there was no stimulus evoked activity, which could be related to ECoG pattern classification in the poststimulus interval, neither in the unfiltered EEP (Fig. 4*A*), nor in the filtered β- and γ-band EEP (data not shown). Obviously, significant poststimulus classification relied on discernible spatial patterns emerging from the ongoing cortical β- and γ-band activity, when animals learned to make use of the site of ICMS for solving the discrimination task. Maximum goodness of classification in the poststimulus/preresponse interval reached significance with classification performance levels already between 70% and 80%. This is consistent with the finding that late patterns occurred at variable latencies, and thus not always within the same time-frames across trials.

At this point, several open questions remain referring to the information used by the classifier. Since classification was always performed in reference to a CS$_{go}$ and a CS$_{no\text{-}go}$ centroid, it must be based on differences in the spatial activity distributions between the go and no-go conditions, and thus should in principle rely on information about the conditioned stimuli. However, for the late patterns, this information might be related only indirectly to the preceding conditioned stimuli, e.g., by reflecting the differential motor responses associated with the go and no-go conditions in the course of learning. By this, the correlation between pattern classification and discrimination performance in the late poststimulus/preresponse interval might be spurious, and rather related to the conditioned go responses occurring with increasing frequency over sessions, than to the learning process itself. As late peaks of pattern classification occurred on average 0.69 s before the actual go responses, and also were not precisely time-locked
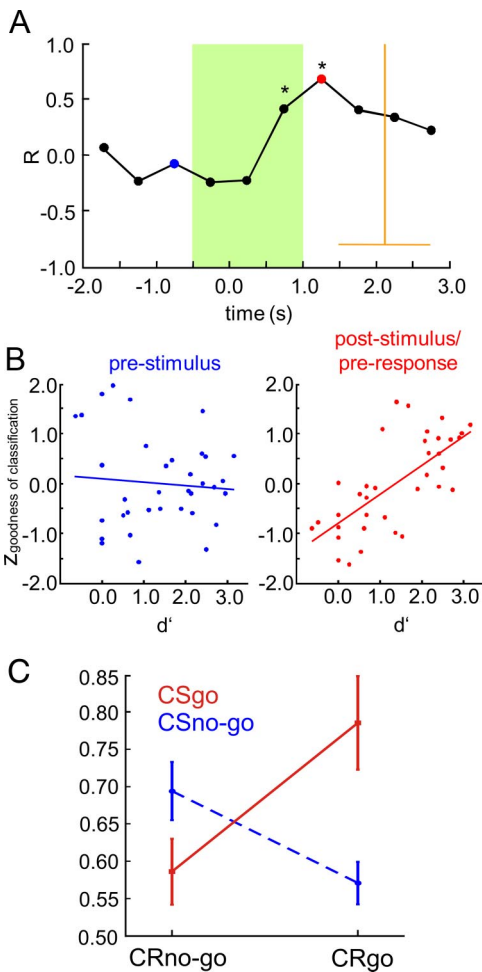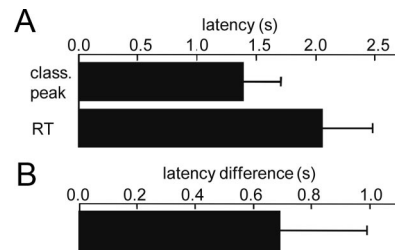
Figure 8. Timing of late spatial patterns relative to stimulus onset and behavioral response. **A**, Mean and SD of hit reaction time (top bar) and the latency of the most significant peak of pattern classification in the poststimulus/preresponse interval relative to stimulus onset (bottom bar). Data were pooled across animals and sessions with significant pattern classification. **B**, Mean and SD of the latency of the most significant peak of pattern classification as in **A**, but relative to the behavioral response in each session (mean of hit reaction times). Data were pooled as in **A**.

Figure 7. Late spatial patterns in the ongoing ECoG and their relation to discrimination learning. **A**, Correlation between behavioral d' and goodness of classification over sessions is plotted as a function of time relative to stimulus onset. Common correlation coefficients across animals were calculated over sessions between d' values and maximum goodness of classification values selected from consecutive 500 ms time windows relative to stimulus onset in each session (z-standardized across sessions for each time window). Highly significant correlation ($p < 0.01$) is marked by asterisks. Statistical testing showed that homogeneity of correlation coefficients across animals held for all time windows (see Materials and Methods). Grand mean and SD of hit reaction times in the last session (pooled across animal) are displayed by vertical and horizontal orange lines, respectively. The time range affected by the electrical stimulus is marked by the green area. **B**, Detailed plots of the correlation for a prestimulus interval (-1.5 to -0.5 s, blue) and for the poststimulus/preresponse interval (1.0 –1.5 s, red). **C**, Interaction plot showing the effects of factors CR (CR_{go} and CR_{no-go}), and CS (CS_{go} and CS_{no-go}) on correct pattern classification. Percentage correct classification was calculated separately for subsets of hit (CR_{go},CS_{go}), miss (CR_{no-go},CS_{go}), false-alarm (CR_{go},CS_{no-go}), and correct-rejection (CR_{no-go},CS_{no-go}) trials, at the latency of the maximum goodness of classification in the poststimulus/preresponse interval of the sessions containing the trials (behavioral response time was quantified by the mean hit reaction time in the session). Subsets consisted of the last n hit, miss, false-alarm, and correct-rejection trials collected across sessions 3–7 in each animal, respectively. To avoid biases due to different set sizes, n was matched to the size of the smallest set in each animal ($n = 18 \pm 9$, mean $\pm$ SD), which always was the false-alarm set. Percentage correct values are plotted as a function of the factor CR, separately for the CS_{go} (red line) and the CS_{no-go} (blue dashed line). Higher percentage correct values were found for hit (mean $\pm$ SE, 0.79 $\pm$ 6%) and correct-rejection (69 $\pm$ 4%) trials compared with false-alarm (57 $\pm$ 3%) and miss (59 $\pm$ 3%) trials. A two-way ANOVA (Huynh-Feldt corrected) with factors CR (CR_{go} and CR_{no-go}) and CS (CS_{go} and CS_{no-go}) was then applied to the percentage correct values across animals based on these trial subsets. A significant interaction was found between the factors CR and CS ($F_{(1,5)} = 7.417$, $p < 0.05$), but no significant main effects, either for the factor CR ($F_{(1,5)} = 1.553$, $p \geq 0.05$) or the factor CS ($F_{(1,5)} = 2.259$, $p \geq 0.05$).

to it, it seems very unlikely that our results were directly related to movements or even movement artifacts occurring with the go responses. Still, our classification might reflect brain signals related to the preparation of the go response, independent from learning. Alternatively, pattern classification might not be based on different patterns discriminating between the two conditioned stimuli, but only on a pattern related to the mere detection of one of the two conditioned stimuli, which also could lead to spurious correlation.

Therefore, we tried to further elucidated the nature of the information used by the classifier by investigating how late pattern classification was influenced by the factors conditioned response (CR) and CS. We determined at the latency of the maximum goodness of classification in the poststimulus/preresponse interval for each trial of a session, whether it was classified correctly. We then selected the last n trials from the sets of hit-, miss-, false-alarm-, and correct-rejection trials across sessions 3–7 in each animal, respectively, and obtained the percentage of correctly classified trials in each of these subsets. To avoid biases due to different set sizes, n was matched to the size of the smallest set in each animal ($n = 18 \pm 4$, mean $\pm$ SE), which always was the false-alarm set. A two-way ANOVA-analysis (Huynh-Feldt corrected) with factors CR (CR_{go} and CR_{no-go}) and CS (CS_{go} and CS_{no-go}) was then applied to the percentage correct values derived in each animal from the subsets of hit (CR_{go},CS_{go}), miss (CR_{no-go},CS_{go}), false-alarm (CR_{go},CS_{no-go}), and correct-rejection (CR_{no-go},CS_{no-go}) trials. We found a significant interaction between the factors CR and CS ($F_{(1,5)} = 7.417$, $p < 0.05$), but no significant main effects, neither for the factor CR ($F_{(1,5)} = 1.553$, p $>$ = 0.05), nor the factor CS ($F_{(1,5)} = 2.259$, p $>$ = 0.05). Thus, correct classification did not depend on the presented stimulus or the type of response made by the animals. From the interaction plot (Fig. 7C) it can be seen that classification performance was better for the trial sets in which animals responded correctly (hits and correct responses), and closer to chance level (50%) in the other trial sets (misses and false-alarms). Obviously, late patterns predominantly occurred in trials in which animals correctly interpreted ICMS (hit and correct-rejection trials). This further suggests that late patterns carried information about the preceding conditioned stimuli, and were associated with a subsequent correct response by the animal. Thus, late patterns apparently reflected the meaning attributed to the presented electrical stimuli by the animal and therefore were closely related to the process of discrimination learning.
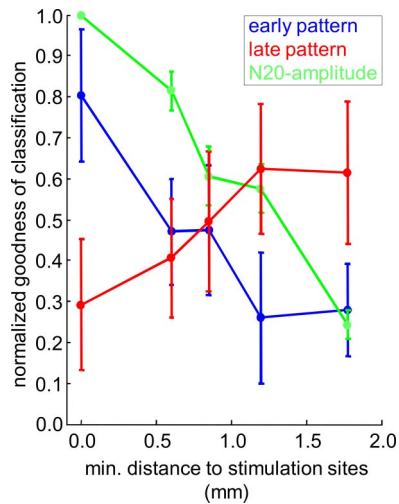
Figure 9. Spatial organization of early and late cortical activity patterns. Distance functions of discriminant pattern information are shown for early (blue) and late (red) patterns. Relative contribution of recording sites to pattern classification (see Materials and Methods) is displayed as a function of their minimal distance to the stimulation sites. Up to distances of 1.2 mm, relative contributions to classification were binned and averaged across channels with the same minimal distance. Larger distances from 1.3 to 3.2 mm were pooled in a single bin before averaging. Shown is the mean of the resulting distance functions across animals interval in the last session of training, for early patterns (blue), i.e., in the time frame of classification temporally overlapping the N20 peak in the EEP, and for the late patterns (red), i.e., in the time frame yielding most significant pattern classification in the poststimulus/preresponse. The green curve shows for comparison the mean and the SE of the distance function of EEP amplitudes at the latency of the N20 peak maximum across animals in the last session of training. Normalized EEP amplitudes were binned in the same way as for the distance function of discriminant information. In each animal, the resulting distance functions of the EEP were averaged across stimulus conditions.

Spatial organization of cortical activity patterns

How does the information carried by early and late cortical activity patterns correspond to the lateral spatial spread of cortical activation evoked by ICMS? To shed light on this question, we analyzed how the information discriminating between the stimulus classes was spatially distributed over the recording array. For this we determined how recording sites contributed to pattern classification in early and late patterns, respectively (see Materials and Methods). This discriminant analysis was performed in the last session of training (session 7), when all animals had reached high levels of discrimination performance (see Materials and Methods). For the early patterns, the relative contribution of recording sites to pattern classification was calculated in the time frame coinciding with the N20 peak of the EEP. For the late patterns, the relative contribution of recording sites was calculated in the time frames yielding maximum of goodness of classification in the poststimulus/preresponse interval (behavioral response time quantified by the mean hit reaction time in the session). The distance functions of discriminant pattern information and EEP amplitude in Figure 9 show how recording sites contributed to classification as a function of their distance to the stimulation sites in dependence on the early EEP amplitude. For the early pattern, recording sites proximal to the stimulation site contributed relatively more information to classification than distal sites. At distances larger than 1.2 mm (Fig. 9, blue curve) the relative contribution to classification reached a lower plateau. In the late patterns, however, recording sites near the stimulation site contributed less information compared with distal recording sites. At distances larger than 1.2 mm, the relative contribution to classification approached an upper plateau (Fig. 9, red curve). For

comparison, an average distance function of the EEP amplitude at the N20 latency across stimulus conditions is shown, which was binned in the same way as the distance functions of discriminant information (Fig. 9, green curve). Repeated measure two-way ANOVA (Huynh-Feldt-corrected) with factors "distance" and "pattern" (early vs late) yielded a weakly significant interaction ($F_{(4,20)} = 2.522$, $p < 0.1$), and no significant main effects. This indicates that in early patterns, recording channels within the spatial spread of the evoked N20 peak (~1.2 mm), where early EEP amplitudes (Fig. 9, green curve) were high, contributed more information than more distal recording sites, where early EEP amplitudes were low. In the late patterns, however, within the spatial spread of ICMS-evoked activation, pattern information was reduced. Surprisingly, more discriminant information was contributed by the cortical sites distal to the stimulation site.

To evaluate the relevance of this finding on the level of individual animals, pattern classification was recalculated with subsets of recording channels, proximal and distal to the stimulation sites. In a first step, the range of the lateral spread of ICMS-evoked activation was quantified for each animal, by estimating the half-widths of a Gaussian function fitted to the distance function of the normalized early EEP amplitude at the N20 latency averaged across session (nonlinear-least-square fitting) (see Electrically evoked potentials). The mean of the estimated half-widths across animals (CS_{go} mean ± SE: 1.25 ± 0.07 mm; CS_{no-go} mean ± SE: 1.28 ± 0.16 mm) lay within the 99% confidence interval of the half-widths estimated from the grand mean of the distance function of the EEP amplitude across animals. Half-widths were averaged across stimulus conditions to quantify the lateral spread of early ICMS-evoked activity in each animal. The (minimal) channel-distance to the stimulation sites closest to the lateral spread ($d\min_{ch}$) was used to define distal and proximal sets of recording sites. The distal set of recording sites included the n_{ch} channels that were at least $d\min_{ch}$ away from the stimulation site (Fig. 10A). The set of proximal recording channels consisted of the n_{ch} channels closest to the stimulation site. By this, distal and proximal channel sets were matched in size, to exclude effects solely based on differences in set size, and not on differences in the distance to the stimulation site. As shown in the example of Figure 6A, significance of pattern classification in the poststimulus interval was largely reduced in the proximal subset of recording channels (Fig. 10A, red arrow). In 5 of the 6 animals, significant pattern classification in the poststimulus/preresponse interval was only found for the distal, but not for the proximal channel-subset (Fig. 10B, right). In animal 3, however, significant pattern classification occurred in both subsets. Interestingly, the spatial maximum of the N20 peak in this animal was located at the ventral edge of the recording array (see Electrically evoked potentials). Maybe, due to a ventral offset, the actual stimulation sites were at a larger distance to the recording sites as determined from the N20 peak maxima, and thus well outside the range of lateral spread of cortical activation. Early pattern classification was always significant with both, distal and proximal subsets (Fig. 10B, left). Only in two animals, it was reduced with the distal compared with the proximal subset. Thus, in the early pattern, proximal as well as distal cortical sites provided sufficient information to discern the patterns. Finally, we analyzed whether pattern information carried by proximal and distal subsets was learning dependent. As with the full set of recording channels (Fig. 7A), correlation analysis between the goodness of classification and d' yielded a highly significant peak of correlation for the distal channel-subset in the poststimulus/preresponse time interval between 1.0 and 1.5 s (Fig. 10C, $p < 0.01$, $r = 0.59$, red curve
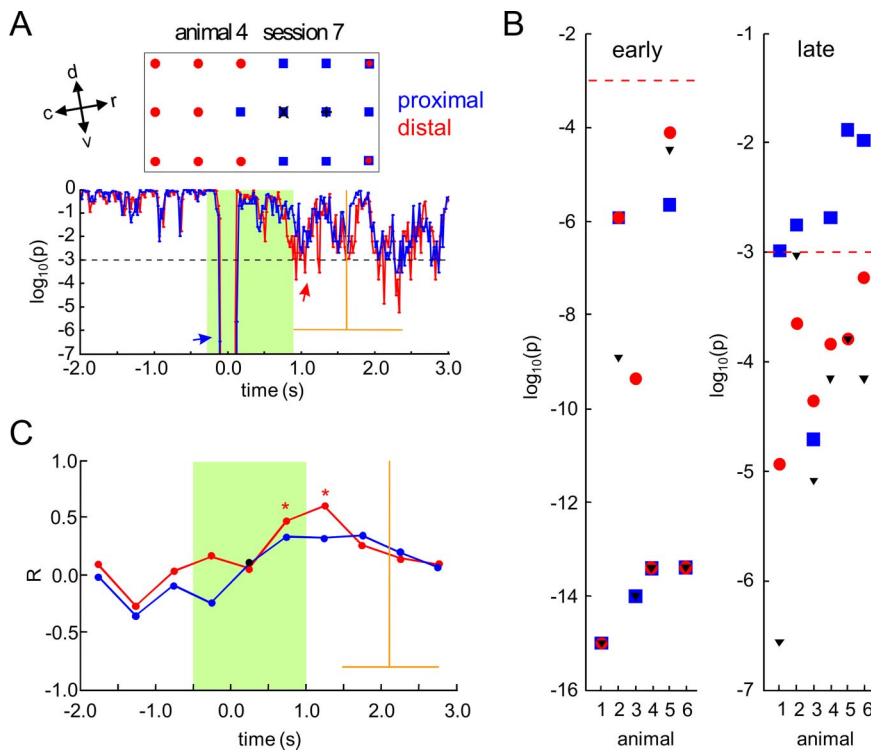
Figure 10. Pattern classification with subsets of recording sites distal and proximal to the stimulation sites. *A*, Typical results of pattern classification with distal and proximal sets of recording channels from a single animal in the last session of training. In the top box, the selected distal and proximal recording channels are marked by red circles and blue squares, respectively. Stimulation sites determined within the coordinates of the recording array by the maxima of the spatial distribution of EEP amplitudes (Fig. 3*B*) are indicated by a + (CS$_{go}$ at the rostral stimulation electrode) and by an x (CS$_{no-go}$ at the caudal stimulation electrode). Distal and proximal sets of channels were defined in each animal on the basis of the lateral spread of ICMS derived from the spatial analysis of the N20 peak in the EEP (see Results). Both sets matched in size. In the bottom, significance of ECoG pattern classification (decadic logarithm of the p value) is shown as a function of time relative to stimulus onset (Fig. 4*A*), separately for the distal (red) and the proximal (blue) set of recording channels. Mean and SD of hit reaction times are indicated by orange lines. *B*, p values obtained with distal (red circles) and proximal subsets (blue squares), and with the full set (black triangles) of recording channels in a time frame of classification temporally coinciding with the N20 component in EEP (left, early patterns), and in the time frame yielding most significant pattern classification in the poststimulus/preresponse interval (right, late pattern). Results are shown for all animals in the last session of training. A predefined level of significance ($p < 0.001$) is indicated by red dashed horizontal lines. *C*, Correlation between behavioral d′ and goodness of classification with distal (red) and proximal (blue) subsets of recording channels as a function of time relative to stimulus onset (Fig. 4*B*). Highly significant common correlation coefficients ($p < 0.01$) are marked by asterisks. Significant deviation from homogeneity across animals was found only for the correlation coefficients in the time interval between 0 and 500 ms after stimulus onset, with the proximal subset of recording channels (black circle).

and asterisks). For the proximal subset, no such highly significant correlation could be found (Fig. 10*C*, blue curve). In summary, these findings suggest that late, learning-dependent pattern information was carried mainly by cortical sites distal to the stimulation sites. Formation of late patterns was suppressed within the range of the lateral spread of activation evoked by ICMS at stimulus onset. Obviously ICMS interfered with the ongoing cortical dynamics from which late spatial activity patterns emerge with learning.

Discussion

Spatial spread and transsynaptic activation by ICMS
In our study, ICMS led to an early, focal, activation of auditory cortex at the stimulation sites ~20 ms after stimulus onset. With the applied current amplitudes (between 50 μA and 70 μA), the radius of direct neural excitation around the stimulation site due to passive spread of current (effective current spread) (Tehovnik, 1996) was expected to range between 0.120 mm for high-threshold neurons (K = 3460 μA/mm^2) and 0.507 mm for low-threshold neurons (K = 272 μA/mm^2), as heuristically

determined from current distance constants K reported for pyramidal neurons (Stoney et al., 1968). However, the range of the lateral spatial spread of the first EEP peak (N20) was ~1.2 mm, and thus largely exceeded the effective current spread. This discrepancy can be explained by the fact that the EEP is generated by synaptic currents, and thus relies on more widespread, transsynaptic effects of direct excitation (Tehovnik et al., 2006). By using the surface EEP, we further might have overestimated the actual spread of transsynaptic activation due to a spatial blur by volume conduction. A considerable over-estimation seems however unlikely, since a highly similar range of lateral spread has been reported previously by Butovas and Schwarz (2003), who analyzed spiking activity of cortical units in rat somatosensory cortex in response to single-pulse ICMS. These authors demonstrated excitatory effects up to a distance of 1.350 mm. Thus, the focal N20 peak in our study most likely reflects an early, excitatory, transsynaptic activation of the cortex by ICMS. If ICMS directly excites mixed populations of bypassing excitatory and inhibitory axons (Nowak and Bullier, 1998a,b; Butovas and Schwarz, 2003), its functional spread would rather be determined by the length of horizontal fibers which is on the order of 1 to 3 mm in gerbil neocortex (Budinger et al., 2000), and by the spatiotemporal interaction of ICMS-evoked excitatory and inhibitory activity, than by the effective current spread alone. This would be consistent with the observed large range of the lateral spread of ~1.2 mm.

The observed offset of the N20 amplitude could be due to the activation of yet undetermined cortical elements that are spatially distributed across large parts of the recording area. Using monkeys, Tolias et al. (2005) reported in an fMRI study that the functional spread in response to ICMS was ~2.6 mm in radius (cf. Tolias et al., 2005, their Fig. 6*A*; 4 s pulse trains, 100 μA amplitude, other parameters identical to our study). Also, in their unit-recordings Butovas and coworkers (Butovas and Schwarz, 2003; Butovas et al., 2006) found inhibitory effects up to distances of 1.8 mm. Alternatively, deep thalamic sources activated by the direct excitation of efferent corticothalamic axons through ICMS (DeYoe et al., 2005) might have led to a widespread offset of surface potentials due to volume conduction.

Behavioral relevance of cortical dynamics in response to ICMS
In our study, stimulation sites were spaced ~0.7 mm, close to reported distance thresholds for ICMS (Schmidt et al., 1996; Otto et al., 2005). Therefore, with a spatial spread of ~1.2 mm, cortical activations at the two stimulation sites were largely overlapping in space (~57% overlap), although the amount of current used

for directly evoking percepts in our study (50 –70 μA) was in a moderate range (Rousche et al., 2003). Still, animals quickly learned to discriminate the two stimulation sites. Whether information was read out from these overlapping activations by spatial integration within sensory cortex, or by other cortical and/or subcortical areas excited via connections preserving information about the stimulation sites (Tehovnik et al., 2006), cannot be resolved by our analysis. However, our study provides important insight in cortical processing during the meaningful, behavioral interpretation of ICMS, which presupposes readout of site information from the evoked cortical activation.

Early cortical activity patterns identified by our multivariate classification approach were largely independent from learning, and time-locked to stimulus onset at a latency corresponding to the first EEP peak (N20). Furthermore, cortical sites near the stimulation sites contributed more information to early patterns than distal sites. Obviously, early patterns reflected the different location of the focus of activation evoked by CS_{go} and CS_{no-go}, respectively. Thus, similar to early patterns found with acoustic stimulation (Ohl et al., 2000, 2003a,b), they represented a topographically organized cortical activation encoding information about physical stimulus properties, in our case the stimulation site.

Late patterns also carried information about the conditioned stimuli. However, in contrast to early patterns, they emerged with discrimination learning from the ongoing activity at variable latencies after the end of the stimulus. Thereby, late patterns predominantly occurred in trials when animals correctly interpreted the conditioned electrical stimuli.

Using a similar classification approach, late spatial patterns have been also identified in the course of auditory discrimination (Barrie et al., 1996; Ohl et al., 2003a,b) and categorization learning (Ohl et al., 2001). As in our study, these patterns emerged from the ongoing poststimulus activity and were rather related to the meaningful interpretation of the stimuli, like their belonging to categories formed by previous experience, than to physical stimulus properties per se like their spectral content. This indicates a similar functional role of late patterns in auditory learning, and in learning with artificial, electrical stimuli. Animal and human studies further suggest that electrical stimulation might generate different pitch-sensations depending on the site of stimulation within the tonotopic map of primary auditory cortex (Dobelle et al., 1973; Otto et al., 2005). Animals in our study might therefore have actually performed pitch discrimination. The question how the late patterns are specifically linked to cognitive functions like perception, decision making, or memory retrieval, needs further investigation (Scheich et al., 2007).

Information of late patterns induced by auditory learning was distributed over a large cortical region in a nontopographical way (Ohl et al., 2003a,b). In contrast, with ICMS, late pattern information was rather confined to cortical sites distal to the stimulation site, outside the range of the lateral spatial spread ($>\sim$1.2 mm) of the first EEP peak (N20). This demonstrates that the stimulated cortical area not only encoded information about the stimulation sites by its focal, stimulus-driven activation, but also provided meaningful signals in its ongoing activity related to the interpretation of ICMS by the learning animal. This involved the stimulated area as a whole, and apparently required large-scale integration in the cortex. Consistently, ICMS in sensory cortex can generate percepts, regardless of the hierarchical level of the stimulated area (Murphey and Maunsell, 2007). As further indicated by the long, variable latencies, late pattern formation pre-

sumably involved feedback interactions with other brain areas like prefrontal cortex, hippocampus, and amygdala.

However, ICMS locally interfered with the ongoing cortical dynamics by suppressing late pattern formation near ($<\sim$1.2 mm) the stimulation sites. As reviewed by Freeman (2000), late patterns are carried by ongoing cortical oscillations in the β- and γ-band that remain transiently coherent for up to $\sim$200 ms over distances of millimeters. Cortical regions strongly affected by ICMS were apparently not included into this global coherent state. This might be due to a local, artificial synchronization or to imbalances between cortical excitation and inhibition caused by ICMS, which actually can lead to highly synchronized responses with each single pulse, to entrainment of cortical activity, prolonged inhibition, and rebound excitation (Butovas and Schwarz, 2003). Also, ICMS might have interfered with plastic changes involved in pattern formation (Recanzone et al., 1992; Freeman, 2000). Hence, it would be important to study how pattern suppression depends on stimulus parameters like pulse rate and amplitude, or location, and whether it is transient within trials or long-lasting across trials.

Despite the observed local pattern suppression, animals learned to discriminate stimulation sites within a few sessions of training, comparable to pure tone discrimination learning (Ohl et al., 1999). Still, pattern suppression might have more subtle effects, not observable by our behavioral measures. Thus, a restriction of pattern formation to subpopulations of neurons distal to the stimulation sites could limit the capacity of the cortex for large-scale integration. Also, behavioral effects might arise from a more widespread pattern suppression, e.g., by simultaneous ICMS at spatially distributed, multiple sites. Although speculative, pattern suppression might relate to the fact that ICMS-evoked percepts are always simple and stereotyped (Dobelle et al., 1973; Schmidt et al., 1996), and that a meaningful interpretation of more complex spatial ICMS patterns seems to be difficult both for humans and animals (Schmidt et al., 1996; Fitzsimmons et al., 2007).

As the interaction between ICMS and ongoing cortical activity is involved in the meaningful interpretation of ICMS, it should be taken into account in the design of ICMS protocols and cortical neuroprostheses. For example, a proper timing between incoming electrical stimuli and state changes in the ongoing cortical dynamics might be important for the meaningful interpretation of ICMS, especially with complex stimulation patterns (Freeman, 2000). Thus, the development of bidirectional, interactive cortical interfaces that guarantee proper timing and shaping of ICMS conditional on the momentary cortical state, could improve the functioning of such interfaces (Kipke et al., 2008).

References

Barrie JM, Freeman WJ, Lenhart MD (1996) Spatiotemporal analysis of prepyriform, visual, auditory, and somesthetic surface EEGs in trained rabbits. J Neurophysiol 76:520–539.

Bartlett JR, DeYoe EA, Doty RW, Lee BB, Lewine JD, Negrão N, Overman WH Jr (2005) Psychophysics of electrical stimulation of striate cortex in macaques. J Neurophysiol 94:3430–3442.

Bradley DC, Troyk PR, Berg JA, Bak M, Cogan S, Erickson R, Kufta C, Mascaro M, McCreery D, Schmidt EM, Towle VL, Xu H (2005) Visuotopic mapping through a multichannel stimulating implant in primate V1. J Neurophysiol 93:1659–1670.

Brindley GS, Lewin WS (1968) The sensations produced by electrical stimulation of the visual cortex. J Physiol 196:479–493.

Budinger E, Heil P, Scheich H (2000) Functional organization of auditory cortex in the Mongolian gerbil (*Meriones unguiculatus*). III. Anatomical subdivisions and corticocortical connections. Eur J Neurosci 12:2425–2451.

Butovas S, Schwarz C (2003) Spatiotemporal effects of microstimulation in rat neocortex: a parametric study using multielectrode recordings. J Neurophysiol 90:3024–3039.

Butovas S, Hormuzdi SG, Monyer H, Schwarz C (2006) Effects of electrically coupled inhibitory networks on local neuronal responses to intracortical microstimulation. J Neurophysiol 96:1227–1236.

Cohen MR, Newsome WT (2004) What electrical microstimulation has revealed about the neural basis of cognition. Curr Opin Neurobiol 14:169–177.

DeYoe EA, Lewine JD, Doty RW (2005) Laminar variation in threshold for detection of electrical excitation of striate cortex by macaques. J Neurophysiol 94:3443–3450.

Digital Signal Processing Committee (1979) Lowpass interpolation algorithm 8.1. In: Programs for digital signal processing. New York: IEEE.

Dobelle WH, Stensaas SS, Mladejovsky MG, Smith JB (1973) A prosthesis for the deaf based on cortical stimulation. Ann Otol Rhinol Laryngol 82:445–463.

Donoghue JP (2002) Connecting cortex to machines: recent advances in brain interfaces. Nat Neurosci 5 (Suppl):1085–1088.

Fernández E, Pelayo F, Romero S, Bongard M, Marin C, Alfaro A, Merabet L (2005) Development of a cortical visual neuroprosthesis for the blind: the relevance of neuroplasticity. J Neural Eng 2:R1–R12.

Fitzsimmons NA, Drake W, Hanson TL, Lebedev MA, Nicolelis MA (2007) Primate reaching cued by multichannel spatiotemporal cortical microstimulation. J Neurosci 27:5593–5602.

Freeman W (1975) Mass action in the nervous system, pp 249–252. New York: Academic.

Freeman WJ (2000) Mesoscopic neurodynamics: from neuron to brain. J Physiol Paris 94:303–322.

Freeman WJ, Baird B (1987) Relation of olfactory EEG to behavior: spatial analysis. Behav Neurosci 101:393–408.

Green DM, Swets JA (1966) Signal detection theory and psychophysics. New York: Wiley.

Hotelling H (1953) New light on the correlation coefficient and its transformation. J R Stat Soc B 15:193–232.

Huynh H, Feldt LS (1976) Estimation of the Box correction for degrees of freedom from sample data in randomised block and split-plot designs. J Educ Stat 1:69–82.

Kipke DR, Shain W, Buzsáki G, Fetz E, Henderson JM, Hetke JF, Schalk G (2008) Advanced neurotechnologies for chronic neural interfaces: new horizons and clinical opportunities. J Neurosci 28:11830–11838.

Leal-Campanario R, Delgado-García JM, Gruart A (2006) Microstimulation of the sornatosensory cortex can substitute for vibrissa stimulation during Pavlovian conditioning. Proc Natl Acad Sci U S A 103:10052–10057.

Murphey DK, Maunsell JH (2007) Behavioral detection of electrical microstimulation in different cortical visual areas. Curr Biol 17:862–867.

Normann RA, Maynard EM, Rousche PJ, Warren DJ (1999) A neural interface for a cortical vision prosthesis. Vis Res 39:2577–2587.

Nowak LG, Bullier J (1998a) Axons, but not cell bodies, are activated by electrical stimulation in cortical gray matter. I. Evidence from chronaxie measurements. Exp Brain Res 118:477–488.

Nowak LG, Bullier J (1998b) Axons, but not cell bodies, are activated by electrical stimulation in cortical gray matter. II. Evidence from selective inactivation of cell bodies and axon initial segments. Exp Brain Res 118:489–500.

Ohl FW, Scheich H (2005) Learning-induced plasticity in animal and human auditory cortex. Curr Opin Neurobiol 15:470–477.

Ohl FW, Wetzel W, Wagner T, Rech A, Scheich H (1999) Bilateral ablation of auditory cortex in Mongolian gerbil affects discrimination of frequency modulated tones but not of pure tones. Learn Mem 6:347–362.

Ohl FW, Scheich H, Freeman WJ (2000) Topographic analysis of epidural pure-tone-evoked potentials in gerbil auditory cortex. J Neurophysiol 83:3123–3132.

Ohl FW, Scheich H, Freeman WJ (2001) Change in pattern of ongoing cortical activity with auditory category learning. Nature 412:733–736.

Ohl FW, Deliano M, Scheich H, Freeman WJ (2003a) Early and late patterns of stimulus-related activity in auditory cortex of trained animals. Biol Cybern 88:374–379.

Ohl FW, Deliano M, Scheich H, Freeman WJ (2003b) Analysis of evoked and emergent patterns of stimulus-related auditory cortical activity. Rev Neurosci 14:35–42.

Otto KJ, Rousche PJ, Kipke DR (2005) Microstimulation in auditory cortex provides a substrate for detailed behaviors. Hear Res 210:112–117.

Press W, Teukolsky SA, Vetterlin WT, Flannery BP (2007) Numerical recipes: the art of scientific computing, pp 773–836. New York: Cambridge UP.

Recanzone GH, Merzenich MM, Dinse HR (1992) Expansion of the cortical representation of a specific skin field in primary somatosensory cortex by intracortical microstimulation. Cereb Cortex 2:181–196.

Romo R, Hernández A, Zainos A, Brody CD, Lemus L (2000) Sensing without touching: psychophysical performance based on cortical microstimulation. Neuron 26:273–278.

Rousche PJ, Otto KJ, Reilly MP, Kipke DR (2003) Single electrode microstimulation of rat auditory cortex: an evaluation of behavioral performance. Hear Res 179:62–71.

Scheich H, Breindl A (2002) An animal model of auditory cortex prostheses. Audiol Neuro-Otol 7:191–194.

Scheich H, Brechmann A, Brosch M, Budinger E, Ohl FW (2007) The cognitive auditory cortex: task-specificity of stimulus representations. Hear Res 229:213–224.

Schmidt EM, Bak MJ, Hambrecht FT, Kufta CV, O'Rourke DK, Vallabhanath P (1996) Feasibility of a visual prosthesis for the blind based on intracortical microstimulation of the visual cortex. Brain 119:507–522.

Stoney SD Jr, Thompson WD, Asanuma H (1968) Excitation of pyramidal tract cells by intracortical microstimulation: effective extent of stimulating current. J Neurophysiol 31:659–669.

Tehovnik EJ (1996) Electrical stimulation of neural tissue to evoke behavioral responses. J Neurosci Methods 65:1–17.

Tehovnik EJ, Tolias AS, Sultan F, Slocum WM, Logothetis NK (2006) Direct and indirect activation of cortical neurons by electrical microstimulation. J Neurophysiol 96:512–521.

Tolias AS, Sultan F, Augath M, Oeltermann A, Tehovnik EJ, Schiller PH, Logothetis NK (2005) Mapping cortical activity elicited with electrical microstimulation using fMRI in the macaque. Neuron 48:901–911.