EJN fens European Journal of Neuroscience

European Journal of Neuroscience, Vol. 37, pp. 1830–1842, 2013 doi:10.1111/ejn.12180

COGNITIVE NEUROSCIENCE

Stable bottom-up processing during dynamic top-down modulations in monkey auditory cortex

Roohollah Massoudi,[1] Marc M. Van Wanrooij,[1,2] Sigrid M. C. I. Van Wetter,[1] Huib Versnel[3] and A. John Van Opstal[1]

[1]Department of Biophysics, Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Heyendaalseweg 135, 6525 AJ, Nijmegen, The Netherlands

[2]Department of Otorhinolaryngology, Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands

[3]Department of Otorhinolaryngology and Head & Neck Surgery, Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, Utrecht, The Netherlands

Keywords: awake, electrophysiology, primate, reaction time, ripple, spectrotemporal receptive field

Abstract

It is unclear whether top-down processing in the auditory cortex (AC) interferes with its bottom-up analysis of sound. Recent studies indicated non-acoustic modulations of AC responses, and that attention changes a neuron's spectrotemporal tuning. As a result, the AC would seem ill-suited to represent a stable acoustic environment, which is deemed crucial for auditory perception. To assess whether top-down signals influence acoustic tuning in tasks without directed attention, we compared monkey single-unit AC responses to dynamic spectrotemporal sounds under different behavioral conditions. Recordings were mostly made from neurons located in primary fields (primary AC and area R of the AC) that were well tuned to pure tones, with short onset latencies. We demonstrated that responses in the AC were substantially modulated during an auditory detection task and that these modulations were systematically related to top-down processes. Importantly, despite these significant modulations, the spectrotemporal receptive fields of all neurons remained remarkably stable. Our results suggest multiplexed encoding of bottom-up acoustic and top-down task-related signals at single AC neurons. This mechanism preserves a stable representation of the acoustic environment despite strong non-acoustic modulations.

Introduction

To establish perceptual invariance, the auditory system faces the dichotomy of either adapting its acoustic representations when perceptual errors call for learning and plasticity, or preserving a stable acoustic representation despite changes of the behavioral context. Much of the low-level (bottom-up) acoustic processing seems to take place at subcortical stages (Nelken *et al.*, 2003; Palmer, 2007), with evidence of higher-order, non-acoustic (top-down) cognitive (Griffiths *et al.*, 2004; Gutschalk *et al.*, 2005; Snyder *et al.*, 2006; Nelken & Bar-Yosef, 2008) and multisensory (Fu *et al.*, 2003; Lakatos *et al.*, 2007; Ghazanfar *et al.*, 2008; Kayser *et al.*, 2008, 2010) processing in the auditory cortex (AC).

As the AC occupies a central position within the acoustic and non-acoustic processing pathways (Aertsen *et al.*, 1981; Edeline *et al.*, 2001; Edeline, 2003; Fritz *et al.*, 2003, 2005a,b; Brosch *et al.*, 2005; Schroeder & Foxe, 2005; Polley *et al.*, 2006; Elhilali *et al.*, 2007; King *et al.*, 2007; Riecke *et al.*, 2007; Atiani *et al.*, 2009), it is expected to play a role in stable perception of the auditory environment, but evidence is lacking. Recent studies have indicated that consistent attention to a specific acoustic feature of a target sound (e.g. a cued frequency) can induce changes in the spectrotemporal tuning characteristics of AC neurons toward the attended feature (Fritz *et al.*, 2003, 2005a,b; Elhilali *et al.*, 2007; Atiani *et al.*, 2009). However, if such changes are contingent upon general auditory task performance, the question arises how the auditory system maintains a stable representation of the acoustic environment under different behavioral conditions.

To address this, we investigated whether AC cells can preserve the bottom-up encoding of a sound's spectral–temporal acoustics, despite prominent top-down signals from non-acoustic, context-dependent sources. We analysed single-unit activity in monkey AC for three different situations, one in a passive condition, and two in an active listening condition. In the passive condition, we merely exposed animals to sounds that started automatically with 500 ms static noise, followed by a spectral–temporal dynamic ripple, and determined the spectrotemporal receptive field (STRF) from the ripple-evoked responses (Depireux *et al.*, 2001; Eggermont, 2011). In the active listening condition the animal had to manually react to the change from static noise to dynamic ripple to obtain a reward. In half of the trials static noise lasted for 500 ms (A500), whereas in the other half it lasted for 1000 ms (A1000). As both trial types were randomly interleaved, the change occurred unexpectedly in A500 trials. However, animals could

Correspondence: A. John Van Opstal, as above.
E-mail: j.vanopstal@donders.ru.nl

Received 31 January 2013, accepted 11 February 2013

© 2013 Federation of European Neuroscience Societies and Blackwell Publishing Ltd

fully predict ripple onset in A1000 trials after more than 500 ms static noise had passed, which added a cognitive factor to the active paradigm. This predictive effect was clearly demonstrated by the animal's response-reaction times and neural responses. Our recordings demonstrated that, in active trials, neural activity to static noise and dynamic ripples differed markedly from passively evoked responses. We showed that the STRFs nonetheless remained identical for all paradigms, in line with perceptual stability, and that the neural modulations in the active paradigms reflected a true task-related top-down signal.

Materials and methods

Subjects

Neurophysiological recordings were performed in the left AC of two adult male rhesus monkeys (*Macaca mulatta*; Monkey J, 7–9 kg; Monkey T, 8–10 kg). Each monkey participated in the recording sessions for about 2 years. They were trained to respond to the onset of spectral–temporal modulations of a sound to receive a drop of water as a reward. Experiments were conducted in accordance with the European Communities Parliament and Council Directive (September 22, 2010, 2010/63/EU). All experimental protocols were approved by the Ethics Committee on Animal Research of the Radboud University Nijmegen (Radboud University Dier Experimenten Commissie). Monkeys were pair-housed to promote normal interactive behavior. Our procedures followed the water-restriction protocol of the Animal Use and Care Administrative Advisory Committee of the University of California at Davis (2001). At about 24 h before the start of an experimental session, water intake was limited to 20 mL/kg. In the experiment, the monkey earned a small water reward of 0.2 mL per successful trial. We ensured that monkeys earned at least the minimum of 20 mL/kg on an experimental day. After an experimental session, water was supplemented to the required minimum amount, if needed, and the animal received additional pieces of fruit. At weekends, the animal's fluid intake was increased to 400 mL daily. To monitor the animal's health status, we kept records of body weight, and water and food intake. Expert veterinarian assistance was available on site. Quarterly testing of hematocrit values ensured that the animal's kidney function remained within the normal physiological range. Whenever an animal showed signs of discomfort, or illness, experiments were stopped and the animal was treated until the problem was solved.

Surgical procedures

When the initial sound-change detection training was completed (day-to-day performance level became 80%, or better), surgery was performed under full anesthesia and sterile conditions. Anesthesia was maintained by artificial respiration (0.5% isoflurane and N_2O), and additional pentobarbital (IV; 3 mg/kg/hour), ketamine (IM; 0.1 mL/kg), and fentanyl (IV; 20 μg/kg/hour) were administered. A stainless-steel recording chamber (12 mm diameter) was placed over a trepaned hole in the skull (10 mm diameter), the coordinates of which had been determined on the basis of magnetic resonance images. The chamber allowed a vertical approach of the AC on the left side. A stainless-steel bolt, embedded in dental cement on the skull, allowed firm fixation of the head during recording sessions.

Experimental setup

The head-restrained monkey sat in a primate chair within a completely dark and sound-attenuated room (2.45 × 2.45 × 3.5 m), while a glass-coated tungsten microelectrode (impedance 1–2 MΩ; Alpha Omega, Ubstadt-Weiher, Germany) was carefully positioned and lowered into the brain through a stainless-steel guide tube by an electronically-driven stepping motor (MM-3M-F-1; National Aperture Inc.). Electrode signals were grounded to a contact in the recording chamber. The analog electrode signal was amplified (model A-1; BAK Electronics), band-pass filtered between 0.1 and 15 kHz (custom-built eighth-order LP Butterworth filter; HP with Krohn-Hite; model 3343), and monitored through a speaker and oscilloscope. The raw signal was then digitised (at 25 kHz, A/D convertor, TDT2 system; module AD-1; Tucker-Davis Technologies). An automated spike-detection algorithm (BRAINWARE, V 9.07 for Tucker-Davis Technologies), run on a PC (WINDOWS 98; DELL), isolated action-potential waveforms and determined their moments of occurrence. Data analysis and spike sorting were performed offline in MATLAB (version 7.9.0, R2009b; MathWorks, Natick, MA, USA).

Sound stimuli

Sound stimuli were digitally generated at a sampling rate of 100 kHz and delivered via BRAINWARE sofware and TDT2 hardware. A trigger provided by a TG6 module started sound presentation (DA1, low-pass filtered at 20 kHz through an FT6 module), and spike data acquisition. The sounds were presented in the free field from the frontal central position at a distance of 100 cm from the monkey by a speaker (Blaupunkt PCxg352, flat frequency characteristics within 5 dB between 0.2 and 20 kHz) at a fixed sound level of 60 dB [A-weighted (dBA), set by two programmable attenuators, PA4, and measured with a calibrated sound amplifier and microphone, BK2610/BK4134, Bruel and Kjaer].

The ambient background acoustic noise level was ~ 30 dBA. Reflections above 500 Hz were effectively absorbed by acoustic foam that was mounted on the walls, floor, ceiling, and every large object present. Ripple stimuli were used to determine the spectro-temporal properties of the recorded cells, and to investigate how detection of a change from a static noise to a dynamic ripple influenced the neuron's responses. They consisted of a broadband complex of 126 components equally distributed (20/octave) from 250 Hz to 20 kHz (Depireux *et al.*, 2001; Versnel *et al.*, 2009). All components had random phase. The amplitude of each component was described as follows

$$S(t,x) = 1 + \sin(2\pi wt + \Omega x) \quad (1)$$

where t is time (s), x is position of the spectral component in octaves above the lowest frequency (250 Hz), ω is ripple velocity (Hz) and Ω is ripple density (cycles/octave). The ripple stimuli followed static noise of either 500 or 1000 ms duration, which consisted of the same components as the ripple but all with equal amplitude.

The set of 55 different ripples used in our study consisted of all combinations of 11 different ripple densities (Ω in [−2.0 : 0.4 : +2.0] cycles/octave) and five different velocities (ω in [8 : 8 : 40] Hz). A negative density corresponds to an upward direction of the spectral envelopes, a positive density to a downward direction, and Ω = 0 means a pure amplitude-modulated sound (see Fig. 1 in Versnel *et al.*, 2009). The modulation depth was fixed at 100%. The ripple stimuli were used to compute the STRF of each cell, according to the procedure described below and illustrated in Fig. 3. The sound level of the static noise and ripple was 60 dB sound pressure level.

© 2013 Federation of European Neuroscience Societies and Blackwell Publishing Ltd

14609568, 2013, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ejn.12180 by Radboud University Nijmegen, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

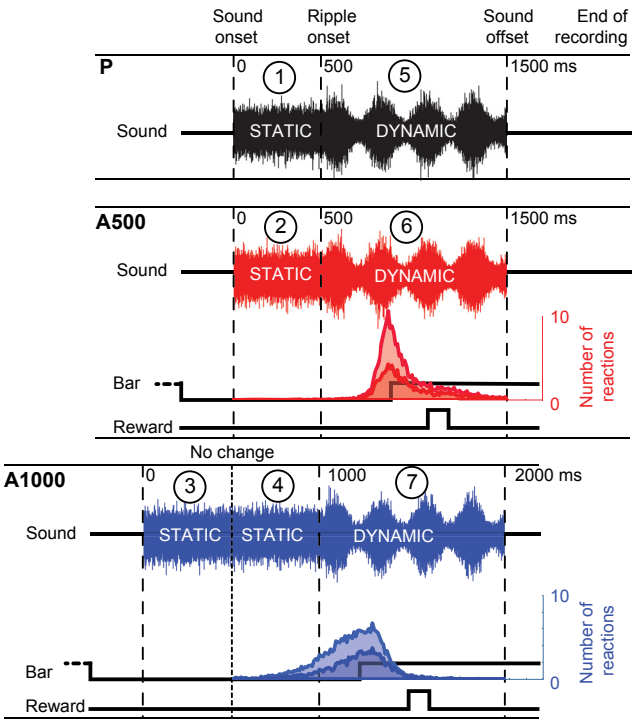

FIG. 1. Experimental paradigms. Tasks: passive sound exposure (black), and active [A500 (red) and A1000 (blue)] trials, with relevant events in the trial indicated. Encircled numbers identify recording epochs during presentation of static noise (1–4) and dynamic ripples (5–7) for the different trial types, used for the correlation analyses throughout this article. For illustrative purposes, the ripple is presented as an amplitude-modulated waveform. Histograms: reaction-time distributions for active conditions show faster responses in the A1000 trials than the A500 trials (light shading, Monkey T; dark shading, Monkey J).

Pure tones were presented passively in a separate block at the start of a cell's recording, and were used exclusively to extract the frequency-tuning properties of the recorded cells (Recanzone *et al.*, 2000). They lasted for 150 ms and were presented at 13 frequencies from 250 to 16 000 Hz at four different sound levels (10, 30, 50, and 70 dB sound pressure level), set by the programmable PA4 attenuators. Trials were presented at least four times in a randomised manner. The spontaneous firing rate was determined during a 300 ms pre-stimulus period across all ($\geq$ 208) presentations. Driven responses were defined as the average firing rate of all ($\geq$ 4) presentations of each stimulus that was greater than the mean plus two SDs of the spontaneous activity. The best frequency of each neuron was defined as the frequency that produced a driven response at the lowest intensity (threshold). The cell's response onset latency was defined as the moment after pure-tone onset at which there was a driven response for at least three consecutive 2 ms time bins. The tuning bandwidth was measured as the difference in octaves between the low- and high-frequency edges of the driven responses, linearly interpolated at 10 dB above the threshold.

Experimental paradigms

Single-unit responses to ripples were measured in two different paradigms (Fig. 1): (i) passive sound exposure, in which the monkey was exposed to the rippled sounds without a task, and the straight-ahead fixation light was off; and (ii) active listening, in which upon presentation of a red fixation light at straight ahead, the monkey initiated a trial by pressing a bar. It had to attend to the same sounds as presented during passive exposure, as it had to respond to the ripple-modulation onset within 100–700 ms in order to receive a drop of water as reward. Trials in which the monkey did not detect the ripple onset were repeated at a randomly selected trial within the same recording block.

The passive and active conditions were presented in two different blocks, typically starting with the passive condition. In some experiments, the passive condition was repeated after the active condition to check the cell's response and STRF stability (not shown). In the passive condition the trial started automatically with the static noise (Fig. 1, sound onset). After 500 ms (Fig. 1, ripple onset), the ripple was presented for 1000 ms (Fig. 1, sound offset). Recording duration was 2500 ms, which started 300 ms before the sound onset, and ended 700 ms after the sound offset (Fig. 1, black). The number of ripple repetitions was between 4 and 10.

The active listening block contained A500 and A1000 trials (based on the duration of the static noise epochs, at 500 and 1000 ms, respectively) that were randomly interleaved in the experiment at equal probability. The static noise was presented at 300 ms after the monkey pressed the bar, and then changed into a ripple (at a pseudorandomly selected velocity and density) that lasted for 1000 ms. We used A500 and A1000 trials for two reasons: (i) to prevent the monkeys from learning that the ripple started after a constant time, which might have prompted them to start counting instead of listening; and (ii) to add a cognitive (i.e. predictive) aspect to the behavioral trials; the well-trained monkeys could anticipate a change at 1000 ms in A1000 trials with a probability of 1.

The recording duration was 2500 ms in both A500 and A1000 trials, and the number of repetitions depended on the monkey's performance (usually no repetition). The monkey's performance scores were ~80% (A500, 80%; A1000, 81%); Monkey J had slightly better performance (A500, 81%; A1000, 85%) than Monkey T (A500, 78%; A1000, 76%).

Data selection

We recorded neural responses to the dynamic ripples in both the passive and active conditions from 128 cells in the AC of the two monkeys (Monkey J, $n = 65$; Monkey T, $n = 63$). For 23 of 128 cells, we could not collect one full repetition for all ripples in either or both active trial types. Thus, we had, in total, 105 cells with complete recordings (Monkey J, $n = 52$; Monkey T, $n = 53$). For most analyses in this study we further selected cells on the basis of an STRF-based quality criterion that excluded neurons with noisy responsiveness in the active condition. We therefore calculated the signal-to-noise ratio (SNR) in the STRF for the pooled active trials by dividing the maximum of the entire active STRF by its SD (σ) between 75 and 125 ms (note that the STRFs of AC cells typically do not contain acoustic tuning information after 75 ms; see also Results)

$$SNR = \frac{\max(\mathrm{STRF})}{\sigma} \qquad (2)$$

The distribution of log(SNR) values for the active STRFs (pooled for A1000 and A500 trials) across the entire population of cells is shown in Fig. 2. The SNR values in our recorded population show up as a bimodal distribution (low vs. high SNR values); the border between the two distributions was at an SNR of about 3.3. We included 51 cells (Monkey J, $n = 31$; Monkey T, $n = 20$) with SNRs above 3.3 for both the individual and population-based analyses. In the Results we also show the population-based

© 2013 Federation of European Neuroscience Societies and Blackwell Publishing Ltd

14609568, 2013, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ejn.12180 by Radboud University Nijmegen, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

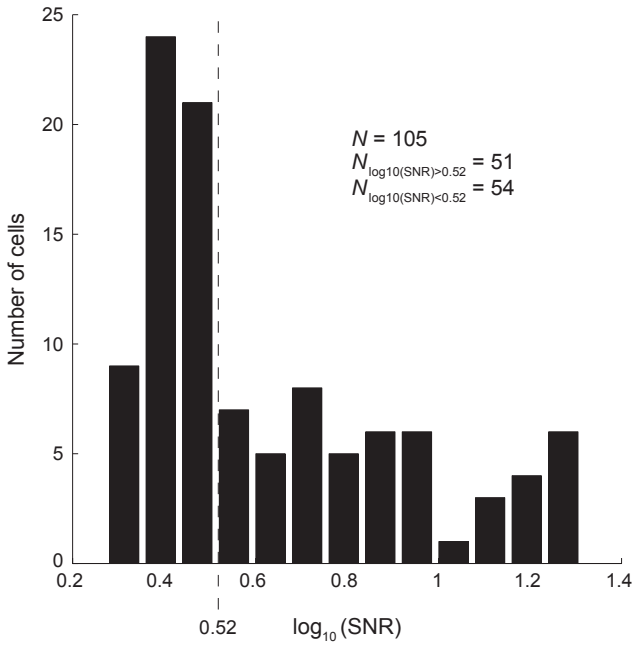

FIG. 2. The logarithmic distribution of SNR values across the entire population of cells. SNRs can be divided into low vs. high SNR values; the border between the two distributions was set at SNR = 3.3 (dashed line).

analysis for the 54 cells with SNRs below 3.3. Low SNR values in active trials could result because the number of presentations for each ripple in the active tasks had to be limited (typically to one), as otherwise the monkey would become satiated before the end of the experiment.

Characterisation of recording sites

Although we cannot with certainty identify the exact AC subdivision(s) in which we encountered individual neurons, we are confident that we recorded from the AC core [primary AC (A1) and rostral area (R) of the AC] and caudomedial field (CM) for the following reasons: (i) magnetic resonance imaging scans were used for stereotaxic placement of the recording chamber, and the subsequent coordinates of the successful recording sites within the chamber corresponded closely to the stereotaxic coordinates of A1 as provided by the atlas of the Rhesus monkey brain by Paxinos *et al.* (2000); (ii) before reaching an AC recording site there was a physiologically silent period, corresponding to the gap between the upper and lower parts of the lateral sulcus (Kaas & Hackett, 2000); (iii) tone-onset latency of the recorded sites was 22.6 ± 5.9 and 23.6 ± 5.6 ms for Monkey J and Monkey T, respectively; (iv) all neurons responded well to pure tones (best frequency, 250–16000 Hz); (v) the pure-tone tuning bandwidths for Monkey J and Monkey T were 1.5 ± 1.2 and 1.5 ± 1.3 octaves, respectively; and (vi) the pure-tone threshold for Monkey J was 21 ± 13 dB sound pressure level, and for Monkey T was 23 ± 12 dB sound pressure level. All of the above-mentioned tuning characteristics fall in the same ranges as reported by Recanzone *et al.* (2000) for behaving monkeys in AC areas A1, rostral area of the AC and caudomedial field.

Recording stability was verified by spike-waveform variability during the different behavioral tasks. When available, the passive STRFs could be obtained from pooling data recorded both before and after the active listening condition. Due to recording stability, this often led to a better estimate of the STRF.

Data analysis

Spike-density function

To convert each raster plot into a continuous spike-density function, we first binned the recorded spike times into a digital sequence at a time resolution of 1 ms, and then convolved the data with a Gaussian kernel with a SD of 5 ms.

Spectrotemporal receptive field

We estimated a cell's STRF from ripple stimuli by using the same off-line method as described by Versnel *et al.* (2009). First, detected spikes (Fig. 3A) were sorted and binned into peri-stimulus time histograms. We then wrapped the 900 ms response window (the ripple duration, excluding the first 100 ms to exclude transient onset responses) into 32 bin period histograms, in which the ripple velocity determined the period as 1/ω (Fig. 3B). We subsequently performed a fast Fourier transform on the period histograms. The magnitude $A(\omega, \Omega)$ (spikes/s) and phase $\Phi(\omega,\Omega)$ (rad) (Figs 3C and 3D) of the period histograms were derived from the first harmonic of the resulting Fourier spectrum to generate the spectrotemporal transfer function

$$T(\omega, \Omega) = A(\omega, \Omega) \cdot \exp(i \cdot \varphi(\omega, \Omega)) \quad (3)$$

The two-dimensional inverse Fourier transformation of $T(\omega,\Omega)$ then produces the spectrotemporal response field, or STRF (Fig. 3E), of the cell

$$\mathrm{STRF}(t, x) = \mathrm{FFT}^{-1}[T(\omega, \Omega)] \quad (4)$$

where x is the frequency in octaves (between 0 and 2.5 in 0.25 octave steps), and t is running from 0 to 125 ms (at 12.5 ms resolution). The spectral dimension (abscissa) of the STRF reflects the frequency tuning, and the temporal direction (ordinate) reflects the cell's linear impulse response. Note that the frequency range of the STRF is determined by the step size of the ripple densities employed in the experiment: [range x] = 1/[step size Ω]. Typically, the step size was 0.4 cycles/octave, resulting in a frequency range of 2.5 octaves. Likewise, the temporal range is determined by the resolution in applied ripple velocities: [range t] = 1/[step size ω], leading to 125 ms. Note that the position of the frequency range (in Hz) is ambiguous, as the lower frequency of the STRF could be either 250 Hz (the lowest component in the ripple stimuli), or multiples of 2.5 octaves above 250 Hz (i.e. at 1414 or 8000 Hz). The pure-tone responses were used to resolve this ambiguity. We shifted the STRF 1 octave higher or lower if the excitatory or inhibitory response areas were located near the edges of the STRF frequency domain, to relocate the center of the STRF close to the middle of the frequency axis.

Analysis of spectrotemporal receptive field changes

We compared the active STRFs ($\mathrm{STRF_A}$) with the passive STRFs ($\mathrm{STRF_P}$) to see whether the spectrotemporal tuning changes systematically in the different listening conditions. To that end, we first aligned the frequency bin (x) at which the maximum of the passive STRF was obtained, with the center of the spectral range (x'), and shifted the active STRFs by the same amount ($\Delta x = x' - x$). We then performed a linear regression analysis for each cell k by comparing

© 2013 Federation of European Neuroscience Societies and Blackwell Publishing Ltd

14609568, 2013, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ejn.12180 by Radboud University Nijmegen, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

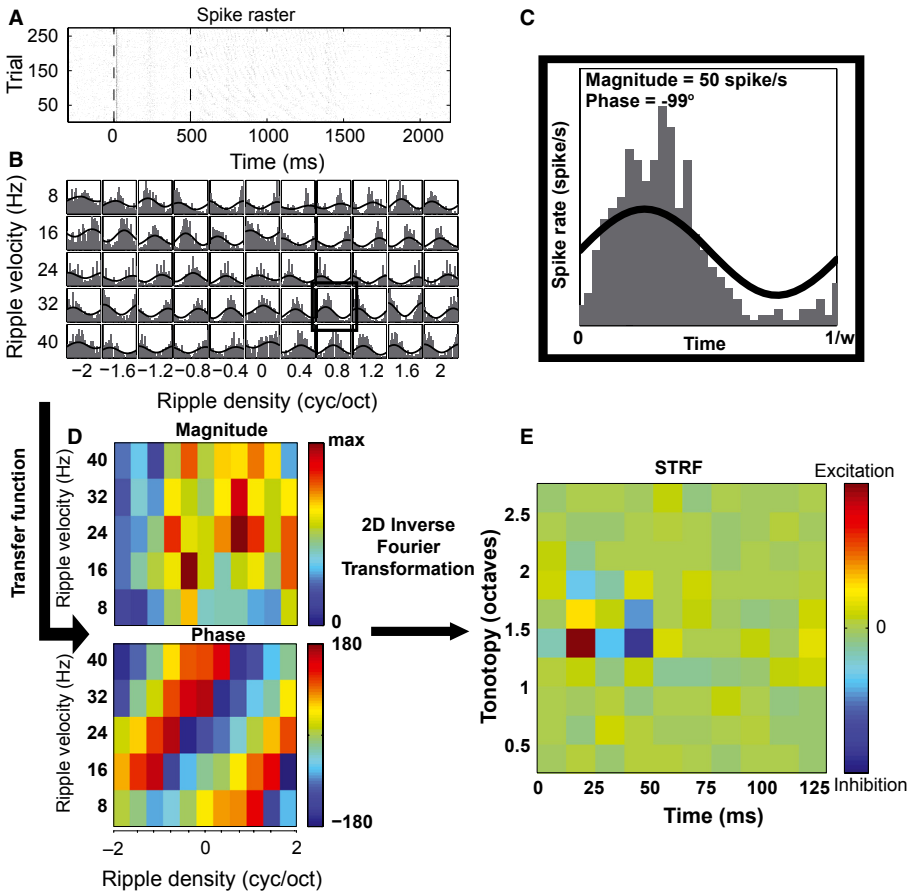

FIG. 3. Procedure to derive the STRF from neural responses to the 55 ripples that varied over the full velocity–density range. (A) Sorted spike trains of AC neuron J93 to ripples with densities between −2.0 and +2.0 cycles/octave, and velocities between 8 and 40 Hz. Each row represents one trial and each dot is a spike. Note the clear phase locking to ripple velocity. (B) Period histograms (32 bins) for all 55 stimuli. The period is defined by the inverse of ripple velocity. (C) Magnitude (in spikes/s) and phase (in degrees) were calculated from the best-fit cosine to the period histogram (here for the highlighted ripple at 0.8 cycles/octave and 32 Hz in B). (D) Magnitude and phase characteristics of the total ripple transfer function. (E) STRF obtained from two-dimensional inverse Fourier transformation of the complex transfer function.

the active and passive STRFs within ±0.25 octaves around the center of the spectral range, over a time window within 12.5–37.5 ms (called the local STRF)

$$\mathrm{STRF}_{\mathrm{A},k}(t,x') = g_k \mathrm{STRF}_{\mathrm{P},k}(t,x') + b_k \tag{5}$$

where g_k is a fixed gain that reflects the change in contrast between excitation and inhibition during active listening, b_k is a constant bias (either increase in overall excitation, or inhibition) of the local STRF, and x' is the shifted spectral coordinate. We then normalised the entire (non-local) active STRF by

$$\widehat{\mathrm{STRF}}_{\mathrm{A},k}(t,x') = \frac{\mathrm{STRF}_{\mathrm{A},k}(t,x') - b_k}{g_k} \tag{6}$$

Finally, the difference STRF for cell k was determined by

$$\Delta\mathrm{STRF}_k(t,x') = \widehat{\mathrm{STRF}}_{\mathrm{A},k}(t,x') - \mathrm{STRF}_{P,k}(t,x') \tag{7}$$

which is the residual of the linear regression. Any systematic pattern in $\Delta\mathrm{STRF}_k(t, x')$ now reflects a change in spectrotemporal tuning of the neuron, irrespective of any gain change. Finally, the difference STRF was normalised by its root mean square value, to allow averaging across neurons with very different firing rates (Eqn 8):

$$\overline{\Delta\mathrm{STRF}}(t,x') = \frac{1}{N}\sum_{k=1}^{N}\frac{\Delta\mathrm{STRF}_k(t,x')}{\mathrm{rms}_k} \tag{8}$$

Correlation analysis

To determine whether the temporal firing patterns of neurons differed for the three types of listening trials, we calculated the correlation coefficient of the spike-density functions (at 1 ms time resolution). For these calculations, we skipped the first 100 ms of each epoch to avoid artificially high correlations due to a persistent transient onset response.

To quantify the similarity between raw STRFs derived from different behavioral trials, we reshaped each recorded STRF (a matrix, consisting of 10 × 10 entries) into a vector (100 × 1 elements), and calculated Pearson's correlation coefficients for the three possible pairs of two vectors.

Results

Task

As described in the Materials and methods (and Fig. 1), the different durations employed in the static epochs (and hence the timing of the ripple onset to initiate bar release) added a cognitive (i.e. predictive)

© 2013 Federation of European Neuroscience Societies and Blackwell Publishing Ltd

14609568, 2013, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ejn.12180 by Radboud University Nijmegen, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

aspect to the behavioral trials. During the first 500 ms of A500 and A1000 trials, the probability that the change into a dynamic ripple would occur at 500 ms was 0.5 for both trial types. However, the well-trained monkeys could anticipate the change at 1000 ms with a probability of 1 in an A1000 trial, if it did not occur at 500 ms. This aspect of the task is indeed strongly reflected in the animal's behavior. In A500 trials the median manual reaction times (Monkey J, 390 ms; Monkey T, 376 ms; Fig. 1, red histograms) were substantially longer than in A1000 trials (Monkey J, 218 ms; Monkey T, 141 ms; Fig. 1, blue histograms), in which clear predictive responses also occurred (i.e. reaction times below 200 ms, sometimes even before ripple onset). We verified that the reaction times did not systematically change over (and during) recording sessions, and that A1000 trials always elicited faster median reaction times than A500 trials. Therefore, the experiment did not induce perceptual learning, which would have led to faster responses as time progressed, and reaction times could be pooled across sessions.

To assess whether the differences in reaction times for A500 and A1000 trials, as well as the considerable reaction-time variability (see distributions in Fig. 1), could somehow be attributed to certain acoustic features of the different ripples in the two active trial types, we analysed and compared the reaction times for the individual ripples. If animals would always respond to a particular acoustic feature in the ripple stimuli, the pattern of reaction times should be similar for A500 and A1000 trials, albeit perhaps shifted by a constant amount toward lower reaction times for A1000 trials. Because the reaction-time distributions of the two animals were very similar (Fig. 1), we pooled their reaction-time data. Figure 4 shows the mean manual reaction times for each ripple of the monkeys (color coded) obtained during the recordings of all 51 cells. In A500 trials (Fig. 4A), animals detected the ripple onsets on the basis of both spectral and temporal modulations, as a clear spectral–temporal pattern of reaction times emerged from this analysis. Along the spectral dimension (ripple density), reaction times increased with increasing (positive and negative) density, which indicates that higher spectral modulations were consistently more difficult to perceive. At the high-density ripples on both ends, high velocities were more difficult to detect than lower ripple velocities; for these high-density/high-velocity ripples we consistently obtained the longest reaction times. In summary, even though the modulation depth was 100% for all stimuli, the animals used both spectral and temporal acoustic modulations to detect the ripple onsets in the A500 trials.

In contrast, for A1000 trials (Fig. 4B) the clear acoustic pattern observed in Fig. 4A is virtually lost. This strongly suggests that in this case both animals mainly responded to the ripple on the basis of prediction, rather than on the ripple's acoustic characteristics.

Table 1 summarises the performance of the two monkeys in the active trials. Trials were counted 'correct' (and hence rewarded) when the manual reaction times fell between 100 and 700 ms (see Materials and methods). Overall, the correct trial percentages were similar (~ 80%) for both active trials. In about 4% of trials the animals did not respond with a manual reaction ('misses'); these trials were excluded from the database for further analysis. Note that, whereas in 15.3% of the A500 trials animals reacted later than 700 ms, and hardly ever responded too early, for the A1000 trials we obtained the opposite result: 14.5% of the trials had reaction times shorter than 100 ms, but animals hardly ever responded too late. These results further corroborate our interpretation that, for A500 trials, the animal's reactions were based on the sound characteristics, and as some ripples (high-velocity/high-density) were more difficult to detect than others, reaction times could be long; in contrast, for A1000 trials animals often predicted the upcoming change, leading to very short (sometimes even negative; see Fig. 1) reaction times.

Our experimental design contained seven different epochs (Fig. 1) with varying acoustic (static epochs 1–4 vs. dynamic epochs 5–7) and/or behavioral (passive vs. active-unpredictive vs. active-predictive) states. The rationale of our paradigms is summarised in Fig. 5, which shows the predictions of a pairwise correlational analysis between the different epochs, given that neural responses would be modulated by the animal's behavioral state in the active conditions. We reasoned that (i) by comparing neural activity in these epochs for the different behavioral conditions, we could identify potential top-down modulations, and (ii) by comparing the STRFs obtained

TABLE 1. Behavioral performance. Data from both monkeys is pooled. Reaction times were between 100-700 ms for *correct* trials, < 100 ms for *early* trials, and > 700 ms in *late* trials. In the *missed* trials, the monkeys did not respond.

Active trial	Total no.	Correct (%)	Missed (%)	Late (%)	Early (%)
A500	4866	79.60	4.10	15.30	1.00
A1000	4819	80.90	4.00	0.60	14.50

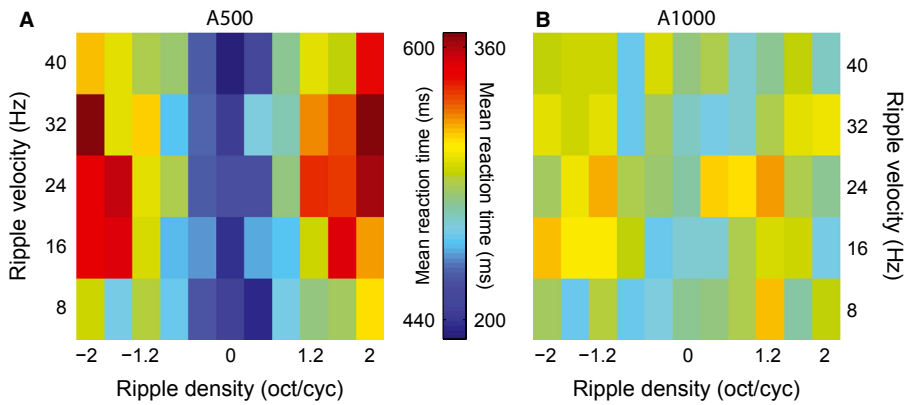

FIG. 4. Mean reaction time of each ripple during A500 (A) and A1000 (B) trials. Because of the resemblance of the two monkeys' results, we pooled the data. There is a clear pattern of reaction time in A500 from lower to higher modulations, which is absent in A1000 trials.

© 2013 Federation of European Neuroscience Societies and Blackwell Publishing Ltd

14609568, 2013, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ejn.12180 by Radboud University Nijmegen, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

from the dynamic epochs, we could determine the effect of top-down factors on the cell's acoustic tuning. Thus, if top-down factors were to systematically influence neural activity of AC cells, responses between passive and active conditions should differ (yielding low correlations, red). Even in the simple detection task these top-down factors might include: attention, sound-change prediction (in epoch 4), motor preparation (handle bar release in epochs 6 and 7), and reward prediction (epochs 6 and 7). Note that the comparison of static epochs 2 and 3 (Fig. 5A) is a crucial control, because the acoustics and the monkey's behavioral states in these epochs are equal; the same sound is presented (500 ms of static noise), the monkey performs a task, and the animal is equally uncertain about the ripple onset. Therefore, despite the influence of top-down factors on AC activity, responses in epochs 2 and 3 should still correlate well (green). Depending on whether or not the bottom-up tuning characteristics change under the influence of task performance, an *a priori* prediction of the correlations between STRFs cannot be made (Fig. 5C, yellow).

Task modulation of auditory cortex activity

Our correlation analysis is based on the recordings from 51 cells (Monkey J, $n = 31$; Monkey T, $n = 20$). All selected cells possessed well-defined STRFs for the passive exposure condition, were selectively tuned to pure tones (best frequencies ranged from 250 Hz to 16 kHz, with the large majority $\leq$ 2 kHz), and recordings were stable during the entire session (see Materials and methods for selection criteria).

Task participation substantially changed neuronal response behavior, as exemplified by the spike-raster plots and the associated average spike densities of three representative AC cells during the dynamic epochs (Fig. 6). For both active listening conditions (Fig. 6A, epochs 6 and 7), the mean spike density was typically higher than for passive sound exposure (Fig. 6A, epoch 5; mean firing rate gain for three cells: A500/P, 1.30; A1000/P, 1.31; A1000/A500, 1.10). However, the increase in firing rate was not merely due to a simple magnitude scaling of the passive response, as the temporal patterns also differed substantially. This resulted in poor correlations between the spike-density functions of these dynamic epochs (Fig. 6B, cf. Fig. 5B, center).

The observed changes in the spike trains during the acoustic detection task suggest an influence of non-acoustic (top-down) factors. To fully quantify this influence on the neuronal responses for each recorded cell and for every response epoch, we analysed the temporal firing patterns, and the acoustic spectrotemporal tuning characteristics as determined by the STRFs for the Passive, A500 and A1000 trials.

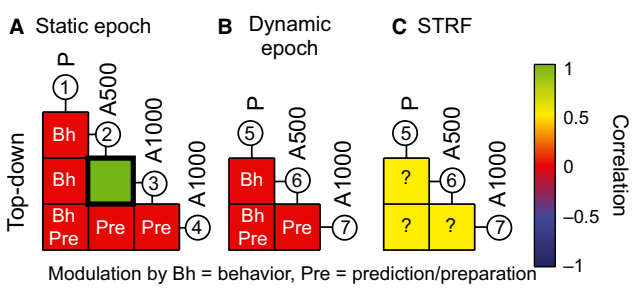

FIG. 5. Rationale of the experiments. If top-down signals modulate AC activity, correlations (A and B) between corresponding activity patterns are reduced for all but two epochs: the first 500 ms of static noise of A500 (2) and A1000 (3). In these epochs the animal attends to the sound, but is equally uncertain about the upcoming change. Bh and Pre indicate potential top-down signals in the active epochs: Bh, signals related to non-specific behavior (e.g. vigilance, attention); Pre, modulations specific to change prediction, reward prediction or motor preparation. Whether and how STRFs would change (C) due to top-down modulations is not known *a priori*. Expected correlation values are indicated by the color bar.

We first describe this analysis for example neuron J67 (Fig. 7). The spike rasters for the 55 ripples and the average spike-density functions for the Passive (Fig. 7A, left), A500 (Fig. 7A, center) and A1000 (Fig. 7A, right) trials demonstrate that task participation had a considerable influence on the neural response patterns during the entire trial. During passive exposure the neuron exhibited low baseline activity during silence ($t < 0$ ms), followed by a strong phasic response (about 100 spikes/s), starting at ~17 ms after sound onset (at $t = 0$) and some additional response peaks during static-noise epoch 1. In contrast, during the silent period of the two active listening trials the neuron ramped up its firing rate before sound onset. This prelude activity was followed by a high phasic response (about 100 spikes/s), followed by sustained activity during the remainder of static epochs 2–4.

Passive exposure during dynamic epoch 5 ($t = 500$–1500 ms) showed a fine temporal structure in the spike rasters that was locked to the ripples' repetition rates (1/velocity). At sound offset ($t = 1500$ ms), activity quickly dropped back to baseline. For the two active-listening trials (dynamic epochs 6 and 7) the neuron produced a much stronger response than in the passive condition, and after ripple offset (at 1500 and 2000 ms, respectively) it persisted with sustained firing.

The correlations comparing the firing patterns of the various epochs were low or insignificant (Fig. 7D, left and center, cf. Fig. 5), except when comparing the first 500 ms of static noise for the active-listening trials (epochs 2 and 3). The high correlation between epochs 2 and 3 ($r = 0.93$) implies that the changes in response firing rates in all epochs were specific to the same acoustical, mental, and behavioral state of the monkey, and were not simply due to random response variations.

Next, we derived the cell's STRF (Fig. 7C; see Materials and methods, Fig. 3) from the cell's first-order phase and magnitude responses in the dynamic epochs (Eqns 3 and 4). Interestingly, the phase of the cell's ripple transfer function (Fig. 7B) was very similar for the three conditions ($r = 0.97$–0.98). This indicates that the cell phase-locked to the acoustic modulations of the ripple envelope in a similar fashion for all three conditions, despite the substantial differences in spike rates. The magnitude transfer functions differed to some extent ($r = 0.50$–0.57), because in the active-listening paradigms the magnitude transfer functions contained stronger responses (red colors) for several, but not all, ripples than when the monkey passively heard these stimuli (Fig. 7B; mean/max magnitude transfer function: Passive, 10/17 spikes/s; A500, 24/42 spikes/s; A1000, 22/41 spikes/s).

As the magnitude transfer functions varied across conditions, the resulting STRFs (Fig. 7C) differed slightly in their peak and trough magnitudes for this cell (max/min STRF: Passive, 7/−4 spikes/s; A500, 15/−7 spikes/s; A1000, 14/−7 spikes/s; dark-red/blue colors, respectively, Fig. 7C). However, as a result of the robust phase-locking behavior, the patterns of excitation and inhibition in the STRFs were very similar ($r = 0.85$–0.90 for the three comparisons; Fig. 7D, right).

Preservation of acoustic tuning

Similarly, the cells shown in Fig. 6 preserved their acoustic encoding features during active listening (Fig. 8A). Their STRFs were virtually unaffected by task performance, despite the marked changes in their firing behavior (Fig. 5). The cells showed a clear excitation for a

© 2013 Federation of European Neuroscience Societies and Blackwell Publishing Ltd

14609568, 2013, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ejn.12180 by Radboud University Nijmegen, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

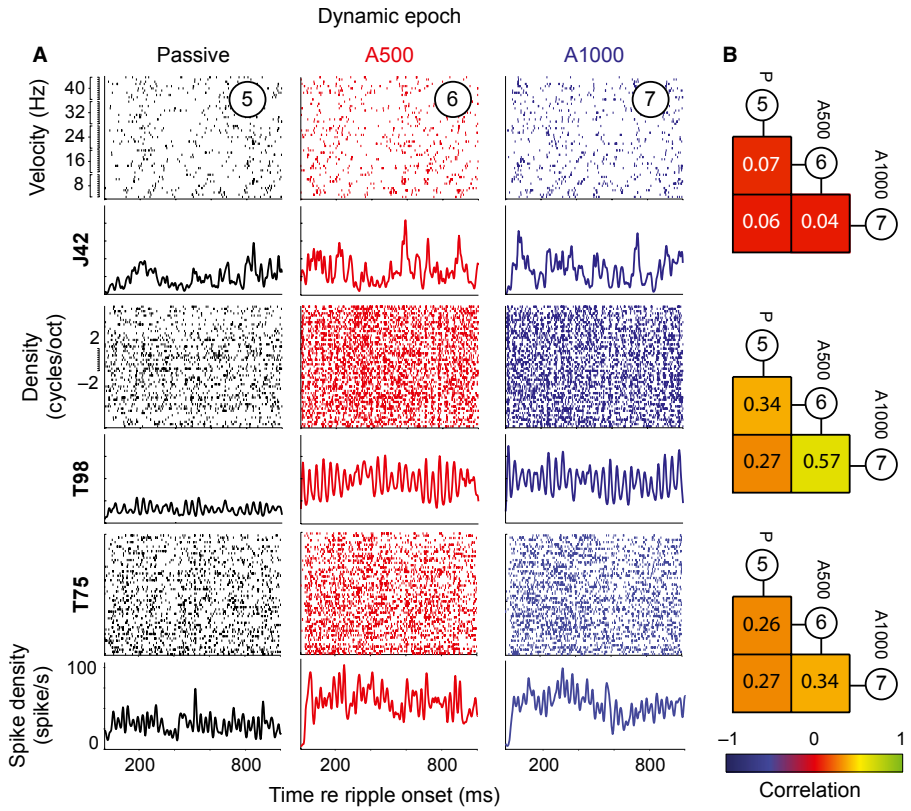

FIG. 6. (A) Spike-raster plots, sorted according to, first, ascending ripple velocity, and then ascending ripple density, and trial-averaged spike density functions for three representative cells (J42, T98 and T75) during dynamic ripples (duration 1000 ms) in the passive (epoch 5, black), A500 (epoch 6, red) and A1000 (epoch 7, blue) trials. Each dot in the raster plot represents one spike. Small tick marks at the top left-hand panel correspond to the applied ripple densities, the values of which are provided at the central left-hand panel. (B) Correlations (color-coded; numbers in boxes) between the spike-density functions of epochs 5–7.

narrow frequency range, with latencies between 10 and 40 ms, typically surrounded by a complex of spectral–temporal inhibitory sidebands. The excitatory (red) and inhibitory (blue) segments were located at exactly the same time-frequency bins for the three listening conditions. Also, the other parts of the STRF remained unaffected, yielding high mutual STRF correlations for these cells (Fig. 8B).

Despite the strong resemblance of the STRFs for the different hearing conditions, subtle but systematic changes might potentially be missed by the correlation analysis. To check for any such consistent spectral–temporal changes in the active STRFs in comparison to the passive ones, we applied the linear regression analysis of Eqns 5–8 (see Materials and methods). Figure 9A represents the average passive-evoked STRF of all 51 cells, showing an excitatory region followed by inhibition. Figure 9B–D shows the average normalised difference STRFs (Eqn 8) between A500-P, A1000-P, and A1000-A500 trials, respectively. Systematic patterns in the difference STRFs are absent, from which we conclude that there were no discernible shape changes in the STRFs during active listening. The distribution of the gain changes computed for all cells was unimodal and on average about 1 (mean ± SD of the gains: A500/P, 1.11 ± 0.85; A1000/P, 1.06 ± 0.71; A1000/A500, 0.81 ± 0.33). The bias values were around zero (mean ± SD of biases: A500/P, 0.09 ± 0.37; A1000/P, 0.15 ± 0.46; A1000/A500, 0.11 ± 0.37).

We also applied the population regression analysis to the 54 cells with low SNR, which is shown in Fig. 9E–H. Again, no systematic changes were observed for the STRFs obtained in active trials. The results for cells yielding low SNRs and high SNRs were therefore quite similar.

For the vast majority of cells (Fig. 10C), the STRF correlations between the three behavioral conditions were high (median $r = 0.74$–0.81). Low correlation values, as found for a few cells, could be ascribed to a relatively low signal-to-noise ratio. Thus, in spite of the large differences in firing behavior, as demonstrated by the low correlations between the spike-density functions during the dynamic epochs (Fig. 10B), and higher responses in the active condition (mean ± SD of firing rate gain for all 51 cells: A500/P, 1.85 ± 0.93; A1000/P, 2.05 ± 1.14; A1000/A500, 1.21 ± 0.2), AC cells faithfully preserved their bottom-up acoustic-tuning characteristics.

During the static-noise epochs (1–4) the stimulus acoustics were the same, but the behavioral states of the animal differed (except for epochs 2 and 3). As a result, the correlations between the mean spike densities for passive and active listening in these epochs were very low for five out of six comparisons (Fig. 10A). In line with the rationale of our paradigm (Fig. 5A), evoked responses for the large majority of neurons during the first 500 ms of A500 (epoch 2) and A1000 trials (epoch 3) were highly correlated (median $r = 0.81$; Fig. 10A cf. Fig. 7D, left). The high correlations between epochs 2 and 3 demonstrate that the activity changes across the other epochs did not result from random noise. Thus, the decorrelation of acoustically identical, yet behaviorally different epochs indicates involvement of a top-down signal.

Top-down signals are time-locked to behavior

During the dynamic epoch, neural responses for passive sound exposure and the two active trial types differed profoundly, as demon-

© 2013 Federation of European Neuroscience Societies and Blackwell Publishing Ltd

14609568, 2013, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ejn.12180 by Radboud University Nijmegen, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

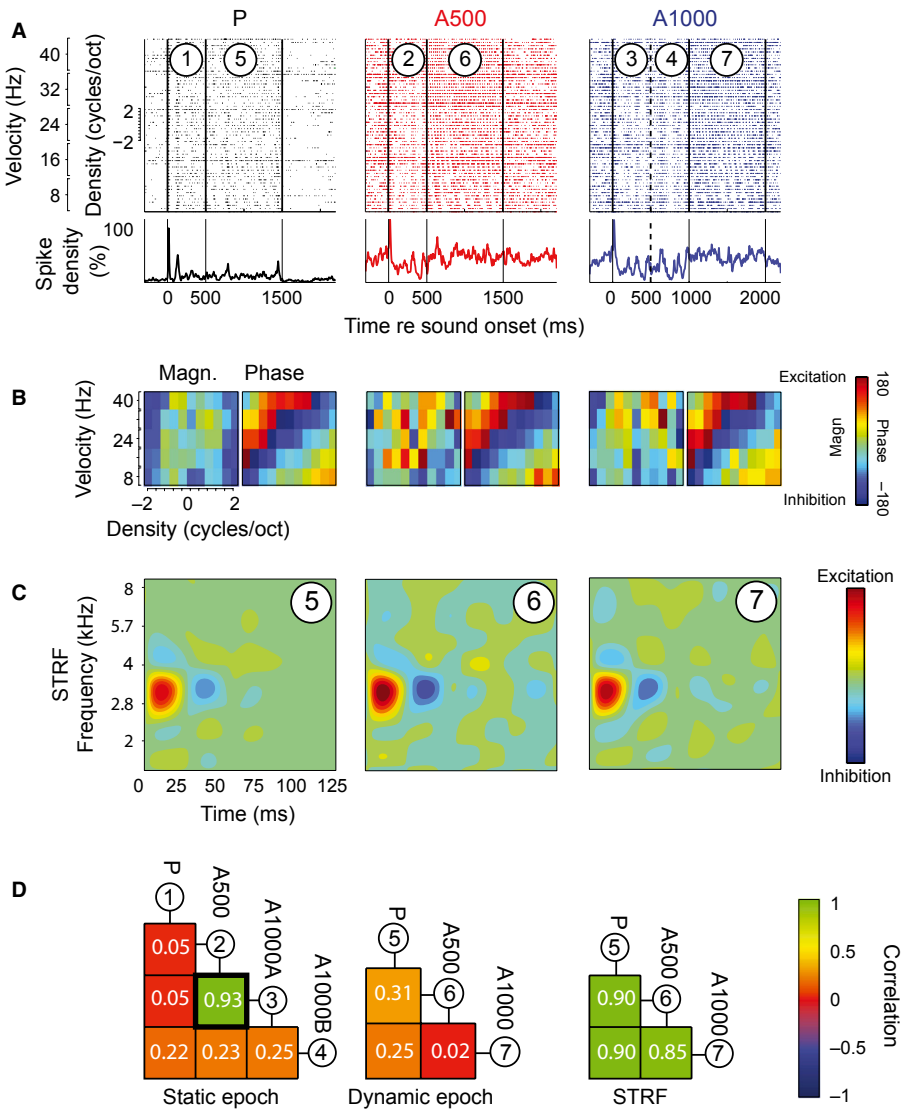

FIG. 7. Full quantitative analysis of example cell J67. (A) Raster plots (top) and spike density functions (bottom) for the three listening conditions during the trial. Vertical lines denote the different recording epochs (cf. Fig. 1). The trials are sorted according to, first, ascending ripple velocities, and second, ascending ripple densities (see tick marks). Averaged spike-density functions are normalised to the peak passive response (set to 100%). Note higher activity in active trials than in passive trials (mean A500/*P* = 5.7, A1000/*P* = 5.4, A1000/A500 = 0.94). (B) Magnitude (left) and phase (right) ripple transfer functions and corresponding STRFs (C) for the three listening conditions. The similar phase transfer functions in different conditions indicate robust phase locking to ripple envelope modulations. (D) Correlations between spike density functions for static (left) and dynamic (center) epochs, and STRFs (right). A1000A denotes the first 500 ms of the static noise in A1000 trials, and corresponds to epoch 3 in Fig. 1; A1000B represents the second 500 ms of the static noise (epoch 4 in Fig. 1).

strated by a low correlation between the average spike-density functions (Fig. 10B). Upon detecting the ripple in the active listening task, the monkey prepares a manual response as fast as possible to obtain a reward. Therefore, an important factor in the active-listening paradigm, so far ignored in the analysis, is the considerable inter-trial variability of reaction times, within and between the two active trials (Figs 1 and 4). This variability could potentially conceal any systematic motor-related or predictive component in the cell's firing rate. Here we test whether this component was present in the activity of AC cells by realigning the spike rasters to the animal's reaction time.

The resulting average spike-density functions of example cell J67 (same cell as in Fig. 7) show that the correlations between the dynamic A500 and A1000 trials increased markedly from 0.13, when aligned to ripple onset (Fig. 11A), to 0.64, when aligned to reaction time (Fig. 11B). Indeed, this correlation improved for 47 of 51 cells ($T_{50} = -9.5$, $P < 10^{-6}$) (Fig. 11E).

This result also generalises to the population averages of the spike-density functions, which show a marked improvement in correlation when the AC responses are aligned to reaction time (Fig. 11D) vs. ripple onset (Fig. 11C). Note that the population average started to increase well before (approximately 400 ms) bar release, thus representing a top-down signal that could be potentially related to change prediction, motor preparation, or to the prediction of the upcoming reward. The present paradigm cannot dissociate these different factors, as reward delivery immediately followed bar release.

In summary, the absence of a correlation between the neural activity patterns during active listening (Fig. 10B) was not due to random noise in the spike timings. Instead, the changes in firing rate

© 2013 Federation of European Neuroscience Societies and Blackwell Publishing Ltd

14609568, 2013, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ejn.12180 by Radboud University Nijmegen, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

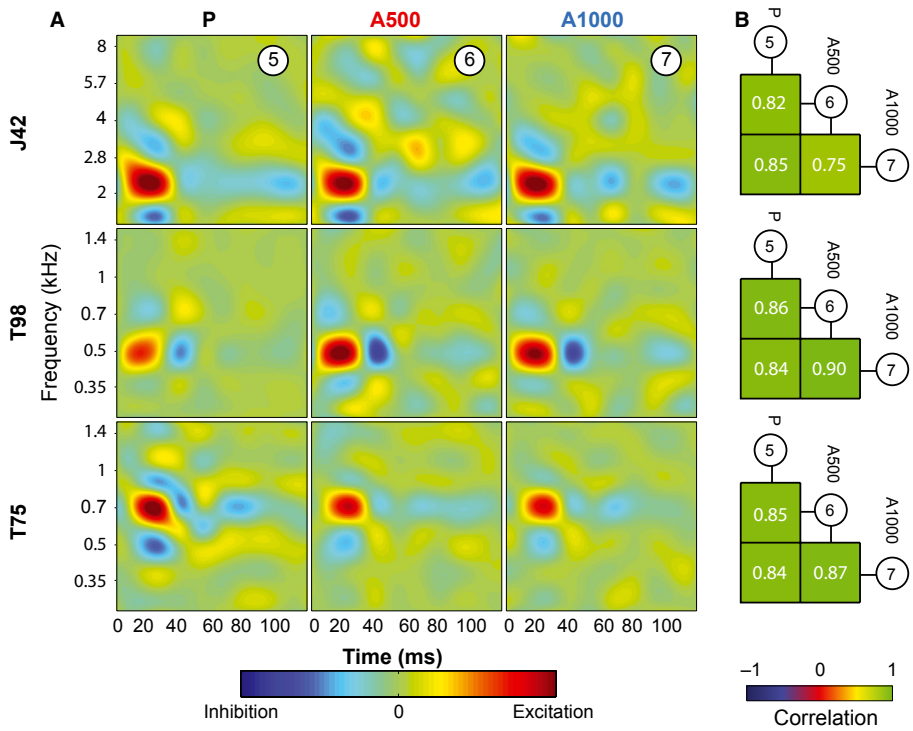

FIG. 8. (A) STRFs of the three cells shown in Fig. 6 for the passive (epoch 5), A500 (epoch 6) and A1000 (epoch 7) trials (see Fig. 1). (B) Correlation coefficients for the STRFs for each pair of epochs. The STRFs for the three listening conditions are very similar for each of the three cells.

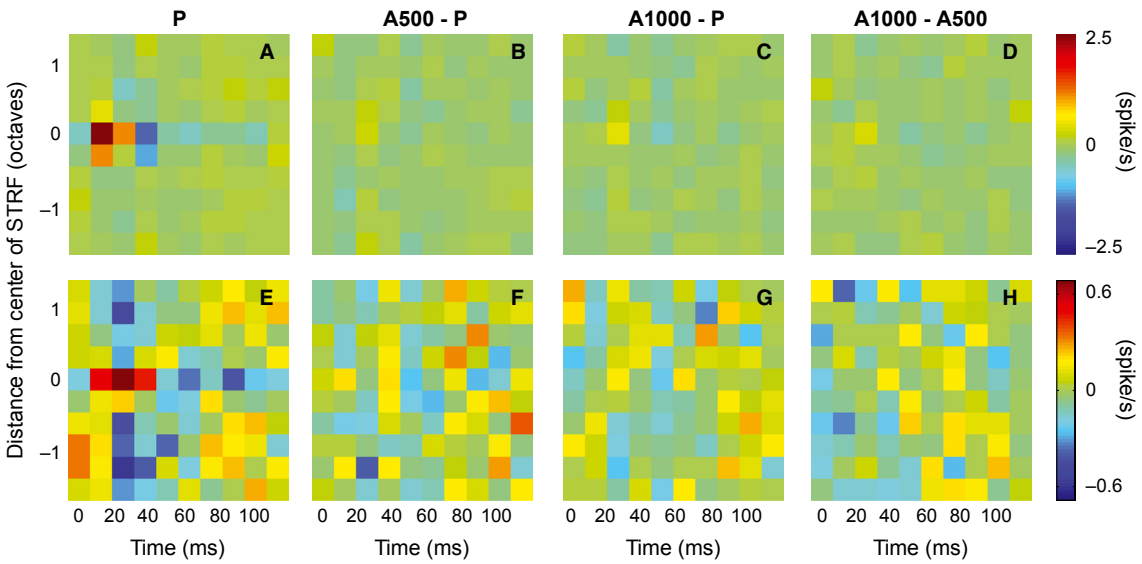

FIG. 9. STRF shape-change analysis for high-SNR cells (A–D) and low-SNR cells (E–H). (A) Shifted passive STRF, centered at the spectral octave range and averaged across all high-SNR cells (n = 51). (B) Normalised average difference STRF (Eqn 8) between passive and A500 listening. Note absence of a systematic response, indicating no systematic change in the STRF of the neurons. (C) Normalised difference between passive and A1000 listening. (D) Normalised difference between A500 and A1000 listening. Color in A-C is scaled to the averaged passive STRF. Color in D is scaled to the averaged A500 STRF. (E–H) Results for the low-SNR cells are qualitatively similar to the high-SNR data shown in A–D. Note differences in scale (color bars) for low- and high-SNR cells, which roughly correspond to the average firing rates for the selected populations.

prior to bar release in both animals restored the neural response correlations, and can therefore be attributed to top-down signals related to movement preparation and/or reward prediction signals.

Discussion

This study is the first to quantify robust spectrotemporal tuning of AC neurons despite strong task-related modulations in their firing rates. Our results show that AC responses differed systematically between active and passive listening conditions, demonstrating clear non-acoustic top-down processes in monkey AC.

The data indicate that, during passive exposure, AC responses were primarily driven by bottom-up, stimulus-driven, acoustic signals, resulting in low pre- and post-stimulus background activity, short onset responses (e.g. Fig. 7A), and often strong phase locking to the ripple envelope (e.g. Fig. 7B). In the active listening

© 2013 Federation of European Neuroscience Societies and Blackwell Publishing Ltd

14609568, 2013, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ejn.12180 by Radboud University Nijmegen, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

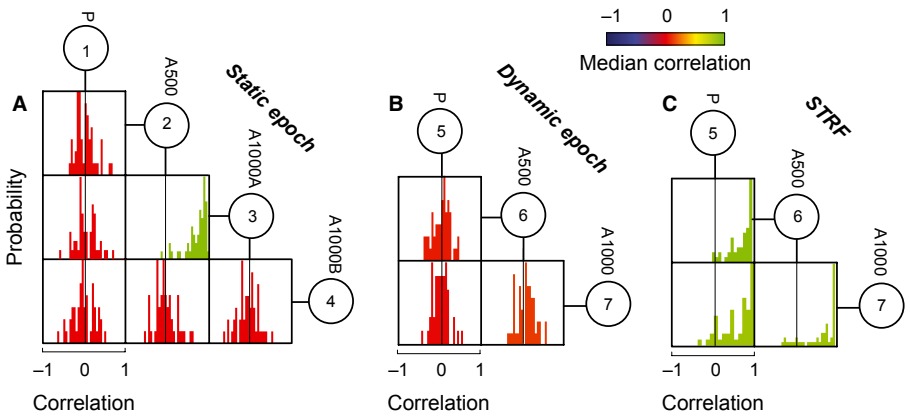

FIG. 10. Population correlation analysis. (A) Correlation coefficients between the spike density functions of the static epochs (1–4) for the three listening conditions of all 51 neurons, color coded according to the median value (color bar). (B) Correlation coefficients between the spike density functions of the dynamic epochs (5–7). (C) Correlations between the STRFs obtained from the three behavioral situations. Correlations are high only for the STRF comparisons and the epoch 2–3 comparison.

condition, the first-order phase locking remained unaltered (e.g. Fig. 7B). As a consequence, the shapes of the STRFs were virtually identical for the active and passive conditions (Figs 7C, 8 and 10C), as there was no systematic shift and/or broadening in the active STRF when compared with the passive one (Fig. 9).

In contrast, the firing patterns of the different dynamic stimulus epochs became decorrelated (Figs 6, 7A and 10B). These differences did not simply reflect random noise, e.g. sounds produced by animal movement, licking behavior, or changes in the animal's vigilance. First, the modulations were highly systematic when responses were aligned to the reaction time (e.g. Fig. 11); and second, in line with the rationale of our paradigms (Fig. 5), although the static noise-evoked firing patterns for passive and active conditions were uncorrelated, spike-density functions for the two active-listening trials during the first 500 ms epochs (2 and 3) were very similar (Figs 7D and 10A). In this situation, both the stimulus acoustics (randomly generated noise) and the monkey's behavioral state (engagement in, and attention to the task and the sound, and equal uncertainty about the upcoming ripple) were identical. Although the acoustics did not change in the second half (epoch 4) of the A1000 static noise either, neural responses did change. We attribute this to differences in the monkey's predictive state that led to decreased reaction times for A1000 trials (Figs 1 and 4). A comparison of each ripple's reaction time during A500 and A1000 trials showed that the monkey's decisions were not triggered by the same acoustic features (Fig. 4). In A500 trials, the animal's reaction was based on both spectral and temporal modulations (Fig. 4A), whereas in A1000 trials change detection was predominantly based on prediction, rather than on acoustic ripple modulations (Fig. 4B). This predictive state led to apparently different spike-density functions for the A500 and A1000 dynamic epochs (Figs 6 and 10B). The correlation patterns shown in Fig. 10 did not change when trials with negative reaction times in the A1000 trials were excluded from the analysis (not shown). Thus, the low correlations between static-noise epochs 3 and 4 did not result from a signal related to motor execution at variable reaction times during the static noise in epoch 4. Realigning the neural responses to reaction time, however, revealed the same preparatory signal and a concomitant increase in response correlation for the A500 and A1000 trials in both animals (Fig. 11B and D). A similar preparatory signal was reported for monkey AC by Brosch *et al.* (2005) and Niwa *et al.* (2012b), and for rat AC by Jaramillo & Zador (2011).

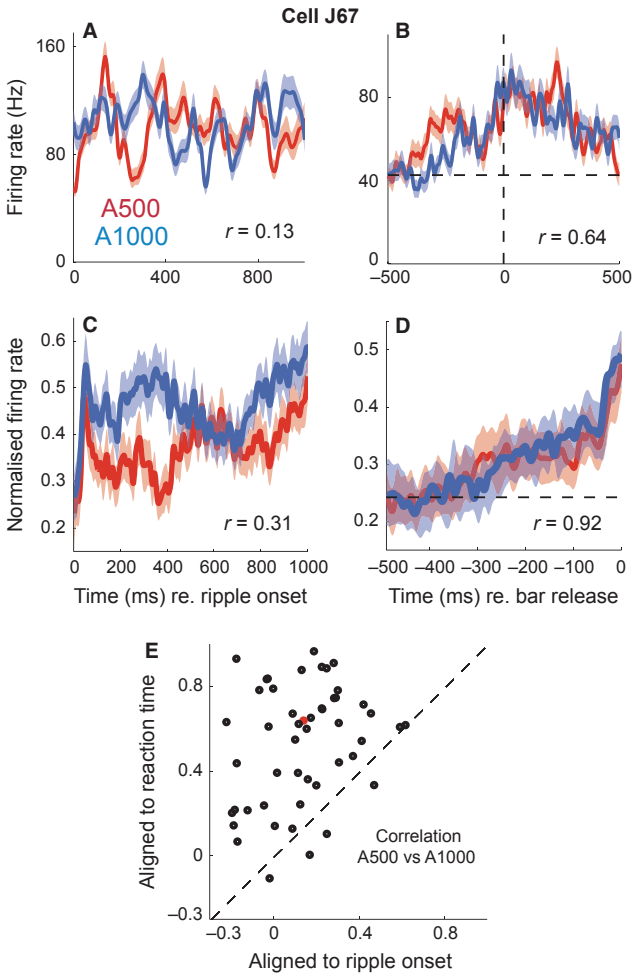

FIG. 11. Top-down signals. (A and C) Average spike density functions for all A500 (red) and A1000 (blue) trials aligned to ripple onset, and (B and D) aligned to bar release for (A and B) cell J67 and (C and D) the population averages. Mean and SE are denoted by solid lines and transparent patches, respectively. The firing activity was normalised by dividing the firing rates by the response range (maximum–minimum) for each cell. (E) Correlation coefficients between spike density functions of dynamic epochs 6 and 7 of A500 and A1000 trials, when aligned to reaction time (ordinate) vs. when aligned to ripple onset (abscissa). The red dot corresponds to example cell J67.

© 2013 Federation of European Neuroscience Societies and Blackwell Publishing Ltd

14609568, 2013, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ejn.12180 by Radboud University Nijmegen, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Our sound-detection paradigm identified several top-down modulations in the active responses of AC neurons. Typically, neural baseline activity was higher for active listening during the entire trial (Figs 6A and 7A) and persisted even during the silent epochs, i.e. before sound onset and after sound offset (Fig. 7A), which was also demonstrated recently by Niwa *et al.* (2012a). This was also apparent from the slight increase in the magnitude transfer characteristics (e.g. Fig. 7B). This overall increase could be due to a fixed, static attentional signal that was present only when the monkeys were actively engaged in an auditory task. Most studies have reported that attentional demands, including arousal and vigilance, increase cortical responses (Edeline *et al.*, 2001; Steriade *et al.*, 2001; Niwa *et al.*, 2012a). Our results are consistent with these findings [see, however, Otazu *et al.* (2009) who reported a decreasing AC response due to task engagement in rats]. Other top-down modulations were highly dynamic, changing on a millisecond time scale from a task-preparation, to a stimulus-onset prediction signal that started before sound onset (Fig. 7A) (in line with Selezneva *et al.*, 2006), and finally to a motor-preparation or reward-prediction signal during the ripple (Fig. 11B and D). These results confirm and extend previous reports on response enhancement in AC neurons prior to motor action (Brosch *et al.*, 2011b; Niwa *et al.*, 2012b), and in relation to reward expectation (Brosch *et al.*, 2011a). Given the different interpretations of the non-acoustic modulations in AC cells, we conjecture that they arose from different cortical sources that could involve auditory belt and parabelt regions (acoustic task) (Durif *et al.*, 2003; Brosch *et al.*, 2005), and parietal and frontal cortical areas (prediction, reward) (Romanski *et al.*, 1999; Fritz *et al.*, 2010), as well as the pre-motor cortex (motor preparation) (Lemus *et al.*, 2009).

Although we demonstrated no changes in STRFs in task-related listening conditions (Figs 7C, 8, 9 and 10C), numerous studies have shown changes in STRFs resulting from various behavioral conditions. A tone-detection task modified STRFs of A1 in trained ferrets by enhancing neural responses at the attended target frequency (Fritz *et al.*, 2003). Moreover, tone discrimination and variation of signal-to-noise ratio in these tasks could both induce changes in the STRF by selectively enhancing sensitivity to the target frequency, and selectively suppressing tuning to the reference frequency (Fritz *et al.*, 2005a,b; Atiani *et al.*, 2009). Consistent focused attention to a particular acoustic feature (the target tone) may have been the key trigger to initiate the receptive field changes (Fritz *et al.*, 2007). Factors other than attention could also induce receptive field plasticity in A1, such as learning (Ohl & Scheich, 1997, 2005; Kilgard *et al.*, 2001, 2002; Kilgard & Merzenich, 2002), conditioning (Bakin & Weinberger, 1996; Ji *et al.*, 2001), expectation (Jaramillo & Zador, 2011), reward (David *et al.*, 2012), sound localisation (Lee & Middlebrooks, 2011), and electrical microstimulation (Suga *et al.*, 2002; Suga & Ma, 2003). These findings suggest that a variety of non-acoustic factors can affect spectrotemporal tuning of A1 cells. However, if auditory task performance by itself were to modify STRFs, the AC would seem ill-suited to encode the acoustic environment. In our experiments monkeys did not attend to, or were learning, a specific acoustic feature of any given target ripple, as all ripples were equally likely and important. Instead, the animals were merely awaiting (in A500; Fig. 4A) or predicting (in A1000; Fig. 4B) their appearance. Although, after the change detection, animals were not required to listen to the ripples any more, AC cells still followed their spectral–temporal dynamics with high fidelity. This suggests that AC plasticity may not be linked to behavioral relevance *per se*, but rather requires behavioral relevance of specific acoustic features.

Model studies have suggested that multiplexing different signals at the population level of narrowly tuned cells allows the target structures of the population to selectively extract the relevant variable(s) for further processing (Pouget & Sejnowski, 1995; Van Opstal & Hepp, 1995). Such a mechanism provides an efficient and flexible encoding scheme to deal with multiple task constraints and representations. Multiplicative encoding ('gain fields') has been reported for various sensorimotor processing stages in the monkey brain. For example, changes in static eye position gain-modulate visual responses in the lateral intraparietal sulcus without affecting visual receptive fields (Andersen *et al.*, 1985). Similarly, multiplicative eye-position modulations were reported for saccadic eye-movements responses in the midbrain superior colliculus (Van Opstal *et al.*, 1995), and for the acoustic responses of a subpopulation of cells in the inferior colliculus (Groh *et al.*, 2001; Werner-Reiss *et al.*, 2003; Zwiers *et al.*, 2004). A recent study also indicated a multiplicative mechanism for bottom-up encoding of sound attributes in A1 cells of anesthetised ferrets (Walker *et al.*, 2011). Although attentional signals have been shown to gain-modulate visual responses in primate (McAdams & Reid, 2005) and cat (Kara *et al.*, 2002) V1, to our knowledge, a multiplicative interaction of acoustic and behavioral signals in the AC has not been reported before. Our results imply that primate AC neurons can efficiently represent rapidly changing bottom-up and top-down signals in a multiplexed fashion. This allows for a stable representation and perceptual invariance of acoustic signals in the presence of strong non-acoustic top-down modulations.

Acknowledgements

We are grateful to Hans Kleijnen, Dick Heeren, Günter Windau, and Stijn Martens for their valuable technical assistance and to Alex Hanssen and the central animal facility for excellent monkey care. This research was supported by a VICI grant of the Earth and Life Sciences of NWO (ALW 865.05.003; A.J.V.O. and R.M.) and the Radboud University Nijmegen (A.J.V.O., R.M., S.M.C.I.V.W., M.M.V.W. and H.V.).

Abbreviations

A1, primary auditory cortex; A500, active500 trials; A1000, active1000 trials; AC, auditory cortex; SNR, signal-to-noise ratio; STRF, spectrotemporal receptive field.

Author contributions

R.M. and S.M.C.I.V.W. performed the experiments. R.M. and M.M.V.W. analysed the data. H.V. and A.J.V.O. designed the paradigms. R.M., M.M.V.W., H.V. and A.J.V.O. co-wrote the manuscript.

References

Aertsen, A.M., Olders, J.H. & Johannesma, P.I. (1981) Spectro-temporal receptive fields of auditory neurons in the grassfrog. III. Analysis of the stimulus-event relation for natural stimuli. *Biol. Cybern.*, **39**, 195–209.

Andersen, R.A., Essick, G.K. & Siegel, R.M. (1985) Encoding of spatial location by posterior parietal neurons. *Science*, **230**, 456–458.

Atiani, S., Elhilali, M., David, S.V., Fritz, J.B. & Shamma, S.A. (2009) Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron*, **61**, 467–480.

Bakin, J.S. & Weinberger, N.M. (1996) Induction of a physiological memory in the cerebral cortex by stimulation of the nucleus basalis. *Proc. Natl. Acad. Sci. USA*, **93**, 11219–11224.

Brosch, M., Selezneva, E. & Scheich, H. (2005) Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J. Neurosci.*, **25**, 6797–6806.

Brosch, M., Selezneva, E. & Scheich, H. (2011a) Representation of reward feedback in primate auditory cortex. *Front. Syst. Neurosci.*, **5**, 5.

Brosch, M., Selezneva, E. & Scheich, H. (2011b) Formation of associations in auditory cortex by slow changes of tonic firing. *Hearing Res.*, **271**, 66–73.

© 2013 Federation of European Neuroscience Societies and Blackwell Publishing Ltd

14609568, 2013, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ejn.12180 by Radboud University Nijmegen, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

David, S.V., Fritz, J.B. & Shamma, S.A. (2012) Task reward structure shapes rapid receptive field plasticity in auditory cortex. *Proc. Natl. Acad. Sci. USA*, **109**, 2144–2149.

Depireux, D.A., Simon, J.Z., Klein, D.J. & Shamma, S.A. (2001) Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. *J. Neurophysiol.*, **85**, 1220–1234.

Durif, C., Jouffrais, C. & Rouiller, E.M. (2003) Single-unit responses in the auditory cortex of monkeys performing a conditional acousticomotor task. *Exp. Brain Res.*, **153**, 614–627.

Edeline, J.M. (2003) The thalamo-cortical auditory receptive fields: regulation by the states of vigilance, learning and the neuromodulatory systems. *Exp. Brain Res.*, **153**, 554–572.

Edeline, J.M., Dutrieux, G., Manunta, Y. & Hennevin, E. (2001) Diversity of receptive field changes in auditory cortex during natural sleep. *Eur. J. Neurosci.*, **14**, 1865–1880.

Eggermont, J.J. (2011) Context dependence of spectro-temporal receptive fields with implications for neural coding. *Hearing Res.*, **271**, 123–132.

Elhilali, M., Fritz, J.B., Chi, T.S. & Shamma, S.A. (2007) Auditory cortical receptive fields: stable entities with plastic abilities. *J. Neurosci.*, **27**, 10372–10382.

Fritz, J., Shamma, S., Elhilali, M. & Klein, D. (2003) Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat. Neurosci.*, **6**, 1216–1223.

Fritz, J., Elhilali, M. & Shamma, S. (2005a) Active listening: task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex. *Hearing Res.*, **206**, 159–176.

Fritz, J.B., Elhilali, M. & Shamma, S.A. (2005b) Differential dynamic plasticity of A1 receptive fields during multiple spectral tasks. *J. Neurosci.*, **25**, 7623–7635.

Fritz, J.B., Elhilali, M. & Shamma, S.A. (2007) Adaptive changes in cortical receptive fields induced by attention to complex sounds. *J. Neurophysiol.*, **98**, 2337–2346.

Fritz, J.B., David, S.V., Radtke-Schuller, S., Yin, P. & Shamma, S.A. (2010) Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. *Nat. Neurosci.*, **13**, 1011–1019.

Fu, K.M., Johnston, T.A., Shah, A.S., Arnold, L., Smiley, J., Hackett, T.A., Garraghty, P.E. & Schroeder, C.E. (2003) Auditory cortical neurons respond to somatosensory stimulation. *J. Neurosci.*, **23**, 7510–7515.

Ghazanfar, A.A., Chandrasekaran, C. & Logothetis, N.K. (2008) Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. *J. Neurosci.*, **28**, 4457–4469.

Griffiths, T.D., Warren, J.D., Scott, S.K., Nelken, I. & King, A.J. (2004) Cortical processing of complex sound: a way forward? *Trends Neurosci.*, **27**, 181–185.

Groh, J.M., Trause, A.S., Underhill, A.M., Clark, K.R. & Inati, S. (2001) Eye position influences auditory responses in primate inferior colliculus. *Neuron*, **29**, 509–518.

Gutschalk, A., Micheyl, C., Melcher, J.R., Rupp, A., Scherg, M. & Oxenham, A.J. (2005) Neuromagnetic correlates of streaming in human auditory cortex. *J. Neurosci.*, **25**, 5382–5388.

Jaramillo, S. & Zador, A.M. (2011) The auditory cortex mediates the perceptual effects of acoustic temporal expectation. *Nat. Neurosci.*, **14**, 246–251.

Ji, W., Gao, E. & Suga, N. (2001) Effects of acetylcholine and atropine on plasticity of central auditory neurons caused by conditioning in bats. *J. Neurophysiol.*, **86**, 211–225.

Kaas, J.H. & Hackett, T.A. (2000) Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl. Acad. Sci. USA*, **97**, 11793–11799.

Kara, P., Pezaris, J.S., Yurgenson, S. & Reid, R.C. (2002) The spatial receptive field of thalamic inputs to single cortical simple cells revealed by the interaction of visual and electrical stimulation. *Proc. Natl. Acad. Sci. USA*, **99**, 16261–16266.

Kayser, C., Petkov, C.I. & Logothetis, N.K. (2008) Visual modulation of neurons in auditory cortex. *Cereb. Cortex*, **18**, 1560–1574.

Kayser, C., Logothetis, N.K. & Panzeri, S. (2010) Visual enhancement of the information representation in auditory cortex. *Curr. Biol.*, **20**, 19–24.

Kilgard, M.P. & Merzenich, M.M. (2002) Order-sensitive plasticity in adult primary auditory cortex. *Proc. Natl. Acad. Sci. USA*, **99**, 3205–3209.

Kilgard, M.P., Pandya, P.K., Vazquez, J., Gehi, A., Schreiner, C.E. & Merzenich, M.M. (2001) Sensory input directs spatial and temporal plasticity in primary auditory cortex. *J. Neurophysiol.*, **86**, 326–338.

Kilgard, M.P., Pandya, P.K., Engineer, N.D. & Moucha, R. (2002) Cortical network reorganization guided by sensory input features. *Biol. Cybern.*, **87**, 333–343.

King, A.J., Bajo, V.M., Bizley, J.K., Campbell, R.A., Nodal, F.R., Schulz, A.L. & Schnupp, J.W. (2007) Physiological and behavioral studies of spatial coding in the auditory cortex. *Hearing Res.*, **229**, 106–115.

Lakatos, P., Chen, C.M., O'Connell, M.N., Mills, A. & Schroeder, C.E. (2007) Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron*, **53**, 279–292.

Lee, C.C. & Middlebrooks, J.C. (2011) Auditory cortex spatial sensitivity sharpens during task performance. *Nat. Neurosci.*, **14**, 108–114.

Lemus, L., Hernandez, A. & Romo, R. (2009) Neural encoding of auditory discrimination in ventral premotor cortex. *Proc. Natl. Acad. Sci. USA*, **106**, 14640–14645.

McAdams, C.J. & Reid, R.C. (2005) Attention modulates the responses of simple cells in monkey primary visual cortex. *J. Neurosci.*, **25**, 11023–11033.

Nelken, I. & Bar-Yosef, O. (2008) Neurons and objects: the case of auditory cortex. *Front. Neurosci.*, **2**, 107–113.

Nelken, I., Fishbach, A., Las, L., Ulanovsky, N. & Farkas, D. (2003) Primary auditory cortex of cats: feature detection or something else? *Biol. Cybern.*, **89**, 397–406.

Niwa, M., Johnson, J.S., O'Connor, K.N. & Sutter, M.L. (2012a) Active engagement improves primary auditory cortical neurons' ability to discriminate temporal modulation. *J. Neurosci.*, **32**, 9323–9334.

Niwa, M., Johnson, J.S., O'Connor, K.N. & Sutter, M.L. (2012b) Activity related to perceptual judgment and action in primary auditory cortex. *J. Neurosci.*, **32**, 3193–3210.

Ohl, F.W. & Scheich, H. (1997) Learning-induced dynamic receptive field changes in primary auditory cortex of the unanaesthetized Mongolian gerbil. *J. Comp. Physiol. A.*, **181**, 685–696.

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

Otazu, G.H., Tai, L.H., Yang, Y. & Zador, A.M. (2009) Engaging in an auditory task suppresses responses in auditory cortex. *Nat. Neurosci.*, **12**, 646–654.

Palmer, A.R. (2007) Anatomy and physiology of the auditory brainstem. In Burkard, R.F., Eggermont, J.J. & Don, M. (Eds), *Auditory Evoked Potentials: basic Principles and Clinical Application*. Lippincott Williams & Wilkins, Baltimore, pp. 200–228.

Paxinos, G., Huang, X.F. & Toga, A.W. (2000) *The Rhesus Monkey Brain in Stereotaxic Coordinates*. Academic Press, San Diego, CA.

Polley, D.B., Steinberg, E.E. & Merzenich, M.M. (2006) Perceptual learning directs auditory cortical map reorganization through top-down influences. *J. Neurosci.*, **26**, 4970–4982.

Pouget, A. & Sejnowski, T.J. (1995) Spatial representations in the parietal cortex may use basis functions. *NIPS Conf. Proc.*, **7**, 157–164.

Recanzone, G.H., Guard, D.C. & Phan, M.L. (2000) Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. *J. Neurophysiol.*, **83**, 2315–2331.

Riecke, L., van Opstal, A.J., Goebel, R. & Formisano, E. (2007) Hearing illusory sounds in noise: sensory-perceptual transformations in primary auditory cortex. *J. Neurosci.*, **27**, 12684–12689.

Romanski, L.M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P.S. & Rauschecker, J.P. (1999) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci.*, **2**, 1131–1136.

Schroeder, C.E. & Foxe, J. (2005) Multisensory contributions to low-level, 'unisensory' processing. *Curr. Opin. Neurobiol.*, **15**, 454–458.

Selezneva, E., Scheich, H. & Brosch, M. (2006) Dual time scales for categorical decision making in auditory cortex. *Curr. Biol.*, **16**, 2428–2433.

Snyder, J.S., Alain, C. & Picton, T.W. (2006) Effects of attention on neuroelectric correlates of auditory stream segregation. *J. Cognitive Neurosci.*, **18**, 1–13.

Steriade, M., Timofeev, I. & Grenier, F. (2001) Natural waking and sleep states: a view from inside neocortical neurons. *J. Neurophysiol.*, **85**, 1969–1985.

Suga, N. & Ma, X. (2003) Multiparametric corticofugal modulation and plasticity in the auditory system. *Nat. Rev. Neurosci.*, **4**, 783–794.

Suga, N., Xiao, Z., Ma, X. & Ji, W. (2002) Plasticity and corticofugal modulation for hearing in adult animals. *Neuron*, **36**, 9–18.

Van Opstal, A.J. & Hepp, K. (1995) A novel interpretation for the collicular role in saccade generation. *Biol. Cybern.*, **73**, 431–445.

Van Opstal, A.J., Hepp, K., Suzuki, Y. & Henn, V. (1995) Influence of eye position on activity in monkey superior colliculus. *J. Neurophysiol.*, **74**, 1593–1610.

Versnel, H., Zwiers, M.P. & van Opstal, A.J. (2009) Spectrotemporal response properties of inferior colliculus neurons in alert monkey. *J. Neurosci.*, **29**, 9725–9739.

Walker, K.M., Bizley, J.K., King, A.J. & Schnupp, J.W. (2011) Multiplexed and robust representations of sound features in auditory cortex. *J. Neurosci.*, **31**, 14565–14576.

Werner-Reiss, U., Kelly, K.A., Trause, A.S., Underhill, A.M. & Groh, J.M. (2003) Eye position affects activity in primary auditory cortex of primates. *Curr. Biol.*, **13**, 554–562.

Zwiers, M.P., Versnel, H. & Van Opstal, A.J. (2004) Involvement of monkey inferior colliculus in spatial hearing. *J. Neurosci.*, **24**, 4145–4156.

© 2013 Federation of European Neuroscience Societies and Blackwell Publishing Ltd

14609568, 2013, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ejn.12180 by Radboud University Nijmegen, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License