EJN European Journal of Neuroscience

FENS Federation of European Neuroscience Societies

European Journal of Neuroscience, Vol. 39, pp. 1538–1550, 2014 doi:10.1111/ejn.12532

COGNITIVE NEUROSCIENCE

Task-related preparatory modulations multiply with acoustic processing 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 Medical Centre Nijmegen, Nijmegen, The Netherlands

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

Keywords: amplitude-modulated noise, awake, electrophysiology, reaction time, ripple, top-down

Abstract

We characterised task-related top-down signals in monkey auditory cortex cells by comparing single-unit activity during passive sound exposure with neuronal activity during a predictable and unpredictable reaction-time task for a variety of spectral-temporally modulated broadband sounds. Although animals were not trained to attend to particular spectral or temporal sound modulations, their reaction times demonstrated clear acoustic spectral-temporal sensitivity for unpredictable modulation onsets. Interestingly, this sensitivity was absent for predictable trials with fast manual responses, but re-emerged for the slower reactions in these trials. Our analysis of neural activity patterns revealed a task-related dynamic modulation of auditory cortex neurons that was locked to the animal's reaction time, but invariant to the spectral and temporal acoustic modulations. This finding suggests dissociation between acoustic and behavioral signals at the single-unit level. We further demonstrated that single-unit activity during task execution can be described by a multiplicative gain modulation of acoustic-evoked activity and a task-related top-down signal, rather than by linear summation of these signals.

Introduction

The mammalian auditory cortex (AC) not only encodes the spectral-temporal acoustic properties of sounds, but also processes a variety of other signals. For example, AC neurons have been shown to be modulated by multisensory integration (Schroeder *et al.*, 2001; Foxe *et al.*, 2002; Fu *et al.*, 2003; Brosch *et al.*, 2005; Ghazanfar *et al.*, 2005; Kayser *et al.*, 2007, 2008, 2009, 2010; Lakatos *et al.*, 2007), attentive behavior (Hubel *et al.*, 1959; Fritz *et al.*, 2003, 2005, 2007; Brechmann & Scheich, 2005; Sussman *et al.*, 2005, 2007; Scheich *et al.*, 2007; Yin *et al.*, 2008; Atiani *et al.*, 2009; Niwa *et al.*, 2012a,b; Massoudi *et al.*, 2013; Bizley *et al.*, 2013; Dong *et al.*, 2013), auditory learning and conditioning (Bakin *et al.*, 1996; Ji *et al.*, 2001; Kilgard & Merzenich, 2002; Kilgard *et al.*, 2001, 2002; Ohl & Scheich, 1997; Ohl and Scheich 2005), expected reward (Jaramillo & Zador, 2011; David *et al.*, 2012), the sound's location (Recanzone, 2000; Lee & Middlebrooks, 2011), and changes in eye position (Werner-Reiss *et al.*, 2003; Maier & Groh, 2010).

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

Received 27 January 2014, accepted 28 January 2014

Non-acoustic top-down signals could interact with acoustic signal processing in a variety of ways. For a purely additive signal interaction, the neural response in a given trial under active listening is a linear combination of the response to the acoustic signal response and a non-acoustic top-down signal. In the case of a purely multiplicative interaction, the latter combines with the acoustically-evoked response through a non-linear, multiplicative gain modulation. These different modes of interaction are usually not dissociated (Maier & Groh, 2010).

Here, we determined the type of interaction at the single-neuron level in the AC during a simple auditory reaction-time task. In such a task the acoustic tuning properties of AC neurons are unaffected by task performance, although evoked neural spike trains during active listening differ substantially from those recorded during passive sound exposure (Massoudi *et al.*, 2013). A stable acoustic representation by AC cells could facilitate the formation of an invariant percept of the acoustic environment when changing between passive and active listening conditions. Although the systematic differences in neural activity patterns suggested the presence of a top-down signal, the precise nature of this signal remained elusive (Massoudi *et al.*, 2013).

In this study, we characterised this signal by analysing the type of interaction between sound-evoked and top-down signals in AC

responses, as well as the potential contribution of acoustic and non-acoustic factors to this interaction. To that end, we compared single-unit activity to a wide range of acoustic stimuli [dynamic spectral–temporal ripples, and amplitude-modulated (AM) Gaussian noises], presented during either passive sound exposure, or during an auditory reaction-time task. As the latter contained unexpected (in non-predictable trials) as well as predictable onsets of spectral and/or temporal sound modulations that were randomly interleaved, it added a cognitive component to the paradigm (the presence or absence of prediction). We analysed the monkeys' reaction times as a function of the spectral and temporal modulation rates and of modulated sound-onset predictability, and identified the neural signature of these behavioral components on the single-unit activity patterns. Finally, we quantitatively dissociated additive and multiplicative signal-interaction modes in the neural response patterns.

Materials and methods

Subjects

We performed neurophysiological recordings in the left AC of two adult male rhesus monkeys (*Macaca mulatta*; monkey J, 7–9 kg; monkey T, 8–10 kg). Animals participated in the recording sessions for about 2 years. They were trained to manually respond to an acoustic change from a static broadband sound into a temporal and/or spectral modulation to receive a drop of water as a reward for each successful trial. Experiments were conducted in accordance with the European Communities Parliament and Council Directive (22 September 2010, 2010/63/EU). All experimental protocols were approved by the local Ethics Committee on Animal Research of the Radboud University Nijmegen (RU-DEC, 'Radboud University Dier Experimenten Commissie'). Monkeys were pair-housed to facilitate normal interactive behavior. 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 the animals 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 animals' fluid intake was increased to 400 mL daily. To monitor the animals' 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 animals' kidney function remained within the normal physiological range. Our procedures followed the water-restriction protocol of the Animal Use and Care Administrative Advisory Committee of the University of California at Davis (UC Davis, AUCAAC, 2001). Whenever an animal showed signs of discomfort, or illness, experiments were stopped and the animal was treated until the problem was solved.

Surgical procedures

After completing the initial training (when the monkey reached a day-to-day performance level of at least 80%), the animal underwent surgery under full anesthesia and sterile conditions. Anesthesia was maintained by artificial respiration (0.5% isoflurane and N_2O), and additional pentobarbital (3 mg/kg/h, i.v.), ketamine (0.1 mL/kg, i.m.), and fentanyl (20 μg/kg/h, i.v.) were administered. A stainless-steel recording chamber (12 mm diameter) was placed over a trepanned hole in the skull (10 mm diameter). The orientation and coordinates of the chamber were directed to the AC, as determined on the basis of magnetic resonance images. The chamber allowed a vertical approach to the left AC. 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 mounted in the skull. The analog electrode signal was amplified (model A-1, BAK Electronics), band-pass filtered between 0.1 and 15 kHz (custom-built eighth order Butterworth filter; model 3343, Krohn-Hite; 100 Hz high-pass cut-off), 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 algorithm detected individual action potentials (Brainware, V 9.07 for TDT, running on a PC, Windows 98; DELL). Data analysis and spike sorting were performed offline in MATLAB (version 7.14.0.739, R2012a; MathWorks, Natick, MA, USA).

The behavior and wakefulness of the animals were observed via an infrared camera during the experimental and training sessions.

Sound stimuli

Sound stimuli were digitally generated at a sampling rate of 100 kHz and delivered via BRAINWARE software and TDT2 hardware. A trigger, provided by a TG6 module, started the sound presentation (DA1, low-pass filtered at 20 kHz through a TDT2-FT6 module) and spike data acquisition. A speaker (Blaupunkt PCxg352, flat frequency characteristics within 5 dB between 0.2 and 20 kHz), positioned at the frontal central position at a distance of 1.0 m from the monkey, presented the stimuli at a fixed sound level of 60 dB (set by two programmable attenuators, PA4).

The ambient background acoustic noise level was about 30 dB. Acoustic foam that was mounted on the walls, floor, ceiling, and every large object in the room effectively absorbed reflections above 500 Hz.

Tones

Pure tones lasted for 150 ms and were presented over a frequency range from 250 to 16 000 Hz at four different sound levels (10, 30, 50, and 70 dB sound pressure level). Trials were presented at least four times in a randomised manner. The frequency-tuning curve of a neuron was determined from the average firing rate for each tone across all sound level presentations. The best frequency of each neuron was taken at the maximum of the tuning curve. The characteristic frequency was defined as the frequency that produced a response higher than the mean plus 2 SDs of the baseline activity at the lowest intensity. The cell's response-onset latency was defined as the moment after the pure-tone onset at which the firing rate exceeded the mean baseline activity plus twice its SD for the first time for at least 10 ms. The peak latency was the time at which the firing rate reached its maximum.

Ripples

The ripple stimuli consisted of a broadband complex of 126 spectral 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 envelopes of the complex were sinusoidally modulated in the spectrotemporal domain. The amplitude of each component was described as follows

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

where t is time (in s), x is position of the spectral component in octaves above the lowest frequency (250 Hz), ω is ripple velocity (in Hz; temporal modulation) and Ω is ripple density [cycles/octave (c/o); spectral modulation].

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] c/o) 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 corresponds to a downward direction, and $\Omega = 0$ means a pure AM sound (e.g. Versnel *et al.*, 2009). The modulation depth was 100% for all ripples. The static noise was frozen within each block of trials, and was generated setting the ripple modulation depth to 0, i.e. all spectral components had the same amplitude. The sound level of the static noise and ripple was 60 dB sound pressure level.

Amplitude-modulated noise

The AM noise stimuli were temporally modulated in 15 steps, logarithmically spaced from 2 to 256 Hz. Like the dynamic ripples, the modulation depth was 100%. In contrast to the ripples, the broadband noise was generated fresh for each stimulus. The average sound level was 60 dB sound pressure level.

Experimental paradigms

The neural responses were measured while monkeys were exposed to the spectral–temporal ripples and AM noises, each presented in different blocks of trials. Each stimulus started (Fig. 2, sound onset) with a static epoch ($\omega = 0$ Hz, $\Omega = 0$ c/o), which smoothly changed (Fig. 2, dynamic onset) into a dynamic epoch that was either a pseudo-randomly selected ripple that lasted for 1000 ms, or an AM noise that lasted for 700 ms (Fig. 2, sound offset).

Sounds were presented to the animal in two different behavioral paradigms: (i) passive listening (Fig. 2, top), in which the monkey was merely exposed to the sounds without a task, and the fixation light was off; and (ii) active listening, in which, upon presentation of a red fixation light at straight ahead, the monkey had to initiate a trial by pressing a lightweight bar. It heard the same sounds as during passive exposure, but now had to respond to the sound-modulation onset within 100–600 ms to receive a drop of water as reward. Trials in which the monkey did not detect the ripple onset correctly were repeated at a randomly selected trial within the same recording block. The passive and active paradigms were presented in different blocks, typically starting with the passive sound exposure.

In the passive experiment the trial started automatically with the static noise (Fig. 2, sound onset). Data collection started at 300 ms before sound onset. For ripples the trial ended at 700 ms after sound offset, yielding a total recording duration of 2500 ms. For AM noise stimuli the data acquisition ended at 500 ms after AM offset, giving a recording duration of 2000 ms per trial (Fig. 2, top). The number of trial repetitions was between four and 10, and the interstimulus interval was 2.5 s.

The active-listening paradigm contained active500 (A500) (Fig. 2, center) and active1000 (A1000) (Fig. 2, bottom) trials 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 lasted for either 500 ms (A500 trial), or 1000 ms (A1000 trial), upon which it changed into a spectral–temporal ripple (or AM noise) that lasted for 1000 ms (or 700 ms). All trials contained both static and dynamic epochs; there were no catch trials without a dynamic modulation. For each trial the data recording started at the moment of bar press, and finished 2500 ms later. The number of repetitions depended on the monkey's performance (typically no repetition for ripples, and five repetitions for AM).

Characterisation of recording site

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 or auditory area 1 (A1) and its immediate rostral part, area *R*] 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 of Paxinos *et al.* (1999). (ii) For both monkeys we reconstructed the tonotopic organisation of the recorded neurons, which demonstrated a systematic increase in characteristic frequency from anterior to posterior locations over a spatial range that corresponded well to other reports (e.g. Niwa *et al.*, 2012a; Dong *et al.*, 2013). This implied that the majority of recordings were probably taken from A1. As an illustration, Fig. 1 shows the tonotopic map for the recorded cells of monkey J. For monkey T, the characteristic frequency also increased over both the anterior–posterior and medial–lateral axes, which implies that recordings were taken from an area between the rostral area of the AC and A1 (not shown) (Bendor & Wang, 2008). (iii) 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). (iv) The tone-onset latency of the recorded sites was 22.6 ± 5.9 and 23.6 ± 5.6 ms for monkey J and T, respectively. (v) All neurons responded well to pure tones (best frequency: 250–16 000 Hz). (vi) The pure-tone tuning bandwidths for monkey J and T were 1.5 ± 1.2 and 1.5 ± 1.3 octaves, respectively. (vii) The pure-tone thresholds for monkey J and T were 21 ± 13 and 23 ± 12 dB sound pressure level, respectively. These tuning characteristics all fall in the same ranges as reported by Recanzone *et al.* (2000) for behaving monkeys in AC area A1, rostral area of the AC and caudomedial area of the AC.

The recording stability was verified by analysing spike-waveform variability during the different behavioral tasks. We included cells from which we obtained stable single-unit recordings with limited spike variability during presentation of both the ripples and AM noises in the active and passive paradigms. As a result, we obtained 100 cells from the two monkeys (J, $n = 57$; T, $n = 43$).

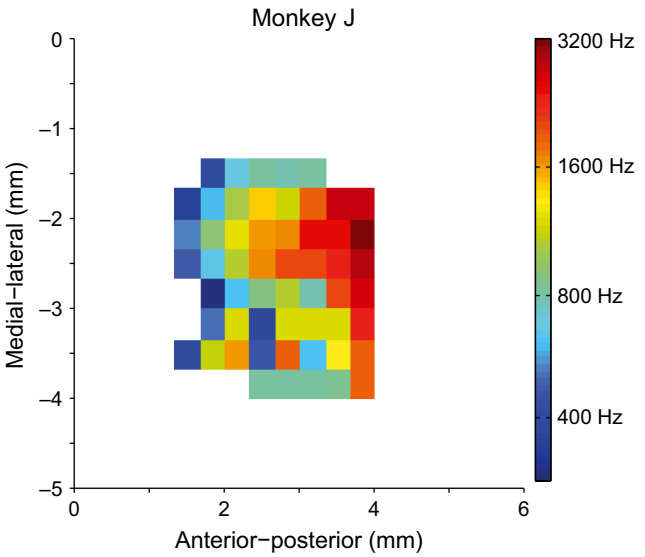

FIG. 1. Reconstructed tonotopic map for monkey J. Characteristic frequency (color coded) increases from anterior to posterior recording sites, which is indicative for A1. Each pixel represents the average characteristic frequency found at that location.

Data analysis

Spike-density function

To convert each spike-raster plot into a continuous spike-density function (a measure for the instantaneous mean firing rate across trials), we first binned the recorded spike times into a digital sequence with a time resolution of 1 ms, and then convolved each spike with a Gaussian kernel with a SD of 5 ms. Finally, the signals were added across trials, and we normalised the resulting function by the number of trials.

By default the data were aligned to sound onset (e.g. Fig. 5A and C). To relate the top-down signals to reaction time, we realigned all neural responses to the bar-release time by shifting the spike timings in each trial by the associated reaction time (e.g. Fig. 5B and D).

Results

Neural responses – passive and active listening

Figure 2 illustrates the experimental paradigms, and shows the neural responses of a representative neuron (J67), exposed to the AM stimuli (AM frequencies between 2 and 256 Hz). All sounds started with freshly generated static broadband Gaussian white noise, which was followed by an amplitude modulation (Fig. 2, sound). In passive hearing (Fig. 2, top), animals were merely exposed to the sounds without a required behavioral response. In that case the neuron was well tuned to the acoustics of the stimuli, as demonstrated by selective, phase-locked, responses to the different AM frequencies (see central group of trials in Fig. 2, top). Similar selective response behavior was obtained for the population of 100 neurons (data not shown here). This was also true for the neural responses to dynamic spectral–temporal ripples (see Massoudi *et al.*, 2013), for which the acoustic modulations varied in the temporal (velocity between 8 and 40 Hz) and spectral (density between −2 and +2 c/o) domains (see Materials and methods).

In the active hearing trials (Fig. 2, center and bottom), the monkey pressed a bar to initiate a trial. This experiment contained

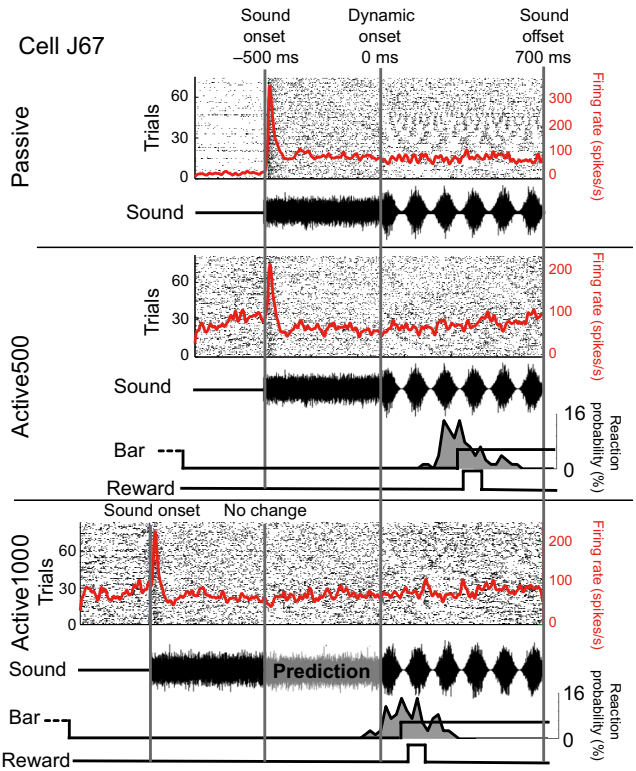

FIG. 2. Responses of cell J67 for AM noise during the passive (top), A500 (center) and A1000 (bottom) paradigms. Each panel shows the different epochs in the stimuli (shown below the spike rasters), indicated by dark-gray vertical lines (at sound onset, dynamic onset and sound offset). Spike rasters are sorted according to the AM frequency (AMF) (bottom, 2 Hz; top, 256 Hz). Each dot corresponds to a spike. Red lines show the trial-averaged spike-density curves (filtered with σ = 5 ms kernel). The passive paradigm elicited clear phase locking to temporal sound modulations in the mid-AMF range. Center and bottom panels: active paradigm to the identical sounds for A500 and A1000 trials, respectively (randomly interleaved in the experiment). The phase locking to sound modulations seems lost. Note the higher firing rates before stimulus onset, and a lower transient peak response. Panels also indicate the behavioral paradigm, i.e. bar press and release, and reward delivery. Gray histograms on the bar lines represent the reaction-time distributions of monkey J for this experiment. In A500 trials reaction times fall around the third stimulus modulation. In the A1000 paradigm, reaction times shift to much shorter values, around the first stimulus modulation, and sometimes even before modulation onset. Gray shading in the static stimulus epoch highlights the prediction period in which the monkey can anticipate the modulation onset.

two different conditions: in A500 trials (Fig. 2, center) the static noise lasted for 500 ms, whereas in A1000 trials (Fig. 2, bottom) it lasted for 1000 ms. To obtain a reward the animal had to react, by releasing the bar within 600 ms, to the perceived sound change from static noise to dynamic modulations. As both trial types were randomly interleaved at equal probabilities, the change occurred unexpectedly (probability 50%) in the A500 trials, but animals could in principle anticipate the sound onset in A1000 trials (with probability 100%), as soon as 500 ms of static noise had passed (highlighted by the light-gray epoch in Fig. 2, bottom).

Being in a task drastically changed the neural responses, which was readily obvious in changes of the mean overall firing rates (e.g. prestimulus firing rate increased for active trials; ripple mean ± SD: passive, 18 ± 17 spikes/s, A500, 32 ± 31 spikes/s, A1000, 31 ± 30 spikes/s; AM mean ± SD: passive, 23 ± 23 spikes/s, A500, 35 ± 36 spikes/s, A1000, 35 ± 36 spikes/s). Here, we will

not focus on these general differences, which might be due to a non-specific attentional signal (like alertness), but will focus instead on the substantial changes in the neural firing patterns related to task execution. During the dynamic response epochs of active trials (Fig. 2, center and bottom) the phase locking to acoustic modulations seemed less obvious in the spike rasters than for the passive trials (Fig. 2, top). This is due to two factors: (i) a higher firing rate introduces additional noise to the firing patterns because of the increased and variable baseline, and (ii) additional activity that is locked to the monkey's reaction time varies considerably from trial to trial, and therefore blurs the phase locking when spike rasters are aligned to sound onset, rather than to bar release. Note, however, that our previous study showed that, despite these additional signals, the underlying phase locking of AC cells to ripples, and hence their spectral–temporal tuning characteristics, remained fully intact for the active trials (Massoudi *et al.*, 2013). Below, we will demonstrate that the changes in firing patterns betray the presence of a dynamic task-related signal that interacts with the neuron's acoustic tuning response, but varies systematically from trial to trial with the animal's behavior. To better appreciate the different task-related factors of the experiment, we first quantify the behavioral effects of the acoustic and non-acoustic aspects of our paradigm.

Behavior

The A500 and A1000 trials led to clearly different behavioral responses of the monkeys (Figs 2 and 3). For example, during the recording session of cell J67 (Fig. 2, gray histograms), the reaction times for AM A1000 trials were on average 198 ms shorter than for AM A500 trials ($t_{\text{d.f.}\,=\,14} = 13.5$, $P < 0.0001$; A1000: median ± SE, 170 ± 11 ms, A500: median ± SE, 368 ± 14 ms). This result was obtained for all recording sessions, both stimulus types, and both animals. In the unpredictable A500 trials the reaction times across sessions were 384 ± 21 ms for ripples, and 357 ± 21 ms for the AM noises, with reaction times typically falling after the sound change (e.g. Fig. 2, gray histogram center). These reaction times are in line with a sensory-evoked hand-movement response (Rogal *et al.*, 1985). The predictability of sound-change onset in the A1000 trials invariably led to a substantial reduction of the reaction times to 188 ± 16 ms for ripples, and 158 ± 16 ms for AM noises, with 17.5% of the responses falling even before the stimulus change (e.g. Fig. 2, gray histogram, bottom). Importantly, although median reaction times varied substantially from day to day (as demonstrated by the considerable SDs), there was no systematic trend in these values over time, or in the difference between non-predictive and predictive trials (data not shown). Thus, animals did not display evidence of procedural learning during the course of the experiments.

Table 1 represents the combined performance of the two monkeys in the active trials for AM and ripple stimuli. Trials were counted 'correct' (and hence rewarded) when the manual reaction times were between 100 and 700 ms (see Materials and methods). In total, the correct trial percentages were around 90% for non-predictable A500 trials and 60–65% for the predictable A1000 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 8.50% of the ripple A500 trials animals reacted later than 700 ms, this decreased to 0.04% for AM A500 trials, which implies that the AM stimuli were easier to detect than the dynamic ripples. Only about 1% of A500 trials were early. Although late responses for A1000 trials were rare, more than 30% of them were early responses. This indicates that animals often released the bar on the basis of sound-modulation onset prediction in A1000 trials.

Reaction times and acoustics

To verify that the predictive and non-predictive differences reflected a cognitive component in the behavior (i.e. uncertainty and predictability of sound-modulation onset), we analysed the reaction times of both animals as a function of the acoustic modulation parameters for the 70 different stimuli (i.e. for 55 different spectral–temporal ripples, and 15 AM stimuli), and the two response conditions (A500 and A1000; data pooled and averaged across animals and cells).

When the modulation onset was unpredictable (i.e. a probability of 0.5 in A500 trials), the monkeys' reaction times strongly depended on both spectral (Fig. 3A; ANOVA; $F_{\text{d.f.}\,=\,5} = 91$, $P < 0.0001$) and temporal (Fig. 3A and C; ANOVA; ripples, $F_{\text{d.f.}\,=\,4} = 7.5$, $P < 0.0001$;

TABLE 1. Behavioral performance

Active trials	Total	Correct (%)	Missed (%)	Late (%)	Early (%)
Ripple A500	6602	87.00	3.20	8.50	1.00
Ripple A1000	6805	65.50	2.00	0.40	32.00
AM A500	7535	95.40	3.60	0.04	0.90
AM A1000	8875	61.40	2.30	0.00	36.10

Data from both monkeys are pooled. Reaction times were between 100 and 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.

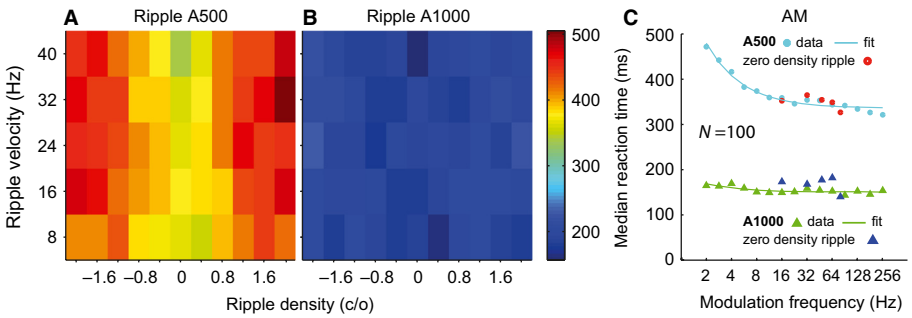

FIG. 3. Median reaction times (color coded) for each of the 55 dynamic ripples during A500 (A) and A1000 (B) trials. Note the clear pattern of faster and slower reaction times for A500 trials as function of temporal (velocity) and spectral (density) modulations, which is absent for the A1000 trials. (C) Median reaction times of each AM noise during A500 (cyan) and A1000 (green) trials. Cyan curve, Eqn 1; green line, Eqn 2. Data points for zero-density ripples in A500 (red) and A1000 (dark blue) are superimposed.

AM, $F_{\text{d.f.}=14} = 121$, $P < 0.0001$) acoustic modulations. Ripple densities of 0 c/o elicited the shortest reaction times of 337 ± 2.5 ms (mean ± SE; Fig. 3A, green), which gradually increased to 458 ± 12 ms for the highest ripple densities (−2 and 2 c/o, Fig. 3A, dark red). Moreover, for the ripples the animals' reaction times also reflected a significant interaction between spectral and temporal modulations (ANOVA; $F_{\text{d.f.}=20} = 1.88$, $P = 0.001$) with the combination of higher velocities and densities yielding an additional increase in reaction times.

In the AM detection paradigms we employed a much larger range (2–256 Hz) of temporal modulations than in the ripple experiments (8–40 Hz). The modulation frequency clearly affected the monkeys' behavioral responses for the AM noises (Fig. 3C, cyan dots), and in a similar fashion for the purely AM ripples at zero density (Fig. 3C, red dots). The monkeys reacted slowest for the lowest modulation frequency at 2 Hz (median ± SE, 483 ± 4 ms), and fastest for the highest modulation frequency at 256 Hz (median ± SE, 315 ± 4 ms). This finding seems at odds with reports suggesting that lower modulation frequencies are easier to perceive and detect than high modulation frequencies (e.g. Chi *et al.*, 1999). However, a possible explanation for this reaction-time effect could be that it simply takes more time for slow acoustic amplitude modulations to reach a perceptual detection threshold than for faster amplitude modulations (e.g. Heil & Neubauer, 2003), even though in general the detection threshold is lower for the lowest modulation frequencies. In line with this idea, we could fit the following simple linear regression model (with respect to the modulation period) to the AM reaction times

$$\text{RT}_{\text{AM500}} = 130/w + 359\,\text{ms} \quad (2)$$

where w is the modulation frequency (in Hz) and RT is the reaction time (in ms). This result ($F_{\text{d.f.}=7241} = 1632$, $P < 0.0001$) indicates that the perceptual detection threshold for AM stimuli (presented at 100% modulation depth) is reached at about 13% of the period of the modulation, and that an additional 359 ms are needed to prepare the manual motor response (Fig. 3C; cyan line).

In contrast, for the predictable modulation onsets (A1000 trials; Fig. 3B and C), reaction times were virtually independent of the acoustic modulation parameters (ANOVA; ripple spectral modulations, $F_{\text{d.f.}=5} = 3.4$, $P = 0.1$; ripple temporal modulations, $F_{\text{d.f.}=4} = 1$, $P = 0.41$; AM, $F_{\text{d.f.}=14} = 1.5$, $P = 0.1$), which becomes especially evident when applying the linear model of Eqn 2

$$\text{RT}_{\text{AM1000}} = 0/w + 162\,\text{ms} \quad (3)$$

The optimal fit to the data ($F_{\text{d.f.}=8282} = 11.1$, $P = 0.0008$) indicated that the manual responses did not depend on the modulation frequency (w), and required only 162 ms of preparation (Fig. 3C; green line). Again, the data obtained for the AM (zero-density) ripples nicely coincided with the AM responses (Fig. 3C, blue triangles). The predictive aspect of the A1000 task was therefore clearly reflected in a considerable reduction of manual reaction times of about 220 ms for the ripples (Fig. 3A and B) and about 200 ms for the AM noises (Fig. 3C, see also above).

Although the reaction-time distributions for A500 and A1000 trials differed substantially, they nonetheless overlapped, which means that a sizeable fraction of the A1000 responses could be considered as relatively slow (e.g. Fig. 2, gray histograms). To test whether these slower A1000 responses depended on spectral–temporal modulations, like the A500 responses, we divided the behavioral responses of all recorded A1000 trials into fast (< 225 ms) and slow (> 225 ms) reaction times, and analysed them as a function of AM frequency (for AM noises), and ripple density. Through this definition, 40% of the reaction times to ripples, and 33% of the reaction times to AM noises were categorised as slow responses for A1000 trials. Figure 4A and B shows the reaction times (relative to the mean A1000 reaction time) for slow (blue), and fast (red) A1000 trials, and for all A500 trials (gray). Interestingly, the slower A1000 responses were controlled by modulation-onset prediction (like fast A1000 responses), as well as by the acoustic modulations (like A500 responses).

Taken together, these findings demonstrate that the animals used all available information in the trial to obtain a reward. For A500 trials responses were exclusively guided by acoustic information. In A1000 trials, however, both acoustic (when the animal had a long reaction time) and non-acoustic (prediction, when the reaction was fast) signals determined the animal's responses.

Top-down neural signals

The neural responses in active trials differed substantially from the passive-evoked acoustic responses to the same sounds (e.g. Fig. 2) (Massoudi *et al.*, 2013). This suggests the presence of an additional task-related signal, which interacts with the acoustically-evoked response. The behavioral results shown in Figs 3 and 4 indicate that the manual reactions were guided by at least two factors – acoustic spectral–temporal modulations in the unpredictable A500 trials as well as in the late A1000 trials, and a non-acoustic predictive signal in all A1000 trials (both early and late). Thus, to identify the top-down neural signals in the active trials, different possibilities should be considered.

First, if the top-down signal is a general attentive signal (e.g. reflecting non-specific task engagement, or vigilance), it could induce an overall change in the firing rate of the active trials, which would relate to neither the acoustic events in the trial, nor the motor events of the monkey's reaction. Second, if the top-down signal is purely acoustic, the change in activity should be locked to the stimulus events (e.g. as a gain-amplitude modulation). Third, if the top-down modulation represents a purely behavioral signal (e.g. the preparation of the manual release response) it should be locked to the manual reaction time and be independent of the sound type. Note that, in principle, the top-down signal could reflect contributions from all of these different factors, in which case there will be

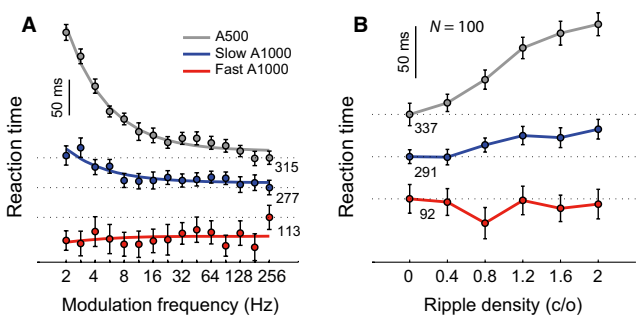

FIG. 4. Fast and slow reaction times in the predictive A1000 paradigm. (A) Mean reaction times [relative to the mean reaction time for AM frequency (AMF) = 256 Hz] as function of AMF for the unpredictable A500 trials (gray), fast (red) and slow (blue) A1000 trials. Note that the slow A1000 responses depend on AMF in a similar way as A500 responses. (B) Mean reaction times (relative to the mean at zero density) as function of ripple density (averaged across ripple velocities). Note that the slow A1000 reaction times vary with ripple density, like the A500 responses.

no single variable that fully explains the extracted neural modulation. To investigate these possibilities, we aligned each trial's response to either the sound-modulation onset, or the manual reaction time.

The results of this analysis for one of the neurons (T103) are shown in Fig. 5. We determined the average spike-density functions across trials (see Materials and methods) for passive (black), unpredictable A500 (red) and predictable A1000 (blue) trial types, and for ripples (Fig. 5A and B) and AM noises (Fig. 5C and D). In the left-hand column the data were aligned to sound onset. Although the average passive firing rate remained almost flat for the two sound types (ripples and AM noises), the A500 and A1000 neural responses changed considerably around the dynamic sound onset (indicated by the red and blue vertical lines, respectively), with roughly similar firing patterns for both sound types. For A500 trials the change in the activity started a few hundred milliseconds after dynamic sound onset, whereas for the A1000 trials the change started even before the sound-modulation onset. As the activity patterns do not align for the A500 and A1000 responses [neither when aligned to sound onset (Fig. 5A and C) nor to modulation onset (not shown)], the top-down signals were clearly not primarily driven by the sound acoustics. Interestingly, the neural responses appeared to follow roughly similar patterns as the underlying reaction-time distributions for the two trial types (filled histograms at the bottom of Fig. 5A and C). This suggests that the neural responses may be better related to the animal's reaction times.

Figure 5B and D shows the average active neural responses when trials were aligned to the reaction time. The firing rates show a steady increase that starts at about 150 ms before the reaction time, with patterns that are highly similar for both unpredictable (A500) and predictable (A1000) trial types. This indicates that the top-down signal could be attributed to a motor-preparation and/or a reward-expectation signal (Fig. 5B and D). The peak at about 20 ms after bar-release has a similar shape and onset latency as the sound-onset peak, and can be attributed to the soft sound produced by the rapid bar-release.

Figure 6 shows the results of this analysis for the population of 100 cells. The observations for neuron T103 persist for the total population; the neural responses in the active trials were much better aligned to the monkey's reaction time than to the acoustic modulations, and demonstrate that the top-down signals reflect a dynamic, non-acoustic signal at the level of single AC cells. The A1000 trials

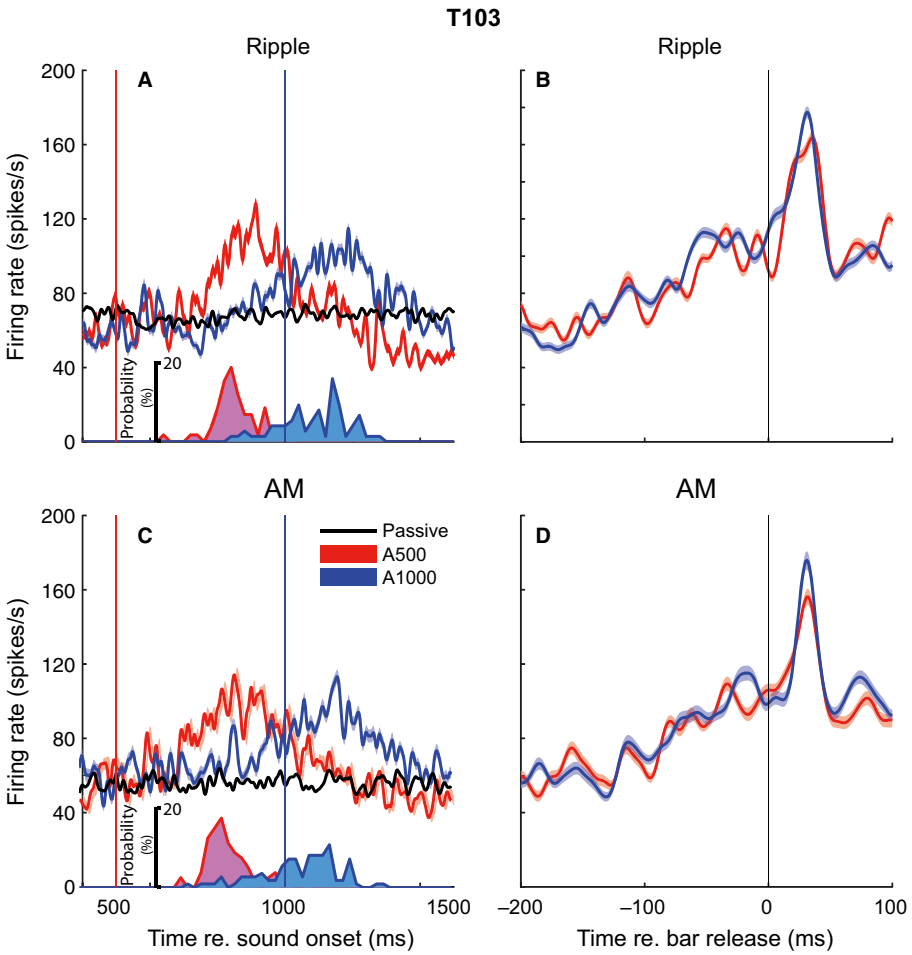

FIG. 5. Spike-density analysis for example cell T103. (A and B) Responses to ripples. (C and D) Responses to AM noises. (A and C) Spike-density functions for all passive (black), A500 (red) and A1000 (blue) trials aligned to sound onset. Mean and SE are denoted by solid lines and transparent patches, respectively. Filled histograms represent reaction-time probability (%) distributions. Note the resemblance between the reaction times and active response modulations. The red and blue vertical lines at 500 and 1000 ms denote modulation onset for passive and A500 trials and for A1000 trials, respectively. (B and D) Spike-density functions aligned to bar release. The curves fully superimpose. The vertical line at 0 ms denotes handle-bar release.

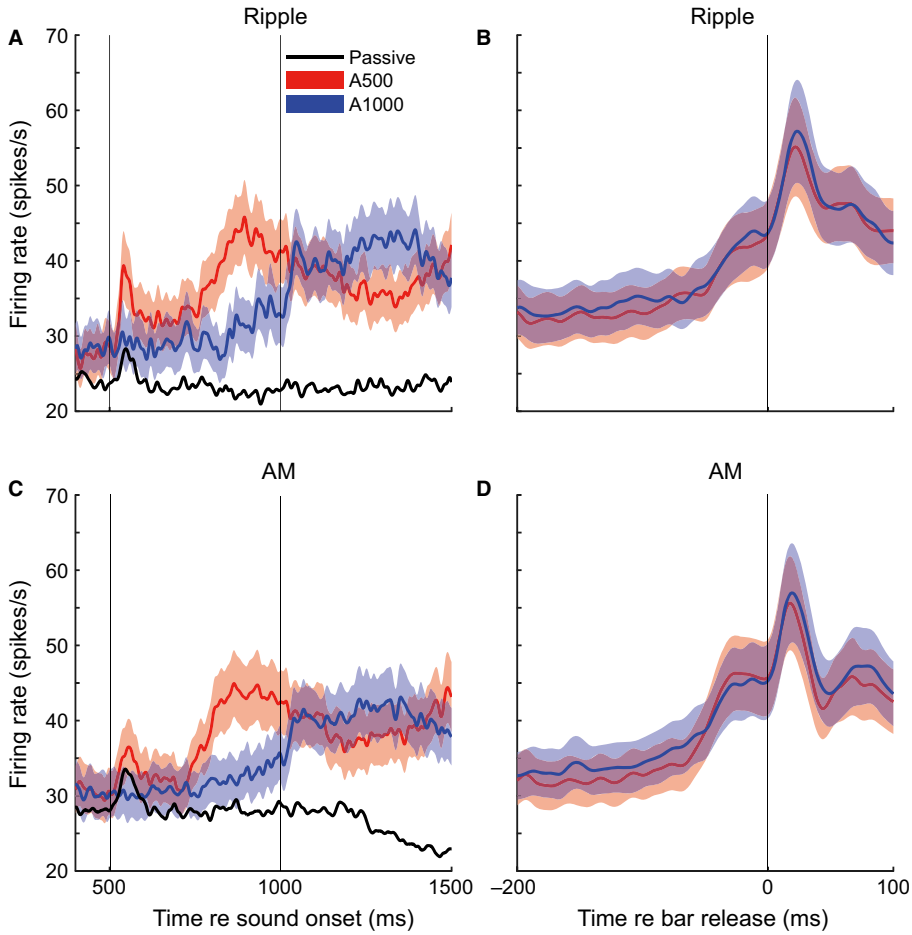

FIG. 6. Population analysis for 100 AC neurons. Average spike-density functions, with SDs (shaded), across cells for the different trial types, sound types, and hearing conditions. The same conventions apply as in Fig. 5.

with early responses (~30% of the A1000 trials; Table 1) had a similar top-down signal (data not shown; there were too few early-response A500 trials to reliably obtain the instantaneous firing rate).

In the analysis so far we averaged neural responses across the different stimuli in the experiment. Although the top-down signals were not locked to the acoustic modulations (Fig. 6A and C), it is conceivable that the signals themselves could somehow depend on the spectral and temporal modulation rates. For example, as the behavioral data (Figs 3 and 4) clearly indicated that the reaction times depended quite reliably on the stimulus acoustics and behavioral context (predictive and non-predictive trials, or fast and slow responses), the amplitude of the top-down signal could depend on the reaction time itself. Alternatively, because the top-down signal rides on top of the sound-evoked activity, it could be stronger for preferred stimuli for each cell.

To investigate the influence of acoustics on the top-down signal, we selected the trials along the spectral and temporal stimulus modulation dimensions, and behavioral contexts (predictive/non-predictive), and calculated the sound-specific top-down signals for the neural population. Figure 7A and B separates the ripple trials for all cells and reaction times into the different ripple densities (pooled across the five ripple velocities, while pooling $\Omega > 0$ with $\Omega < 0$; Fig. 7A) and different ripple velocities (pooled across the 11 ripple densities, Fig. 7B). Figure 7C shows the data for the 15 different AM frequencies of the AM stimuli. The left-hand panels of Fig. 7 correspond to the A500 condition and the right-hand panels to the A1000 condition. The individual curves are remarkably similar within each modulation range (ANOVA over the epoch [−200 to 0] ms; ripple spectral modulations: A500, $F_{\text{d.f.}=10} = 0.31$, $P = 0.98$; A1000, $F_{\text{d.f.}=10} = 0.16$, $P = 1$; ripple temporal modulations: A500, $F_{\text{d.f.}=4} = 0.06$, $P = 0.99$; A1000, $F_{\text{d.f.}=4} = 0.01$, $P = 1$; AM stimuli: A500, $F_{\text{d.f.}=14} = 0.2$, $P = 1$; A1000, $F_{\text{d.f.}=14} = 0.13$, $P = 1$), and between different modulation types. Thus, spectral–temporal modulation rates, or predictive and non-predictive modulation onsets, had no influence on the size and shape of the top-down neural modulation.

Model selection

We here consider two potential models for the neural interaction of acoustic and behavioral signals on single AC neurons – linear summation, and non-linear gain modulation. For the summation model, the neural response for a trial in the active reaction-time task can be described by

$$R_{\text{ADD}}(S;t) = A_{\text{ACT}}(S;t) + B(t - t_R) \quad (4)$$

where $A_{\text{ACT}}(S;\ t)$ is the acoustic-evoked response component to sound S in the active listening condition, and $B(t\ -\ t_R)$ is the behavioral top-down signal, which is referenced to the reaction time, t_R

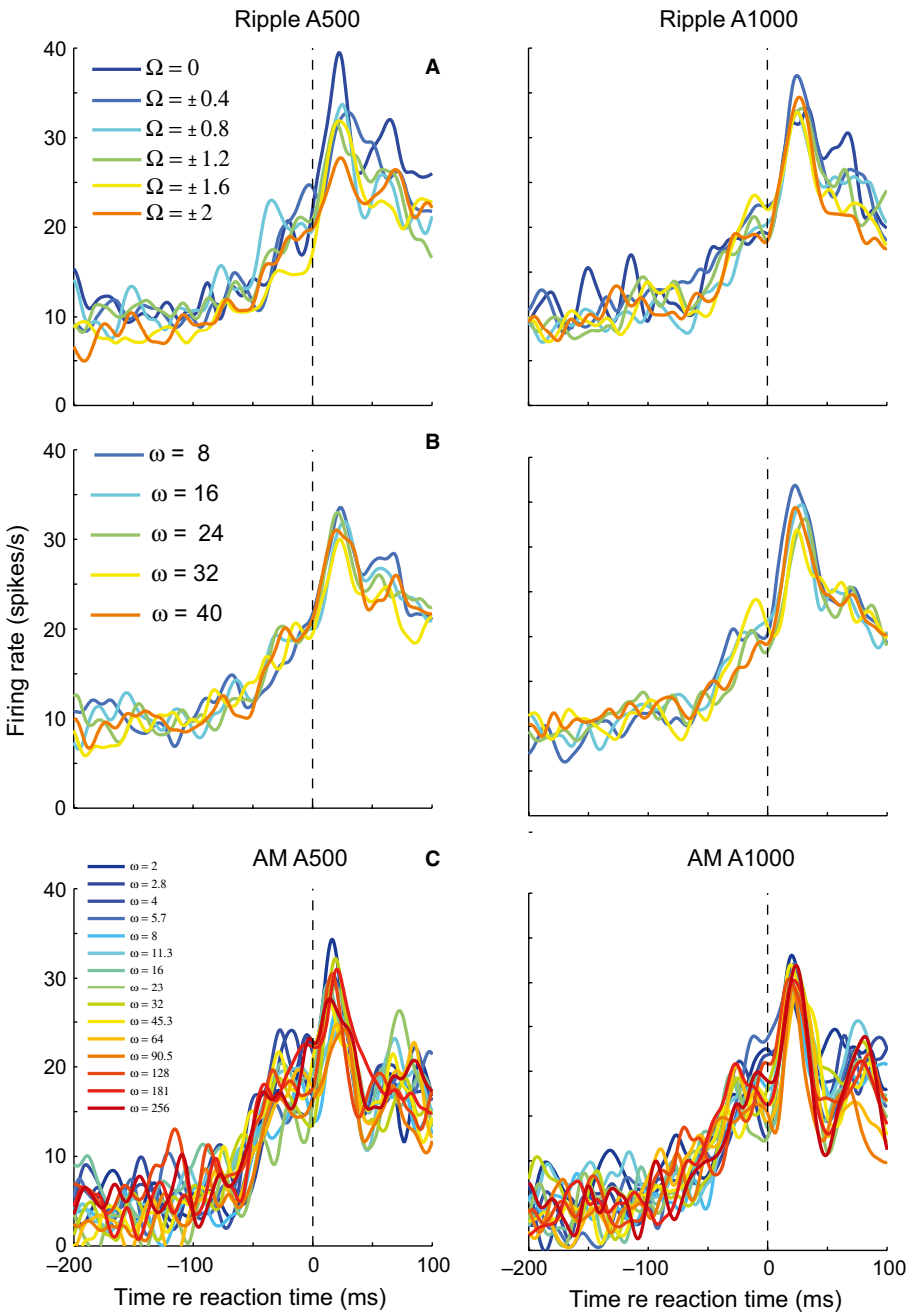

FIG. 7. Population-averaged spike-density functions separated for the different spectral–temporal sound modulations. Responses aligned to reaction time, and averaged across all cells. (A) Responses for ripples at different absolute densities (pooled for five velocities). (B) Responses for ripple velocities (pooled across 11 densities). (C) Responses for 15 different AM frequencies of AM Gaussian white-noise stimuli. Left-hand column, A500 data; right-hand column, A1000 data. Note the similarity of individual response curves across stimulus properties and hearing conditions.

(according to Fig. 6B and D), and taken independent of the acoustics of sound S (demonstrated in Fig. 7). In contrast, the multiplicative gain-modulation model for that trial reads

$$R_{\mathrm{MUL}}(S;t) = (1 + B(t - t_R)) \times A_{\mathrm{ACT}}(S;t) \qquad (5)$$

We employ the property that the acoustic component of the response during active listening is the same as for passive listening, as recently reported by Massoudi *et al.* (2013). In other words, $A_{\mathrm{ACT}}(S,t) = A_{\mathrm{PASS}}(S,t)$. Further, in the absence of a behavioral signal $[B(t - t_R) = 0]$ both models reduce to the passive-evoked response. To dissociate the two models, we use the observation that the mean firing rates of AC cells for the static and dynamic components of the stimuli during passive listening varied considerably from cell to cell (ranging from only a few spikes/s to over 100 spikes/s), but, when averaged across sound modulations, remained roughly constant during the trial (see Figs 5 and 6, black lines). As a measure for this passive sound-evoked activity we calculated, for each cell, the mean firing rate obtained during the passive static-noise epoch, separately for ripples and AM noises: μ_{PASS} = mean $(A_{\mathrm{PASS}}(S;\, t))$. We took $200 < t < 500$ ms to avoid a biased contribution from the transient sound-onset peaks. We then subtracted this

value from the mean active response during the acoustic modulation epoch for each active trial, and aligned the resulting modulatory activity to the animal's reaction time for that trial. We subsequently determined the spike-density function, now pooled across all ripples and AM noises, to obtain the mean dynamic change in activity for each neuron resulting from the top-down signals

$$\Delta R_{\mathrm{MOD}}(t - t_R) = R(t - t_R) - \mu_{\mathrm{PASS}} \quad (6)$$

Applying Eqn 6 to the two interaction schemes (Eqns 4 and 5) predicts for the additive model that

$$\Delta R_{\mathrm{ADD}}(t - t_R) = B(t - t_R) \quad (7)$$

and for the multiplicative model

$$\Delta R_{\mathrm{MUL}}(t - t_R) = \mu_{\mathrm{PASS}} \times B(t - t_R) \quad (8)$$

In other words, when plotting the modulatory firing rate vs. the mean passive firing rate, μ_{PASS}, it would yield a flat line if the additive model is valid, with an offset at the mean top-down modulation-firing rate. In contrast, the multiplicative model predicts a linear relation with a slope that is determined by the average top-down modulation strength. A combination of additive and multiplicative signals on the neurons would give an additional offset to Eqn 8. Figure 8A shows the results of this analysis for the population of 100 cells (pooled for the two monkeys, and across ripples and AM noises). Data points were colored according to their mean passive firing rate (see color bar). The correlation coefficient of $r = 0.62$ is highly significant ($F_{\mathrm{d.f.}\,=\,98} = 61.3$, $P < 0.001$). The linear regression line through the data points has a slope of 0.51 ($t_{\mathrm{d.f.}\,=\,98} = 7.83$, $P < 0.001$), and an offset that is indistinguishable from zero (offset = −1.4 spikes/s, $t_{\mathrm{d.f.}\,=\,98} = -0.44$, $P = 0.65$). Thus, the data do not support the presence of an additive component, and clearly favor the prediction of the multiplicative model (Eqns 5 and 8).

In Fig. 8B we show the instantaneous modulatory firing rates (Eqn 6) aligned to bar release, after sorting the cells for their mean passive firing rates (color bar). The figure illustrates that the higher the cells' mean firing rate, the stronger the change in modulatory activity across the measured epoch (here 150 ms), as predicted by Eqn 8.

Discussion

Summary

This study is the first to compare task-related top-down signals in the monkey AC for different acoustic stimuli during the same behavioral demands. Our results revealed clear top-down components during active hearing that could be extracted from the animal's behavior (Figs 3 and 4), which contained acoustic (spectral–temporal) and non-acoustic (cognitive) factors. We also identified a systematic top-down signal at the level of single-unit activity that was locked to the animal's behavioral response but invariant to the acoustic modulations (Figs 5–7). We demonstrated that the neural top-down component is processed independently from the acoustic signals by AC neurons in a multiplicative, separable, way (gain modulation).

Behavioral response properties

The behavioral results demonstrated that the monkeys' manual reaction times depended on the spectral–temporal sound characteristics in unpredictable trials (Fig. 3A and C) and, interestingly, also in the slower predictable trials (Fig. 4). In all A1000 trials the reaction times were substantially shortened by prediction of the upcoming sound change (Figs 2 and 3) (Jaramillo & Zador, 2011; Massoudi *et al.*, 2013). These results suggest that the animals adopted a single strategy for both trial types, which led them to use all available information to earn a reward as fast as possible. In A500 trials the monkey only had acoustic information to achieve this goal, as all sounds and trial types were randomly selected, and equally likely and important. The results for the ripples and AM noises reflected this acoustic dependence quite clearly. Reaction times to unpredictable AM noises and zero-density ripples were only determined by the AM frequencies, and could be fitted well with the simple cycle-detection regression model of Eqn 1 (Fig. 3C). The manual responses to the unpredictable moving ripples (density ≠ 0) relied on both spectral and temporal modulations, showing a significant interaction between these acoustic features (Fig. 3A).

In the A1000 trials, however, the monkey acquired target-onset predictability as an additional source of information. This non-acoustic cognitive top-down signal appeared to over-ride the dependence on purely acoustic information; reaction times not only shortened by about 200 ms on average, their reliance on acoustic

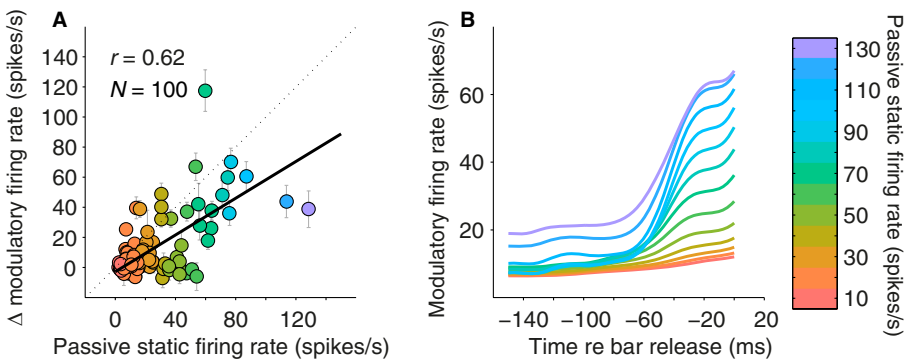

FIG. 8. Evidence for multiplicative signal interaction at AC cells. (A) Change in mean modulatory firing rate as a function of the mean passive firing rate in the static epoch (based on Eqn 6). For each cell ($n = 100$), data were pooled for ripples and AM noises, and for A500 and A1000 trials. Note linear relation with zero offset ($r = 0.62$). (B) The modulatory signal aligned to bar release. For higher passive static firing rates, the total modulatory signal is also stronger. Different colors denote different mean passive firing rates (color bar).

modulation rates had also virtually disappeared (Fig. 3B and C). However, when the animal withheld its immediate response in A1000 trials, the longer reaction times (> 225 ms) also depended on the acoustic modulations. These results demonstrate a clear qualitative difference between the predictable and unpredictable trial types, which was reflected in the single-unit responses of AC neurons (Figs 2, 5A and C, and 6A and C), but not in the strength of the top-down signals (Figs 5B and D, 6B and D, and 7). In other words, the neural response does not reveal an additional prediction signal.

Task-related effects on auditory cortex responses

Executing a sound-detection task considerably changed the neural firing patterns for all AC cells, when compared with the passive condition (Fig. 2). For both sound types, active engagement in the task led to an overall increase of the spontaneous (background) firing rate prior to trial initiation (Fig. 2), and during the acoustic-evoked firing rates, when compared with passive hearing (Figs 5A and C, and 6A and C) (see also Massoudi *et al.*, 2013). However, the pure stimulus-evoked mean activity was hardly affected, as is visible in the responses during the static noise (Figs 5A and C, and 6A and C). The overall increase in mean firing rate during the dynamic epochs is therefore attributable to the behavioral top-down modulation. Most studies have reported that attentional demands, including non-specific effects due to arousal and vigilance, increase cortical firing rates (Edeline *et al.*, 2001; Steriade *et al.*, 2001; Scott *et al.*, 2007; Niwa *et al.*, 2012a,b), whereas some studies reported decreased AC responses during task performance in rats (Otazu *et al.*, 2009) and ferrets (Atiani *et al.*, 2009), and overall inhibition during behavior in the somatosensory cortex of mice (Crochet & Petersen, 2006; Poulet & Petersen, 2008). In contrast to these reports, our data did not reveal strong non-specific effects on the firing rates of AC neurons, which was further supported by the zero offset in the regression analysis of Fig. 8A.

Note that in our paradigms the animals did not, and could not, attend to a particular precued acoustic feature of the upcoming target sound, as they were not instructed to do so, and were merely anticipating any detectable spectral–temporal modulation while listening to the static noise. Neither could the animals predict any particular acoustic modulation in advance, as stimuli were all equally likely and important for obtaining a reward. The only trial-specific feature that mattered in the experiment was the duration of the static interval. Our data clearly demonstrate that the animals indeed used this information, and the neural activity patterns reflected this non-acoustic aspect of the task.

Realigning the active responses of each trial to its associated reaction time revealed a signal starting at about 150 ms before the bar release. This signal was clearly present at the single cell and population level (Figs 5B and D, and 6B and D), and was found to be invariant to the acoustic modulations (Figs 5–7). The signal may be attributed to either motor preparation for bar release, or to reward expectation. These two interpretations could not be dissociated as, in our experiments, reward delivery fell immediately after bar release. Previous studies have also shown the presence of motor-preparation (Brosch *et al.*, 2005, 2011a,b; Selezneva *et al.*, 2006; Yin *et al.*, 2008; Niwa *et al.*, 2012a), and reward expectancy (Yin *et al.*, 2008) signals at the level of AC neurons, although the relative proportions of cells carrying such a signal differ between studies from < 15% (e.g. Yin *et al.*, 2008) to more than 60% (e.g. Brosch *et al.*, 2005). A possible explanation for this apparent discrepancy could be related to the results on the multiplicative effects in Fig. 8. Although many recorded AC cells carry the task-related signal (47% of cells; Wilcoxon signed rank test on the change in modulatory firing rate), its strength, and hence its detectability, are directly proportional to the mean acoustic-evoked firing rate. Variation in these firing rates between studies could thus underlie the different proportions of cells seen to carry multiplicative non-acoustic modulations. These task-related signals could originate from different cortical areas such as parietal and frontal cortical areas (reward) (Romanski *et al.*, 1999; Fritz *et al.*, 2010), or from the premotor cortex (motor preparation) (Lemus *et al.*, 2009).

Although for the fast A1000 trials (Fig. 4, red) the animals responded mainly on the basis of sound-change prediction, we observed no additional neural prediction signal in AC cells, as the A500 and A1000 curves did not systematically differ. This suggests that the prediction signal, possibly originating from the prefrontal cortex (Nobre *et al.*, 1999), induces a temporal shift (by about 200 ms) of the top-down modulation on AC cells.

Additive and multiplicative interactions

Multiplicative interactions between sensory signals, top-down signals, and motor-related signals have been described for a variety of brain areas and sensory-motor systems. For example, reports have described eye-position gain fields for visually responsive cells in the monkey posterior parietal cortex (Andersen *et al.*, 1985), and for eye-movement-related visual-motor cells in the monkey midbrain superior colliculus (Van Opstal *et al.*, 1995). Similarly, in the auditory system, eye-position gain-field tuning, as well as multiplicative spatial sensitivity, has been reported for auditory sensitive cells in the monkey inferior colliculus (Groh *et al.*, 2001; Zwiers *et al.*, 2004). It has also been shown that attention may influence the firing rate of visual (McAdams & Maunsell, 1999; Treue & Martínez Trujillo, 1999; Reynolds *et al.*, 2000; Maunsell & Treue, 2006), and somatosensory (Sripati & Johnson, 2006) cortex cells in a multiplicative fashion. This indicates that attention can change the gain of neuronal responses without changing their underlying response properties, or tuning curve. Different gain mechanisms could be invoked by attention, such as enhancing the attended stimulus strength, changing the neuronal gain of the behaviorally salient target, or sharpening the sensory selectivity of the neuron. Polack *et al.* (2013) recently identified a noradrenergic mechanism underlying gain modulations in mouse V1 cells, which led to a stronger, but less variable, change in the cells' membrane potential during a locomotion task.

Other studies have proposed additive interactions between sensory and non-sensory signals at the level of AC neurons. For example, Maier & Groh (2010) suggested an additive interaction of eye-position signals with auditory-evoked activity in the behaving monkey AC, and Schnupp *et al.* (2001) demonstrated that the spatial selectivity of AC neurons in the anesthetised ferret is well predicted by a linear integration model of sound levels within each frequency band and ear. Additive interactions have also been proposed for attentional top-down signals in the visual cortex (Luck *et al.*, 1997; Williford & Maunsell, 2006; Buracas & Boynton, 2007; Thiele *et al.*, 2009).

The analysis in Fig. 8 demonstrates that it is possible to dissociate additive and multiplicative interactions on AC cells by exploiting the cell-to-cell variability of mean acoustic responses, together with the idea that the sound-evoked responses are unaffected by task performance. We recently verified this latter requirement for the spectral–temporal receptive fields of AC neurons (Massoudi *et al.*, 2013). Indeed, a relatively slow modulation on top of a rapid

acoustic tuning response will not affect the spectrum of the latter. This requirement on the top-down signals is typically fulfilled; the monotonic rising phase of the top-down signal is confined to the low-frequency spectrum (up to a few Hz; Fig. 8B), when compared with the rapid phase-locked response profiles of AC cells to acoustic amplitude modulations (up to 256 Hz in our sample of cells).

Our population analysis revealed that the non-acoustic top-down modulation and the sound-evoked activity combine in a separable way through multiplicative gain modulation (Eqns 5 and 8; Fig. 8). Such multiplexing of different signal types is beneficial for an efficient and reliable transmission of multiple sources of information, which express different types of variables, through the same channel. A relatively simple read-out mechanism, based on linear weighting of the cell contributions in the recruited population, could subsequently decode the individual signals for further processing. Quantitative models for such read-out mechanisms have been described for eye-position, sound-level and sound-location gain fields of cells in the auditory midbrain inferior colliculus (Zwiers *et al.*, 2004), and for eye-position gain fields in the visuomotor midbrain superior colliculus (Van Opstal & Hepp, 1995).

Acknowledgements

We thank Hans Kleijnen, Dick Heeren, Günter Windau, and Stijn Martens for their valuable technical assistance, and Alex Hanssen and the central animal facility for excellent monkey care. This research was supported by a VICI and Open Competition grant of the Earth and Life Sciences of NWO (ALW 865.05.003 and 821.02.009; A.J.V.O. and R.M.), and the Radboud University Nijmegen (A.J.V.O., R.M., M.M.V.W., S.M.C.I.V.W. and H.V.).

Abbreviations

A1000, active1000; A1, primary auditory cortex or auditory area 1; A500, active500; AC, auditory cortex; AM, amplitude-modulated; c/o, cycles/octave.

References

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., South, D.A. & Weinberger, N.M. (1996) Induction of receptive field plasticity in the auditory cortex of the guinea pig during instrumental avoidance conditioning. *Behav. Neurosci.*, **110**, 905–913.

Bendor, D. & Wang, X. (2008) Neuronal response properties of primary, rostral and rostrotemporal core fields in the auditory cortex of marmoset monkeys. *J. Neurophysiol.*, **100**, 888–902.

Bizley, J.K., Walker, K.M., Nodal, F.R., King, A.J. & Schnupp, J.W. (2013) Auditory cortex represents both pitch judgments and the corresponding acoustic cues. *Curr. Biol.*, **23**, 620–625.

Brechmann, A. & Scheich, H. (2005) Hemispheric shifts of sound representation in auditory cortex with conceptual listening. *Cereb. Cortex*, **15**, 578–587.

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) Formation of associations in auditory cortex by slow changes of tonic firing. *Hearing Res.*, **271**, 66–73.

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

Buracas, G.T. & Boynton, G.M. (2007) The effect of spatial attention on contrast response functions in human visual cortex. *J. Neurosci.*, **27**, 93–97.

Chi, T., Gao, Y., Guyton, M.C., Ru, P. & Shamma, S. (1999) Spectro-temporal modulation transfer functions and speech intelligibility. *J. Acoust. Soc. Am.*, **106**, 2719.

Crochet, S. & Petersen, C.C.H. (2006) Correlating whisker behavior with membrane potential in barrel cortex of awake mice. *Nat. Neurosci.*, **9**, 608–610.

David, S.V., Fritz, J.B. & Shamma, S.A. (2012) Task reward structure shapes rapid receptive field plasticity inauditory 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.

Dong, C., Qin, L., Zhao, Z., Zhong, R. & Sato, Y. (2013) Behavioral modulation of neural encoding of click-trains in the primary and nonprimary auditory cortex of cats. *J. Neurosci.*, **33**, 13126–13137.

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.

Foxe, J.J., Wylie, G.R., Martinez, A., Schroeder, C.E., Javitt, D.C., Guilfoyle, D., Ritter, W. & Murray, M.M. (2002) Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J. Neurophysiol.*, **88**, 540–543.

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.B., Elhilali, M. & Shamma, S.A. (2005) 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., Maier, J.X., Hoffman, K.L. & Logothetis, N.K. (2005) Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J. Neurosci.*, **25**, 5004–5012.

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.

Heil, P. & Neubauer, H. (2003) A unifying basis of auditory thresholds based on temporal summation. *Proc. Natl. Acad. Sci. USA*, **100**, 6151–6156.

Hubel, D.H., Henson, C.O., Rupert, A. & Galambos, R. (1959) "Attention" units in the auditory cortex. *Science*, **129**, 1279–1280.

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.

Kayser, C., Petkov, C.I., Augath, M. & Logothetis, N.K. (2007) Functional imaging reveals visual modulation of specific fields in auditory cortex. *J. Neurosci.*, **27**, 1824–1835.

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

Kayser, C., Petkov, C.I. & Logothetis, N.K. (2009) Multisensory interactions in primate auditory cortex: fMRI and electrophysiology. *Hearing Res.*, **258**, 80–88.

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.

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

Luck, S.J., Chelazzi, L., Hillyard, S.A. & Desimone, R. (1997) Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J. Neurophysiol.*, **77**, 24–42.

Maier, J.X. & Groh, J.M. (2010) Comparison of gain-like properties of eye position signals in inferior colliculus versus auditory cortex of primates. *Front. Integr. Neurosci.*, **4**, 1–11.

Massoudi, R., Van Wanrooij, M.M., Van Wetter, S.M.C.I., Versnel, H. & Van Opstal, A.J. (2013) Stable bottom-up processing during dynamic top-down modulations in monkey auditory cortex. *Eur. J. Neurosci.*, **37**, 1830–1842.

Maunsell, J.H.R. & Treue, S. (2006) Feature-based attention in visual cortex. *Trends Neurosci.*, **29**, 317–322.

McAdams, C.J. & Maunsell, J.H. (1999) Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J. Neurosci.*, **19**, 431–441.

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.

Nobre, A.C., Coull, J.T., Frith, C.D. & Mesulam, M.M. (1999) Orbitofrontal cortex is activated during breaches of expectation in tasks of visual attention. *Nat. Neurosci.*, **2**, 11–12.

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.

Paxinos, G, Huang, X.-F. & Toga, A.W. (1999) *The Rhesus Monkey Brain in Stereotaxic Coordinates*. Academic Press, San Diego, CA.

Polack, P.-O., Friedman, J. & Golshani, P. (2013) Cellular mechanisms of brain state-dependent gain modulation in visual cortex. *Nat. Neurosci.*, **16**, 1331–1339.

Poulet, J.F.A. & Petersen, C.C.H. (2008) Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice. *Nature*, **454**, 881–885.

Recanzone, G.H. (2000) Spatial processing in the auditory cortex of the macaque monkey. *Proc. Natl. Acad. Sci. USA*, **97**, 11829–11835.

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.

Reynolds, J.H., Pasternak, T. & Desimone, R. (2000) Attention increases sensitivity of V4 neurons. *Neuron*, **26**, 703–714.

Rogal, L., Reible, G. & Fischer, B. (1985) Reaction times of the eye and the hand of the monkey in a visual reach task. *Neurosci. Lett.*, **58**, 127–132.

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.

Scheich, H., Brechmann, A., Brosch, M., Budinger, E. & Ohl, F.W. (2007) The cognitive auditory cortex: taskspecificity of stimulus representations. *Hearing Res.*, **229**, 213–224.

Schnupp, J.W., Mrsic-Flogel, T.D. & King, A.J. (2001) Linear processing of spatial cues in primary auditory cortex. *Nature*, **414**, 200–204.

Schroeder, C.E., Lindsley, R.W., Specht, C., Marcovici, A., Smiley, J.F. & Javitt, D.C. (2001) Somatosensory input to auditory association cortex in the macaque monkey. *J. Neurophysiol.*, **85**, 1322–1327.

Scott, B.H., Malone, B.J. & Semple, M.N. (2007) Effect of behavioral context on representation of a spatial cue in core auditory cortex of awake macaques. *J. Neurosci.*, **27**, 6489–6499.

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

Sripati, A.P. & Johnson, K.O. (2006) Dynamic gain changes during attentional modulation. *Neural Comput.*, **18**, 1847–1867.

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

Sussman, E.S., Bregman, A.S., Wang, W.J. & Khan, F.J. (2005) Attentional modulation of electrophysiological activity in auditory cortex for unattended sounds within multistream auditory environments. *Cogn. Affect. Behav. Ne.*, **5**, 93–110.

Sussman, E.S., Horváth, J., Winkler, I. & Orr, M. (2007) The role of attention in the formation of auditory streams. *Percept. Psychophys.*, **69**, 136–152.

Thiele, A., Pooresmaeili, A., Delicato, L.S., Herrero, J.L. & Roelfsema, P.R. (2009) Additive effects of attention and stimulus contrast in primary visual cortex. *Cereb. Cortex*, **19**, 2970–2981.

Treue, S. & Martínez Trujillo, J.C. (1999) Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, **399**, 575–579.

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.

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.

Williford, T. & Maunsell, J.H.R. (2006) Effects of spatial attention on contrast response functions in macaque area V4. *J. Neurophysiol.*, **96**, 40–54.

Yin, P., Mishkin, M., Sutter, M. & Fritz, J.B. (2008) Early stages of melody processing: stimulus-sequence and task-dependent neuronal activity in monkey auditory cortical fields A1 and R. *J. Neurophysiol.*, **100**, 3009–30029.

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