Correlates of Auditory Decision-Making in Prefrontal, Auditory, and Basal Lateral Amygdala Cortical Areas

Julia L. Napoli,* Corrie R. Camalier,* Anna-Leigh Brown, Jessica Jacobs, Mortimer M. Mishkin, and Bruno B. Averbeck

Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland 20892-4415

Spatial selective listening and auditory choice underlie important processes including attending to a speaker at a cocktail party and knowing how (or whether) to respond. To examine task encoding and the relative timing of potential neural substrates underlying these behaviors, we developed a spatial selective detection paradigm for monkeys, and recorded activity in primary auditory cortex (AC), dorsolateral prefrontal cortex (dlPFC), and the basolateral amygdala (BLA). A comparison of neural responses among these three areas showed that, as expected, AC encoded the side of the cue and target characteristics before dlPFC and BLA. Interestingly, AC also encoded the choice of the monkey before dlPFC and around the time of BLA. Generally, BLA showed weak responses to all task features except the choice. Decoding analyses suggested that errors followed from a failure to encode the target stimulus in both AC and dlPFC, but again, these differences arose earlier in AC. The similarities between AC and dlPFC responses were abolished during passive sensory stimulation with identical trial conditions, suggesting that the robust sensory encoding in dlPFC is contextually gated. Thus, counter to a strictly PFC-driven decision process, in this spatial selective listening task AC neural activity represents the sensory and decision information before dlPFC. Unlike in the visual domain, in this auditory task, the BLA does not appear to be robustly involved in selective spatial processing.

Key words: amygdala; auditory; auditory cortex; decision-making; prefrontal cortex; selective listening

Significance Statement

We examined neural correlates of an auditory spatial selective listening task by recording single-neuron activity in behaving monkeys from the amygdala, dorsolateral prefrontal cortex, and auditory cortex. We found that auditory cortex coded spatial cues and choice-related activity before dorsolateral prefrontal cortex or the amygdala. Auditory cortex also had robust delay period activity. Therefore, we found that auditory cortex could support the neural computations that underlie the behavioral processes in the task.

Introduction

Spatial selective listening is critical for solving everyday problems including the classic “cocktail party problem,” which requires attending to one sound source amid a noisy background of

competing sources (Cherry, 1953). Common auditory spatial selective listening paradigms used in research with humans include modified Posner paradigms in which subjects detect auditory stimuli after being cued to a spatial location (Spence and Driver, 1994; Alho et al., 1999; McDonald and Ward, 1999; Mayer et al., 2007, 2009; Roberts et al., 2009; Teshiba et al., 2013) and selective listening studies (Ahveninen et al., 2013; Frey et al., 2014; Bidet-Caulet et al., 2015). Previous work in humans has shown that auditory cortex (AC) plays an important role in spatial selective listening tasks, through interactions with prefrontal (Alho et al., 1999) and parietal (Deng et al., 2019) cortices. In addition to a role for these structures, previous studies in the visual domain in nonhuman primates have shown that the basolateral amygdala (BLA) contributes to spatial selective attention (Peck and Salzman, 2014; Costa et al., 2019).

There are only a few studies comparing multiple areas in auditory processes, especially in nonhuman primates, so we lack clear evidence on the relative contributions and timing of information between areas. Auditory processing is characterized by

Received Aug. 20, 2020; revised Nov. 2, 2020; accepted Nov. 26, 2020.

Author contributions: C.R.C. designed research; C.R.C., A.L.B., and J.J. performed research; M.M.M. and B.B.A. contributed unpublished reagents/analytic tools; J.L.N. and C.R.C. analyzed data; J.L.N., C.R.C., and B.B.A. wrote the paper.

*J.L.N. and C.R.C. contributed equally to this work.

This work was supported by the Intramural Research Program of the National Institute of Mental Health (ZIA MH002928). We thank Dr. Richard Krauzlis [National Eye Institute/National Institutes of Health (NIH)], Dr. Barbara Shinn-Cunningham (Carnegie Mellon), and Dr. Brian Scott [National Institute of Mental Health (NIMH)/NIH] for valuable input on task design and training; Dr. Richard Saunders (NIMH/NIH) for surgical assistance; and the NIH Section on Instrumentation for assisting in custom manufacture of recording chambers and grid.

The authors declare no competing financial interests.

Correspondence should be addressed to Corrie R. Camalier at corrie.camalier@duke.edu or Bruno B. Averbeck at bruno.averbeck@nih.gov.

<https://doi.org/10.1523/JNEUROSCI.2217-20.2020>

Copyright © 2021 the authors

Figure 1. Task design and recording locations. **A**, Structure of the spatial selective listening task. Cue conditions (listen left/right) were blocked with two types of trials (match/foil) in each condition. To begin each trial, the animal must depress a lever and maintain fixation at a central point on the screen. After a short delay (2.1–2.4 s), the animal heard a 4 kHz square-wave cue from a speaker on the left or right of its head. A continuous white noise was played 500 ms after the initial cue to make target detection difficult. In match trials, the animal heard a 1 kHz match target (various levels, see Materials and Methods) after some stimulus onset asynchrony (SOA; 500, 800, or 1300 ms) on the same side as the cue. If the animal released the lever within 700 ms of the match target, a fixed juice reward was delivered. In foil trials, the animal heard a 1 kHz foil target on the opposite side as the cue after the same SOA. The animal had to continue to hold down the lever until a 1 kHz match target was presented (after 800 or 1000 ms) on the same side as the cue. If the animal released the bar within 700 ms of the match target, a fixed juice reward was delivered. The “passive listening” control condition was identical to the active task except the monkey listened passively and did not press a lever, fixate, respond, or receive juice. **B**, Recording locations of single neurons across AC, dIPFC, and the BLA. Top, Patch of dIPFC recording area morphed to anatomic landmarks. Middle, AC grid coverage on region A1 of auditory cortex based on topography, latency, and frequency reversals. Bottom, Region of interest highlighted in blue—the entire left basolateral amygdala—targeted by V-Trodes. In all three areas, we selectively recorded from the left hemisphere of the animal. The bottom right image shows a structural MRI with contrast agent (betadine gel) in chamber grid holes for targeting. Recording locations and trajectories were further verified using tungsten electrodes inserted through grid locations to target areas. Yellow lines in each image show an approximate trajectory.

speed, especially relative to the visual system. In nonhuman primates, A1 has response latencies of ~ 20 ms (Camalier et al., 2012), compared with ~ 40 ms for primary visual cortex (Schmolsky et al., 1998). AC is, however, further removed from the peripheral sensory receptors than primary visual cortex. This speed is consistent with a hypothesized role for the auditory system in rapid spatial alerting or orienting. However, the processing depth of A1 has led some authors to suggest that it can also process cognitive factors such as choice, normally attributed to higher-order sensory areas (Näätänen et al., 2001; Nelken, 2004). Certainly, AC has been shown to reflect aspects of auditory decision-making beyond sensory processing (Niwa et al., 2012; Tsunada et al., 2016; Christison-Lagay and Cohen, 2018; Huang et al., 2019), but it is unclear whether this choice information is coming from prefrontal cortex (PFC) or another area (Lee et al., 2009; Plakke et al., 2015). A recent decision-making study in ferrets suggested that sensory information was encoded first in A1, but category information and the decision was encoded first in ferret dorsolateral PFC (dlPFC), which is a premotor area potentially analogous to primate PFC (Yin et al., 2020). This would be consistent with auditory working memory data in nonhuman primates which suggests that a categorical “match” decision may emerge earlier in ventral PFC than in AC (Bigelow et al., 2014). At present, the relative role of AC and dlPFC, especially in spatial decision-making, is unclear. Aside from the cortical sensory and prefrontal pathways, a BLA pathway for spatially selective

processing and decision-making is hypothesized to be fairly fast in the visual domain (Peck and Salzman, 2014; Costa et al., 2019), but whether the BLA is involved in auditory decision-making in nonhuman primates is unknown.

To address these outstanding questions, here we describe an experiment in which we used a spatial selective detection paradigm for monkeys, grounded in spatially cued listening tasks used in the human studies discussed above. To investigate potential neural correlates of this task, we recorded single-unit activity in primary AC (A1), dlPFC, and the BLA while the monkeys conducted the task. The dlPFC recordings were located in dorsal prearcuate cortex (primarily area 8A; see Materials and Methods; Fig. 1B), which is the primary prefrontal target of the auditory “dorsal stream” arising from caudal belt and parabelt, thought to be important for auditory spatial processing (Bon and Lucchetti, 1994; Hackett et al., 1999; Lanzilotto et al., 2013). Specifically, these recordings targeted the zone between the principal sulcus and dorsal arcuate sulcus, at least 1 mm away from the arcuate sulcus, primarily corresponding to area 8Ad, but also potentially including the dorsal bank of 46d, caudal 8Adv, and caudal border of 8b. Recordings in the amygdala targeted the basal and lateral nuclei. Cortical auditory inputs to the amygdala from the caudal parabelt terminate in the larger lateral nucleus (Yukie, 2002). However, the rostral superior temporal gyrus, which also indirectly receives auditory input, projects more broadly to the lateral and basal nuclei (Stefanacci and Amaral, 2002). We

Figure 2. Auditory target sound level and accuracy. Trial performance compared against the target sound level in decibels. Numbers below line indicate the percentage of trials across sessions at that sound level. Note, catch trials (0 dB) are not plotted, so percentages do not add to 1. Bars at each point represent the SE. Mean values were first calculated for each session, and then means were taken across sessions in which the indicated sound level was used. The SEM was calculated across sessions where the number of sessions were: Monkey 1: 16–20 dB ($N = 53$); 21–25 dB ($N = 19$); 26–30 dB ($N = 53$); 36–40 dB ($N = 1$; data not shown); Monkey 2: 26–30 dB ($N = 42$); 31–35 dB ($N = 42$); 36–40 dB ($N = 42$); 41–45 dB ($N = 9$).

examined the strength and latency of signals at the single-cell level related to the task across these areas. In AC and dlPFC, a substantial fraction of neurons was selective to the location of the cue and the subsequent target. We found that AC preceded both dlPFC and BLA in sensory discrimination and also in the decision. Classification analyses of firing rate patterns in error trials indicated that errors during the task were usually the result of a failure to encode the first target stimulus in AC and also in dlPFC. A comparison of responses and timing with a control “passive listening” condition showed that sensory target-related activity in dlPFC was almost completely abolished in the passive task, suggesting task-dependent gating of information to areas beyond sensory cortex.

Materials and Methods

The experiments were conducted using two adult male rhesus macaques (*Macaca mulatta*). The monkeys had access to food 24 h/d and earned their liquid through task performance on testing days. Monkeys were socially pair housed. All procedures were reviewed and approved by the National Institute of Mental Health Animal Care and Use Committee.

Experimental setup. The monkeys were operantly trained to perform a spatial selective listening paradigm. The task was controlled by custom software [System 3: OpenWorkbench and OpenDeveloper, Tucker-Davis Technologies (TDT)], which controlled multispeaker sound delivery and acquired bar presses and eye movements. Eye movements were tracked using the ViewPoint EyeTracker system (Arrington Research) sampled at 1 kHz. Monkeys were seated in a primate chair facing a 19 inch LCD monitor 40 cm from the eyes of the monkey, on which the visual fixation spot was presented. Monkeys performed the task in a darkened, double-walled, acoustically isolated sound booth (Industrial Acoustics). All auditory stimuli were presented from a speaker 10 cm from the left or right of the head of the monkey. Juice rewards were delivered using a solenoid juice delivery system (Crist Instrument).

Task design and stimuli. The monkeys conducted a spatial selective listening task (Fig. 1), modeled after spatially cued tasks used in humans. The task required oculomotor fixation throughout the duration of the trial. Both spatial cues and target stimuli were auditory, and the monkeys were required to respond when they detected a target embedded in

masking noise presented on the cued side. Listening conditions (listen left/right) were blocked with two types of trials (match/foil) in each condition. At the start of each trial, the monkey was prompted to press a lever and fixate on a central point on the screen. After a short delay (2.1–2.4 s), a 50 ms 4 kHz square-wave (70 dB) cue was played from a speaker on the left or right of the midline. Frozen diotic white noise (40 dB) was then played from both the left and right speakers from 500 ms after the initial cue until the lever was released. Following a variable delay after noise onset (500, 800, or 1300 ms), a 300 ms 1 kHz square-wave target sound was played from either the left or right speaker. If the target sound was on the same side as the cue, it was a match trial and the animal had to release the lever within 700 ms to receive a juice reward. If the target sound was on the opposite side from the cue, it was a foil trial and the monkey had to continue to hold the lever. Following a second interval of 800 or 1000 ms in foil trials, a second 1 kHz match target was always played on the same side as the original cue. If the animal correctly released the lever following the second target in foil trials, it was given a juice reward. Thus, both match and foil trials were identical in terms of reward expectation. If the choice was incorrect, there was a long “time-out” period before the next trial could be initiated. As in our previous work (Camalier et al., 2019), the use of square waves (which contain odd harmonics) allowed for wideband stimulation that was perceptually distinct, but whose broad spectral signature robustly activated large swaths of AC in a way that pure tones would not. Thus, similar to human paradigms, the stimuli could be kept identical across all sessions, independent of where recordings were conducted in AC, and data could be collapsed across sessions for analysis.

To achieve maximal effort and selective effects on neurons (as well as to be able to analyze sources of errors), it was important that the targets be difficult to detect. Thus, several psychometric quality controls were included to ensure that the monkeys were consistently performing the task across sessions. The sound level of the target for the two monkeys was individually titrated to maintain performance at ~70–80% correct (exact titration, 71.14%). Thus, the detection was difficult. The cue presentation was blocked to ensure the monkeys were able to maintain high accuracy on the task, as complex auditory tasks in monkeys have traditionally been difficult to condition operantly (Scott and Mishkin, 2016; Rinne et al., 2017). Analysis of the first trial after the cue switched sides showed that animals were correct 76.13% of the time, indicating that the monkeys were primarily using the cue in the task. For monkey 1, target

Figure 3. Example neurons. Left-hand panel shows rasters of single trials, and right-hand panel shows the p value from ANOVA for the indicated factor. Only correct trials are shown. The x -axis for p value plots shows the right-hand edge of the 300 ms bin used for ANOVA. **A**, Example neuron from cue epoch in AC. **B**, Example AC neuron showing responses to target. Plotted p values are for target factor. **C**, Example dIPFC neuron showing responses to cue. **D**, Example dIPFC neuron showing responses to target. Plotted p values are for cue $\times$ target interaction. **E**, Example BLA neuron showing responses to cue. **F**, Example BLA neuron showing response to cue $\times$ target interaction.

tones were delivered at levels between 16 and 40 dB, with most tones in the 17–24 dB range. For monkey 2, target tones were delivered at levels between 27 and 45 dB, with most tones in the 29–36 dB range. Within a session, the target sound varied 0–7 dB from trial to trial to ensure that

the monkeys were responding to the target side and not to consistencies in (or guessing based on) sound level differences between speakers that may have resulted from otherwise undetectable differences in calibration between the two speakers. To further ensure accurate performance,

Figure 4. ANOVA. Recording of single neurons from caudal AC (A1, lateral belt), dIPFC, and BLA while monkeys are performing a spatial selective listening task. A 2×2 factor ANOVA (cue side $\times$ target side, $p < 0.05$) using 300 ms bins sliding at 25 ms. The bin end point was used to align time on the x-axis (i.e., 300 ms is a bin from 0 to 300 ms). Bars above each plot represent the bins in which a statistically significant fraction of neurons encode each factor by color ($p < 0.01$; binomial test). **A, E, I**, During presentation of the cue, neurons respond differentially to the cue location. **B, F, J**, Postcue, a substantial fraction of neurons is selective to cue side, during the delay period, in both AC and dIPFC. **C, G, K**, Post-target presentation, a substantial portion of neurons in all three areas of interest are selective to the choice. **D, H, L**, Post-target presentation analysis including error trials shows choice encoding over and above cue $\times$ target interaction. Thus, choices are not a direct reflection of sensory input.

periodic “catch trials” (~10% with a 0 dB target tone) were included to ensure that the monkeys were responding to the target and not timing their choices relative to the presentation of the cue or noise. To encourage motivation during foil trials (which were of longer duration and were thus more likely to be aborted), the “match/bar release” target after a foil sound was louder (and easier) than typical target sounds for each monkey (monkey 1, 27 or 30 dB; monkey 2, 35 or 40 dB).

Before the task was run, a battery of passive listening and mapping stimuli were played. Within this battery was a control condition of “passive listening.” In this condition, the monkey was presented with the task stimuli, with trial types and timing matched to the selective listening task. However, the animals did not press or release a bar, fixate, or receive juice rewards. This task allowed us to compare sensory responses between active and passive task conditions. Monkeys were cued that this was a passive condition as they did not have access to the lever or juice tube, and it was done as part of a passive-listening mapping battery, consistently before the start of the active task.

Neurophysiological recordings. Monkeys were implanted with titanium headposts for head restraint before data collection began. Custom 45×24 mm acrylic chambers were designed and fitted to the monkeys in a separate procedure. The chamber was aligned with the long axis oriented anterior–posterior. The placement allowed vertical grid access to the left dorsolateral prefrontal cortex (Fig. 1B; dorsal bank of the principal sulcus extending to dorsal arcuate but at least >1 mm away from arcuate sulcus, primarily corresponding to area 46/8Ad, but also potentially including dorsal bank of 46d, caudal 8Adv, and caudal border of 8b), the basal and lateral portions of the amygdala (entire dorsoventral extent), and auditory cortex (primarily A1 but including small portions of surrounding areas). A 1 mm grid was located inside the chamber for targeting (Fig. 1B, bottom right), and all penetrations were dorsoventral. This dorsoventral trajectory was essential for targeting AC tonotopic reversals. The chamber was custom fit to a 3D print of the skull of each monkey generated by a computed tomography scan before

implantation. Recording areas were verified through a T1 scan of grid coverage with respect to underlying anatomic landmarks (Fig. 1B), combined with maps of frequency reversals and response latencies of single neurons to determine A1 location and extent (Camalier et al., 2012, 2019). Recordings were mainly conducted in simultaneous AC and PFC sessions, with BLA sessions occurring later in the experiment, but some data included were from just one or even three simultaneously recorded areas in a given session. We recorded the activity of 2387 single neurons during the task ($N = 847$ (AC), $N = 968$ (dIPFC), and $N = 572$ (BLA) across monkeys 1 ($N = 1540$) and 2 ($N = 847$).

In both monkeys, we recorded using either 16 or 24 channel laminar “V-Trodes” (Plexon; 200–300 μ m contact spacing, respectively). The electrodes allowed for identification of white matter tracts, further allowing the identification of electrode location with respect to sulci and gyri. To ensure that V-Trodes went as straight as possible, sharpened guide tubes for the buried structures (amygdala, AC) were advanced 10–15 mm above the structures. This was not possible for the PFC as it is a surface structure, but a guide tube was used to puncture overlying granulation tissue to permit a V-Trode to advance. Electrodes were advanced through the guide tubes to their target location (NAN Microdrives, NAN Instruments) and allowed to settle for at least 1 h before recording. Neural activity was recorded either primarily simultaneously (AC and PFC) or primarily individually (BLA), although there were some sessions in which all three areas were recorded from.

Multichannel spike and local field potential recordings were acquired with a 64-channel Tucker-Davis Technologies data acquisition system. Spike signals were amplified, filtered (0.3–8 kHz), and digitized at ~24.4 kHz. Spikes were initially sorted online on all channels using real-time window discrimination. Digitized spike waveforms and timestamps of stimulus events were saved for sorting offline (Offline Sorter version 3.3.5, Plexon). Units were graded according to isolation quality (single or multiunit neurons). Single and multiunit recordings were analyzed separately, but patterns were similar, so they

were combined. The acquisition software interfaced directly with the stimulus delivery system, and both systems were controlled by custom software (System 3: OpenWorkbench and OpenDeveloper controlling RZ2, RX8, TDT). For inclusion in analysis, cells had to be present for at least two blocks and 80 trials over the session.

Data analysis. For the initial ANOVA and poststimulus time histogram (PSTH) analysis, all trials on which monkeys released the lever in the correct interval were analyzed (71.14% of all trials). Trials in which the monkey answered incorrectly (28.86% of all trials) were excluded. The average number of correct trials analyzed for the ANOVA and PSTH analyses were 467.05 (AC, 480.11 trials; dIPFC, 471.72 trials; BLA, 439.81 trials). We performed a $2 \times 2 \times 5$ ANOVA (cue $\times$ target $\times$ sound level) on the activity of single neurons. The choice is given by the interaction in this ANOVA. The dependent variable was the firing rates of individual neurons. Trials in which the monkeys correctly released the lever within 700 ms of the target and were not catch trials (target 0 dB) were analyzed. The firing rate of each cell was computed in 300 ms bins advanced in 25 ms increments. We separated the analysis into three different segments of time, locked to the time surrounding the individual presentations of the cue, noise, and first target.

Next, we created a population PSTH for the firing rates of the individual neurons with respect to cue condition (left/right) and trial condition (match/foil). For this analysis, the firing rate of each cell was computed in 1 ms bins and smoothed with a three-bin moving average. Data are plotted using 25 ms bins, but *t* tests, to determine onset latencies, were computed on the 1 ms bins.

For the decoding analyses, we separately analyzed correct and error trials. A trial was considered correct if the monkey released the lever after the presentation of the appropriate target within 700 ms. All other trials were deemed incorrect. The average number of error trials analyzed for decoding was 123.53 (AC, 127.57 trials; dIPFC, 136.99 trials; BLA, 94.77 trials). For the decoding analysis, the firing rate of each cell was computed in 100 ms bins and advanced in 25 ms increments. Decoding analyses were performed using leave-one-out cross-validation to predict which observations belong to each trial condition using the SVM classifier in MATLAB. All decoding was done using pseudopopulations composed of all neurons recorded from a structure across all sessions. Trials were assigned randomly from the different sessions within each condition.

For the ANOVAs, we used 300 ms bins, as this provided additional sensitivity to detect significant effects in neurons with low firing rates. We followed this up with the population analysis, which used 1 ms bins, to optimize the detection of onset latencies. Finally, we used 100 ms bins for the decoding analysis because the large number of neurons used in this analysis increases the signal-to-noise ratio for detecting effects, and therefore a smaller bin than was used for the ANOVA allows us to detect timing effects more accurately.

For the decoding analyses, we calculated significant differences between correct and error trials using a bootstrap analysis (Efron and Tibshirani, 1998). We generated data according the null hypothesis that

Figure 5. Median reaction time split ANOVA. Recording of single neurons from AC, dIPFC, and BLA while monkeys were performing the task. The results are from a 2×2 ANOVA (cue side $\times$ target side, $p < 0.05$) using 300 ms bins sliding at 25 ms. Only the interaction term (response) is plotted on the graph. Bin end point was used to align time on the x-axis. The trials for all neurons were split into fast and slow reaction times by the median reaction time within a session, and separate ANOVAs were run for each set of trials. Paired *t* tests ($p < 0.01$; consecutive bins, ≥ 3) were computed to determine significance between the fraction of significant neurons assessed in each reaction time split. Bars above each plot represent the bins in which a statistically significant difference was seen between response to trials with fast reaction times versus slow reaction times. **A**, Data for AC. **B**, Data for dIPFC. **C**, Data for BLA. Sig., Significant; Diff, different; Avg., average.

there were no differences between correct and error trials. We did this by sampling with replacement, from the combined set of correct and error trials, sets of bootstrap correct and error trials. Both the null correct and error bootstrap sets contained combinations of correct and error trials. We then conducted the decoding analysis using the bootstrap trials to determine the decoding accuracy when correct and error trials were mixed. We did this 1000 times. We calculated the difference in fraction correct between correct and error trials in each time bin, for each set of bootstrap trials. This gave us 1000 differences sampled from the null distribution between correct and error trials in each time bin. We then compared the difference in the actual data to the differences in the null distribution and computed a *p* value, which was the relative rank of the true difference in the null distribution samples. That is to say, if the true difference was larger than, for example, 986 samples in

Figure 6. Contribution of individual neuron selectivity to the population representation. A 2×2 ANOVA ($p < 0.05$) using 300 ms bins sliding at 25 ms conducted on each individual neuron. Neurons are plotted along the y-axis and the time is on the x-axis. Bin end point was used to align time on the x-axis. Colored bars represent the times in which a neuron was significant for that task factor. Blue column displays encoding of the cue during the cue presentation, gray column shows the cue encoding during the delay period, red column shows the target encoding during target presentation and yellow column shows the response encoding during the target period. **A**, Data for neurons recorded in AC. **B**, Data for neurons recorded in dlPFC. **C**, Data for neurons recorded in BLA.

the null distribution, it was significant with a two-sided p value of $2 \times (1000 - 986) / 1000 = 0.028$.

Results

Task and behavior

We recorded neural activity from two monkeys while they conducted a spatial selective listening task (Fig. 1A). At the start of each trial, the monkeys acquired central fixation (Fig. 1A) and pressed a bar. After a baseline hold period, an auditory stimulus (the cue) was presented from a speaker on the left or right of the monkey. After the cue, there was a delay period during which white noise was played, continuing until bar release. Following the delay period, a second target stimulus was presented on the same (match) or opposite (foil) side as the cue, at different sound levels (Fig. 2). The monkeys were trained to release the bar if the cue and target stimulus were on the same side (match trials) and continue to hold if they were not on the same side (foil trials). In foil trials, following a second delay after the target stimulus, a third match target was played that was always on the same side. In match trials, the mean response time was 374.9 ms (SD, 27 ms). Monkey 1 had a slightly faster mean response time of 358.3 ms (SD, 11.5 ms) and monkey 2 had a mean response time of 405.1 ms (SD, 19.9 ms).

Single-cell encoding of task factors

While the animals conducted the task, neural activity was recorded (Fig. 1B), from the following three areas: AC ($N = 847$), dlPFC ($N = 968$), and the BLA ($N = 572$). We found neurons in all structures that responded to the presented cues (Fig. 3A,C,E) and the targets or the interaction of cue and target (Fig. 3B,D,F). We assessed the encoding of each task factor in single neurons across the population by carrying out ANOVAs on correct trials,

Table 1. Neuron selectivity to task variables by brain area

Also encoded	Encoding cue			Encoding delay			Encoding target response
	Delay	Target	Response	Target	Response		
AC	26.70	17.05	5.40	16.90	8.81	5.83	
dlPFC	30.90	10.11	4.49	9.80	3.92	5.80	
BLA	25.00	8.33	2.08	5.77	7.69	8.00	

All results displayed are percentages. A 2×2 ANOVA ($p < 0.05$) was performed using one bin per time period. A single time bin was chosen at the peak of the ANOVA curves (Fig. 4). Cue presentation included 150–450 ms postcue, the delay period included 100–400 ms postnoise, target presentation included 50–400 ms post-target presentation, and the choice period included 200–500 ms post-target presentation. Delay indicates cue encoding during delay.

for each single neuron. With the ANOVA, we examined the effects of cue location, target location, target sound level, and interactions (cue $\times$ target codes decision) using spike counts in a 300 ms window, advanced by 25 ms (Fig. 4). During the cue period, we found that activity discriminated cues rapidly in AC (Fig. 4A). In dlPFC, activity discriminated cues as well, but the effect increased slowly (Fig. 4E). The BLA, however, showed minimal cue discriminative activity, with the number of neurons coding cue location only slightly above chance (Fig. 4I). Note that cue location trials were blocked in the task, which led to small baseline, statistically significant, elevation of cue side encoding before cue presentation. Although the cue side was blocked, performance on the first trial after the cue switched sides was 76.13%, and therefore the animals were attending to the cue. Although encoding peaked in AC and dlPFC following the cue, elevated cue discrimination was maintained during the delay interval, which was not affected by the white noise, in both AC and dlPFC. The BLA showed less delay period activity.

When the target stimulus was presented, it was rapidly and robustly encoded in AC (Fig. 4C). The dlPFC also encoded the

Figure 7. PSTHs. Mean normalized firing rates of neurons plotted using nonoverlapping 25 ms bins smoothed with a 3-point moving average. Only neurons significant for the corresponding factor (i.e., cue, target, or cue $\times$ target) were included in this analysis. The bin midpoint was used to align time on the x-axis. Analysis was conducted to assess the precise timing of changes in neuronal firing rates in AC, dIPFC, and BLA. Paired *t* tests were performed on all bins to determine significant differences in firing rates. Bootstrapping analysis was performed to directly compare timing differences in different brain areas. Vertical bars with stars indicate nonoverlapping 95% confidence intervals for discrimination times between areas. **A, C, E**, Comparison of conditions that are identical in cue side but vary in target side, as a measure of sensory identification. **B, D, F**, Comparison of conditions that are identical in target side but vary in the cue location (left or right). **G, I, K**, Conditions are matched for cue side but vary in target side. **H, J, L**, Conditions shown have opposite cue sides but matched target side.

target stimulus (Fig. 4G), although later than AC, which would be expected. The BLA only weakly encoded the target stimulus and only at about the time of the choice (Fig. 4K). The cue $\times$ target interaction, which defined the choice in correct trials, was encoded first in AC (Fig. 4C), after which it was encoded in dIPFC (Fig. 4G). The cue $\times$ target interaction, unlike the cue and target locations, was robustly encoded in the BLA (Fig. 4K). Sound level was also robustly encoded in AC (Fig. 4C) and less robustly in dIPFC (Fig. 4G) and BLA (Fig. 4K).

We also followed up this ANOVA with an additional ANOVA that included both correct and error trials. This allowed us to dissociate the choice from the sensory processing reflecting the cue $\times$ target interaction. When we conducted this analysis, we found that the choice was more robustly encoded than the cue $\times$ target interaction across all areas (Fig. 4D,H,L), and most of the cue $\times$ target interaction could be accounted for with the choice variable. Overall, all variables, including the delay period activity and the choice, were encoded first and most robustly by AC. The dIPFC did encode all task factors, but after AC. The BLA showed only weak encoding of the cue and the target but robustly encoded the choice.

In the next analysis, we compared encoding of the choice in fast and slow reaction time trials, to see whether encoding of the choice (i.e., the interaction between cue side and target side in the ANOVA) differed (Fig. 5). We performed a median split using the reaction times for all trials, both match and nonmatch, within a session. For nonmatch trials, we used the release time (RT) after the second target as the RT. ANOVAs were run on each neuron twice, once on trials below the median reaction

time, and once on trials above the median reaction time. We found, in all three areas, that the choice was encoded faster when the animals responded quickly than when the animals responded slowly. Only in auditory cortex did the activity related to the choice diverge before the average of the fast reaction times (Fig. 5A). In both dIPFC and the BLA, the activity diverged just before or after the average fast reaction time. Thus, the choice variable from the ANOVA depends on the timing of the motor response and is not completely determined by the timing of the auditory cues.

The results from the ANOVAs show the contribution of the neurons to each task factor. However, they do not illustrate whether single neurons code multiple task factors through time. Therefore, we also examined whether single neurons encoded more than one task factor during each epoch (Fig. 6). It could be seen that many neurons coded more than one task factor and coded cue, for example, in both the cue and delay periods.

To further quantify whether neurons encoded more than one variable, we also estimated the fraction of neurons that encoded multiple factors using a single representative bin for each factor, centered on the time at which the population encoding of each factor peaked (Table 1). Most often, neurons that encoded the cue during the cue presentation continued to encode the cue during the delay interval. In AC, of the neurons that encoded the cue during cue presentation, 26.70% of them also encoded the cue during the delay period. Neurons in dIPFC were most selective to encoding the cue during the cue presentation and through the delay interval, with an overlap of 30.90%. In the BLA, 25.00%

Figure 8. Classification analysis to cue location factor comparing correct and error trials. Analysis was performed using 100 ms bins, sliding at 25 ms. The bin end point was used to align time on the x-axis. Analysis performed using leave-one-out cross-validation to predict which observations belong to each cue condition. Bootstrap test performed with 1000 pseudorandom samples. No significant difference in classification rates was found between correct and error trials in any brain region during any time bin. **A–C**, data from AC relative to cue, noise and target onset. **D–F**, data from dlPFC relative to cue, noise and target onset. **G–I**, data from BLA relative to cue, noise and target onset.

of the neurons encoded the cue during both time periods. Neurons that encoded the cue also often eventually encoded the target, with an overlap of 17.05%, 10.11%, and 8.33% in AC, dlPFC, and BLA, respectively. Most interestingly, while a relatively small portion of neurons encoded the cue through the delay period as well as eventually encoding the target in dlPFC and BLA, AC did this with an overlap of 16.9%. Neurons in AC were most likely to continue to encode other task variables, compared with the dlPFC and the BLA.

Next, we examined finer time-scale encoding of several task factors. The ANOVA used relatively large time windows to calculate sensitive statistics on potentially low-firing rate neurons. These time windows, however, do not allow the determination of precise onset times for task factors. To characterize onset times at a finer time scale, we calculated PSTHs using 1 ms time windows, smoothed with a 3-point moving average, for each neuron (plotted using 25 ms bins; Fig. 7). We then conducted *t* tests ($p < 0.01$, uncorrected) in each bin to estimate the time at which the population in each area discriminated between conditions. We found that the cue was discriminated in AC at 25 ms (Fig.

7B) and in dlPFC at 65 ms (Fig. 7D) after stimulus onset. Using these small bins, the population of BLA neurons did not discriminate cue side, likely because of low firing rates (Fig. 7F). The target was discriminated in AC at 36 ms (Fig. 7G), in dlPFC at 169 ms (Fig. 7I) and in the BLA at 185 ms (Fig. 7K) after tone onset. Finally, the decision was discriminated in AC at 146 ms (Fig. 7H), in dlPFC at 321 ms (Fig. 7J), and in BLA at 266 ms (Fig. 7L) after target onset.

Next, we used a bootstrap analysis to determine whether onset latencies differed significantly between areas (Fig. 7). We pulled samples of 100 neurons for each brain area and computed the time at which the two conditions diverged ($p < 0.05$; consecutive bins, ≥ 6) in each bootstrap sample. We did this 100 times to create a sample distribution. We then calculated a 95% confidence interval for the discrimination times for each area. If the confidence intervals overlapped, the brain areas were not deemed statistically significant. From this analysis, AC preceded both dlPFC and BLA in cue and target discrimination, and AC preceded dlPFC in decision discrimination. AC, however, did not statistically precede BLA in decision discrimination.

Figure 9. Classification analysis to target factor by correct or error. Analysis was performed using 100 ms bins, sliding at 25 ms. The bin end point was used to align time on the x-axis. Gray shaded areas represent timepoints where correct and error classification rates differ ($p < 0.01$, bootstrap). **A–C**, data from AC relative to cue, noise and target onset. **D–F**, data from dIPFC relative to cue, noise and target onset. **G–I**, data from BLA relative to cue, noise and target onset.

Decoding correct and error trials

In the next analyses, we used decoding to examine error trial activity. We were interested in which processes broke down in error trials. To examine this, we used leave-one-out cross-validation on pseudopopulations (see Materials and Methods) to predict, using the neural activity, the side on which the cue was presented (Fig. 8), the side on which the target was presented (Fig. 9), and the choice (Fig. 10). The decoding model was first estimated using only correct trials. We then classified the error trials using the decoding model estimated on correct trials to see whether neural activity in error trials represented the stimuli that were presented and the choice that was made. We found that in correct and error trials the neural population in both AC and dIPFC rapidly predicted the cue location (Fig. 8A,D), and maintained prediction through the delay interval (Fig. 8B,E), consistent with the single-neuron results. The BLA did not discriminate clearly the cue side (Fig. 8G). There were no significant differences between correct and error trials for cue encoding, and this finding was consistent through the delay interval. Therefore, the cue was correctly encoded in error trials.

When we decoded the target side using neural activity, we found that in correct trials the target location was robustly predicted by AC (Fig. 9C) and dIPFC (Fig. 9F). There was minimal prediction of the target in the BLA (Fig. 9I). In error trials, however, the target

was not well predicted by any of the areas (Fig. 9). The correct and error trial predictions diverged ($p < 0.01$ bootstrap) 75 ms after target onset in AC and 125 ms after target onset in dIPFC.

In error trials, animals either released when they should not have or did not release when they should have. When we predicted the choice, relative to what the monkeys should have done, we found an accurate prediction in correct trials in all three areas (Fig. 10C,F,I). Furthermore, in error trials, the predicted choice tended to fall below chance, which indicates that the neural activity is coding the choice the monkey made in error trials, as opposed to the choice the monkey should have made. However, this coding was only significantly below chance late in the choice period in AC (Fig. 10C). We used a smaller time bin step of 5 ms in the rightmost column (Fig. 10C,F,I) to more precisely determine the point at which the curves diverged. Consistent with the other analyses, we found that predictions in error and correct trials diverged statistically in auditory cortex (270 ms after target onset) and subsequently in dIPFC and BLA (275 and 300 ms after target onset).

Next, we examined the position of the population neural activity relative to the discrimination boundary, extracted from the decoding model. For the decoding analysis (Figs. 8–10), this quantity is thresholded in each trial and time bin, and the time bin in that trial is classified as either (e.g., cue left or cue right),

Figure 10. Classification analysis to choice. Gray bar indicates timepoints where correct and error classification rates differ, and red bar indicates timepoints where error trials were significantly below chance (0.5). The bin end point was used to align time on the x-axis. The reaction time for detect trials is shown as a dotted line, with the SD shaded, in **C**, **F**, and **I**. **C**, **F**, **I**, Differences between correct and error trials in AC were from 270 to 500 ms (**C**), in dlPFC from 275 to 500 ms (**F**), and in BLA from 300 to 500 ms (**I**). **A–C**, data from AC relative to cue, noise and target onset. **D–F**, data from dlPFC relative to cue, noise and target onset. **G–I**, data from BLA relative to cue, noise and target onset.

depending on whether the position is positive or negative. However, the average distance to the decoding boundary provides a continuous estimate of how well the population discriminated the conditions versus time (Fig. 11). In general, these analyses were consistent with the thresholded decoding analysis. Cue-related activity diverged in correct and error trials, reflecting the cued side, and the activity in error trials matched the activity in correct trials (Fig. 11A,D,G). The breakdown in activity following target presentation could also be seen (Fig. 11B,E,H). However, there was some maintained coding of the target, particularly in auditory cortex (Fig. 11B), which may also be reflected in the decoding accuracy in error trials (Fig. 9C). Therefore, cue encoding is intact in error trials and target encoding is mostly but not completely absent. The choice-encoding dynamics did reflect the fact that the wrong choice tended to be predicted by population activity (Fig. 11C,F,I). However, it could be seen that the activity diverged less than it did in correct trials, consistent with the lower decoding performance.

Neural responses in the passive task

In a final series of analyses, we analyzed data from a passive task, collected in each session before the main, active task data. The

sensory stimulation in the passive task was identical to the stimulation in the active task, except the animals did not press a bar to initiate a trial, they did not release the bar to indicate their choice, and there was no juice tube so they could not be rewarded. When we examined encoding of cue location, we again found robust coding in AC (Fig. 12A). All of the other signals, however, were much weaker. The cue responses in dlPFC dropped from a peak near 30% in the active task to ~10% in the passive task (Fig. 12D). Interestingly, there was delay activity in the passive task in AC (Fig. 12B), perhaps because the animals were highly overtrained. The delay activity in dlPFC was reduced from ~20% of the population to ~10% (Fig. 12E). There was also a small amount of target encoding in AC (Fig. 12C). Target encoding in dlPFC did not exceed chance (Fig. 12F). Encoding in the BLA only sporadically exceeded chance, perhaps because of type I errors, or low-level encoding (Fig. 12G–I).

We also examined onset times using small time bins (Fig. 13). We found differences in responses in AC that depended on the side of the stimulus for the cue at 32 ms (Fig. 13B) and for the target at 53 ms (Fig. 13G). However, we did not detect population-level differences in responses, using these small time bins, in dlPFC or BLA, which suggests that responses that reached

Figure 11. Distance to classification boundary derived from support vector machine classifier. Analyses were conducted using 200 ms bins sliding at 25 ms. The bin end point was used to align to the x-axis. Error trials were defined as trials in which the animal responded incorrectly or responded outside of the allotted reaction time window. Correct trials are presented in shades of blue, and error trials in shades of red. **C, F, and I.** The reaction time for detect trials is shown as a dotted line, with the SD shaded. **A, D, G.** Conditions were separated by whether a trial was cued on the left or right side of the animal. **B, E, H.** Conditions were separated by whether the first target was presented on the left or right side of the animal. **C, F, I.** Conditions were separated by whether there was a response or no response made by the animal (i.e., if it was a detect or foil trial, respectively). For error trials, conditions were separated by whether there should have been a response or no response, regardless of what the animal chose to do.

significance in the ANOVAs were driven by low firing rates. Overall, beyond cue encoding in AC, responses across all three areas were reduced in the passive task, relative to the active task.

Discussion

We trained monkeys on a selective listening task, based on tasks used in humans. The task required animals to detect a difficult-to-discriminate auditory stimulus, embedded in white noise. We found that AC encoded cues, targets, and decisions before either dlPFC or BLA. In addition, AC had delay activity that coded the location of the initial cue. It is not clear, however, whether the AC delay activity depended on dlPFC delay activity, or even parietal activity that we did not record. Activity in dlPFC closely followed activity in AC. The BLA, on the other hand, only minimally encoded cue and target activity. The BLA was strongly engaged, however, at the time of choice, although the choice-related activity followed activity in AC. Therefore, the AC appears to support many of the functions required for auditory selective listening. This is in contrast to early visual areas, which

represent visual features, but play a minimal role in decision-making aspects of tasks (Britten et al., 1992; Zaksas and Pasternak, 2006).

Previous work has shown that AC neurons can encode non-sensory, choice-related activity (Niwa et al., 2012; Christison-Lagay and Cohen, 2018; Huang et al., 2019). The study by Huang et al. (2019) found that whether a choice was predictable following a cue tone, based on the task condition, affected neural responses in AC to the tone. Therefore, AC encoded whether the response was determined by the first cue. Our results are consistent with this and other studies (Christison-Lagay and Cohen, 2018), in that we show that auditory cortex encodes the necessary response. However, in our task, the choice was not determined by the first cue, so choice-related activity only followed the target. Our paradigm does not allow us to dissociate decision-making from the motor response required to indicate the decision, and therefore our choice coding could be related to either, though note that it begins well before the reaction time of ~400 ms. We also show that encoding in AC precedes encoding in dlPFC, and we dissociated through our fully crossed experimental design,

Figure 12. Passive task data. ANOVA of data from passive condition, which was identical to the task, except the monkey was simply required to sit passively and listen to the sounds of the task structure. Bars above each plot represent the bins in which a statistically significant fraction of neurons encode each factor by color ($p < 0.01$). **A, D, G**, During presentation of the cue, neurons respond differentially to the cue location. **B, E, H**, Postcue, neurons are selective to cue side, during the delay period, in both AC and dlPFC, but weakly, compared with the active task. **C, F, I**, Post-target presentation, only AC encodes the target, and none of the areas encode the response.

encoding of cue location and target location, and the required response. Although it is possible that AC inherits response encoding from a cortical area other than dlPFC, the anatomic organization of this system suggests that it would have to be a nearby area, for example belt or parabelt auditory cortex (Romanski and Averbeck, 2009; Kajikawa et al., 2015; Tsunada et al., 2016). Given that AC is deeper into the neural processing stream than, for example, primary visual cortex (Mizrahi et al., 2014), it is also possible that AC could have sufficiently sophisticated mechanisms to compute the required response locally. Though AC precedes PFC in the encoding of the decision in both correct and error trials, the responses across areas are also quite similar within this task (~50 ms differences). This tight temporal relationship between AC and dlPFC is contextually dependent. When responses during the passive condition were analyzed, the fraction of responsive neurons was reduced, and responses were later in all areas relative to the task-related responses (and BLA was completely unresponsive, consistent with a primary role in reward-guided behavior). Particularly, dlPFC showed a reduction of responses to the cue and delay activity and an abolishment of target-related activity compared

with the active task condition. This is consistent with data from the same animals and areas during a passive oddball task in which dlPFC activity was later (~100 ms) and weaker than in AC (Camalier et al., 2019). Together, it suggests that the strength and timing of the information transfer between AC and dlPFC can be flexibly allocated and are dependent on task demands. Last, comparison of the active and passive conditions highlights the sustained nonsensory motor/reward related activity in “primary” sensory cortex (AC; Knyazeva et al., 2020).

Several of the analyses show that the neural responses recorded in this task were not straightforward sensory responses to the auditory stimuli. This was true across areas. For example, we found that the cue × target interaction, which defines choices in correct trials, was less strongly encoded than the choices, when both correct and error trials were analyzed. In addition, when we split trials into those with fast and slow reaction times, we found that the neural representation of the decision was coded earlier when choices were made earlier. We also saw that much of the task-related neural activity was reduced, although not eliminated, in the passive condition, when animals did not have to respond to the sensory cues.

Figure 13. Passive data; PSTHs. Mean normalized firing rates of neurons plotted using nonoverlapping 25 ms bins smoothed with a 3-point moving average. The bin midpoint was used to align time on the x-axis. Analysis was conducted to assess precise timing of changes in neuronal firing rates in AC, dlPFC, and BLA. Paired *t* tests were performed on all bins to determine significant differences in firing rates. **A, C, E**, Comparison of conditions that are identical in cue side but vary in target side, as a measure of sensory identification. **B, D, F**, Comparison of conditions that are identical in the target side but vary in the cue location (left or right). **G, I, K**, Conditions are matched for cue side but vary in target side. **H, J, L**, Conditions shown have opposite cue sides but matched target sides.

Both prefrontal (Green et al., 2011; Bidet-Caulet et al., 2015) and parietal (Michalka et al., 2016; Deng et al., 2019, 2020) cortex have been shown to play important roles in auditory spatial attention in humans. AC has also been shown to have attention-selective modulation of single neurons when targets and distractors are separated by frequency content (Atiani et al., 2009; Schwartz and David, 2018; O'Sullivan et al., 2019). Although we found clear responses related to the cued side in dlPFC, they followed AC. This was true of not only the sensory responses, but also the decision response. From our data, it is not, however, possible to determine whether the delay period activity, which may represent sustained attention/working memory for the cue location, was sustained by AC, dlPFC, or their interaction. In addition, several of the spatial attention paradigms used in the human work required participants to attend or discriminate sounds in one location, while ignoring sounds on the contralateral side (Deng et al., 2019). It is possible that if we had required the monkeys to carry out complex perceptual discriminations at one location, while ignoring distractors at another location, we would have found stronger engagement of dlPFC. We did use a white masking noise following the cue signal to examine its effects on behavior and neural representations of the cue location. Although we did see some effects of the noise onset in the decoding analysis, effects that were stronger in AC than dlPFC, they were transient and resulted in increased decoding accuracy for the cued location. The increased accuracy may have followed from an overall increase in neural activity, which may have improved decoding performance. Also, we did not record neural activity in parietal cortex, which may also play a role in the sustained delay period activity, although it would be interesting to consider inferior parietal cortex in future studies.

We found that the BLA played little role in encoding the cue location, and responses related to the choice followed responses in AC. This is inconsistent with previous reports of the involvement of BLA in visual-spatial attention (Peck et al., 2013). In these tasks, the amygdala neurons encoded the valence of stimuli, which were saccade targets, during delay periods (Peck and Salzman, 2014). There are several differences between these tasks and ours, however. For example, the tasks used in the study by Peck et al. (2013) were based on visual-spatial paradigms instead of an auditory-spatial paradigm, and they also required eye movements to spatial locations. Although the BLA receives auditory inputs (Yukie, 2002), these inputs may play a smaller role in the primate than they do in rodents (Munoz-Lopez et al., 2010). In rodents, auditory cues can be associated with shock in pavlovian fear conditioning (Romanski and LeDoux, 1992). These studies have shown that the amygdala plays an important role in the associative process between cues and shocks. Although, the amygdala is also involved in reward-guided behavior (Costa et al., 2016, 2019; Averbach and Costa, 2017). We did find a small, although significant, population of amygdala neurons that encoded the auditory cue and the auditory target. They did so, however, at long latencies. Therefore, the BLA appears to play a minimal role in the cognitive process of selective listening under reward-constant trials in highly trained animals. It is, however, possible that if we had primarily recorded from the lateral nucleus, which receives most of the direct auditory inputs (Yukie, 2002), we would have found more neurons related to aspects of our task.

The present study also shows a substantial dissociation of function between the BLA and dlPFC. This dissociation differs from the similarity between these structures seen in

reinforcement learning (RL) tasks, in which both dlPFC and BLA show substantial encoding of the identity of visual stimuli, the reward values associated with those stimuli, and reward outcomes (Costa et al., 2019; Bartolo et al., 2020). The primary difference between the BLA and dlPFC, in RL tasks, is that the dlPFC strongly encodes the direction of eye movements required to saccade to a rewarding visual stimulus (Bartolo et al., 2020), whereas the BLA encodes eye movement directions only at a low level (Costa et al., 2019). Thus, in RL tasks, the BLA and dlPFC show similar responses, which are also similar to those seen in the ventral striatum (Costa et al., 2019) and orbitofrontal cortex (Costa and Averbeck, 2020), with which the BLA is monosynaptically connected. The current study, however, shows that in cognitive, auditory selective listening tasks, the BLA and dlPFC show different responses, until the animal makes a reward-guided choice.

In conclusion, we found that AC encoded cues, targets, and decisions, before dlPFC, in an auditory selective listening task. We also found that AC had delay period activity. The BLA had minimal cue or target activity, although it did encode decision activity. The decision-related activity in the BLA, however, followed decision related activity in AC. Overall, this suggests that AC may carry out most important computations relevant to auditory selective listening. The main caveat is that it is not possible to determine whether delay period activity, which likely critically underlies performance in this task, is supported by AC in the absence of dlPFC or parietal cortex. Future work, for example inactivating dlPFC and/or parietal cortex (Plakke et al., 2015), while recording in AC, could clarify this question.

References

- Ahveninen J, Huang S, Nummenmaa A, Belliveau JW, Hung AY, Jääskeläinen IP, Rauschecker JP, Rossi S, Tiitinen H, Raij T (2013) Evidence for distinct human auditory cortex regions for sound location versus identity processing. *Nat Commun* 4:2585.
- Alho K, Medvedev SV, Pakhomov SV, Roudas MS, Tervaniemi M, Reinikainen K, Zeffiro T, Näätänen R (1999) Selective tuning of the left and right auditory cortices during spatially directed attention. *Brain Res Cogn Brain Res* 7:335–341.
- Atiani S, Elhilali M, David SV, Fritz JB, Shamma SA (2009) Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron* 61:467–480.
- Averbeck BB, Costa VD (2017) Motivational neural circuits underlying reinforcement learning. *Nat Neurosci* 20:505–512.
- Bartolo R, Saunders RC, Mitz A, Averbeck BB (2020) Dimensionality, information and learning in prefrontal cortex. *PLoS Comput Biol* 16:e1007514.
- Bidet-Caullet A, Buchanan KG, Viswanath H, Black J, Scabini D, Bonnet-Brihault F, Knight RT (2015) Impaired facilitatory mechanisms of auditory attention after damage of the lateral prefrontal cortex. *Cereb Cortex* 25:4126–4134.
- Bigelow J, Rossi B, Poremba A (2014) Neural correlates of short-term memory in primate auditory cortex. *Front Neurosci* 8:250.
- Bon L, Lucchetti C (1994) Ear and eye representation in the frontal cortex, area 8b, of the macaque monkey: an electrophysiological study. *Exp Brain Res* 102:259–271.
- Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci* 12:4745–4765.
- Camalier CR, D'Angelo WR, Sterbing-D'Angelo SJ, de la Mothe LA, Hackett TA (2012) Neural latencies across auditory cortex of macaque support a dorsal stream supramodal timing advantage in primates. *Proc Natl Acad Sci U S A* 109:18168–18173.
- Camalier CR, Scarim K, Mishkin M, Averbeck BB (2019) A comparison of auditory oddball responses in dorsolateral prefrontal cortex, basolateral amygdala, and auditory cortex of macaque. *J Cogn Neurosci* 31:1054–1064.
- Cherry EC (1953) Some experiments on the recognition of speech, with one and with two ears. *J Acoust Soc Am* 25:975–979.
- Christison-Lagay KL, Cohen YE (2018) The contribution of primary auditory cortex to auditory categorization in behaving monkeys. *Front Neurosci* 12:601.
- Costa VD, Averbeck BB (2020) Primate orbitofrontal cortex codes information relevant for managing explore-exploit tradeoffs. *J Neurosci* 40:2553–2561.
- Costa VD, Dal Monte O, Lucas DR, Murray EA, Averbeck BB (2016) Amygdala and ventral striatum make distinct contributions to reinforcement learning. *Neuron* 92:505–517.
- Costa VD, Mitz AR, Averbeck BB (2019) Subcortical substrates of explore-exploit decisions in primates. *Neuron* 103:533–545.e5.
- Deng Y, Reinhart RM, Choi I, Shinn-Cunningham BG (2019) Causal links between parietal alpha activity and spatial auditory attention. *Elife* 8:e51184.
- Deng Y, Choi I, Shinn-Cunningham B (2020) Topographic specificity of alpha power during auditory spatial attention. *Neuroimage* 207:116360.
- Efron B, Tibshirani RJ (1998) An introduction to the bootstrap. New York: Chapman & Hall.
- Frey JN, Mainy N, Lachaux JP, Müller N, Bertrand O, Weisz N (2014) Selective modulation of auditory cortical alpha activity in an audiovisual spatial attention task. *J Neurosci* 34:6634–6639.
- Green JJ, Doesburg SM, Ward LM, McDonald JJ (2011) Electrical neuroimaging of voluntary audiospatial attention: evidence for a supramodal attention control network. *J Neurosci* 31:3560–3564.
- Hackett TA, Stepniewska I, Kaas JH (1999) Prefrontal connections of the parabelt auditory cortex in macaque monkeys. *Brain Res* 817:45–58.
- Huang Y, Heil P, Brosch M (2019) Associations between sounds and actions in early auditory cortex of nonhuman primates. *Elife* 8:e43281.
- Kajikawa Y, Frey S, Ross D, Falchier A, Hackett TA, Schroeder CE (2015) Auditory properties in the parabelt regions of the superior temporal gyrus in the awake macaque monkey: an initial survey. *J Neurosci* 35:4140–4150.
- Knyazeva S, Selezneva E, Gorkin A, Ohl FW, Brosch M (2020) Representation of auditory task components and of their relationships in primate auditory cortex. *Front Neurosci* 14:306.
- Lanzilotto M, Perciavalle V, Lucchetti C (2013) Auditory and visual systems organization in Brodmann Area 8 for gaze-shift control: where we do not see, we can hear. *Front Behav Neurosci* 7:198.
- Lee JH, Russ BE, Orr LE, Cohen YE (2009) Prefrontal activity predicts monkeys' decisions during an auditory category task. *Front Integr Neurosci* 3:16.
- Mayer AR, Harrington DL, Stephen J, Adair JC, Lee RR (2007) An event-related fMRI Study of exogenous facilitation and inhibition of return in the auditory modality. *J Cogn Neurosci* 19:455–467.
- Mayer AR, Mannell MV, Ling J, Elgie R, Gasparovic C, Phillips JP, Doezema D, Yeo RA (2009) Auditory orienting and inhibition of return in mild traumatic brain injury: a FMRI study. *Hum Brain Mapp* 30:4152–4166.
- McDonald JJ, Ward LM (1999) Spatial relevance determines facilitatory and inhibitory effects of auditory covert spatial orienting. *J Exp Psychol Human* 25:1234–1252.
- Michalka SW, Rosen ML, Kong L, Shinn-Cunningham BG, Somers DC (2016) Auditory spatial coding flexibly recruits anterior, but not posterior, visuotopic parietal cortex. *Cereb Cortex* 26:1302–1308.
- Mizrahi A, Shalev A, Nelken I (2014) Single neuron and population coding of natural sounds in auditory cortex. *Curr Opin Neurobiol* 24:103–110.
- Munoz-Lopez MM, Mohedano-Moriano A, Insausti R (2010) Anatomical pathways for auditory memory in primates. *Front Neuroanat* 4:129.
- Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I (2001) "Primitive intelligence" in the auditory cortex. *Trends Neurosci* 24:283–288.
- Nelken I (2004) Processing of complex stimuli and natural scenes in the auditory cortex. *Curr Opin Neurobiol* 14:474–480.
- Niwa M, Johnson JS, O'Connor KN, Sutter ML (2012) Activity related to perceptual judgment and action in primary auditory cortex. *J Neurosci* 32:3193–3210.
- O'Sullivan J, Herrero J, Smith E, Schevon C, McKhann GM, Sheth SA, Mehta AD, Mesgarani N (2019) Hierarchical encoding of attended auditory objects in multi-talker speech perception. *Neuron* 104:1195–1209.e3.

- Peck CJ, Salzman CD (2014) Amygdala neural activity reflects spatial attention towards stimuli promising reward or threatening punishment. *eLife* 3:e04478.
- Peck CJ, Lau B, Salzman CD (2013) The primate amygdala combines information about space and value. *Nat Neurosci* 16:340–348.
- Plakke B, Hwang J, Romanski LM (2015) Inactivation of primate prefrontal cortex impairs auditory and audiovisual working memory. *J Neurosci* 35:9666–9675.
- Rinne T, Muers RS, Salo E, Slater H, Petkov CI (2017) Functional imaging of audio-visual selective attention in monkeys and humans: how do lapses in monkey performance affect cross-species correspondences? *Cereb Cortex* 27:3471–3484.
- Roberts KL, Summerfield AQ, Hall DA (2009) Covert auditory spatial orienting: an evaluation of the spatial relevance hypothesis. *J Exp Psychol Hum Percept Perform* 35:1178–1191.
- Romanski LM, LeDoux JE (1992) Equipotentiality of thalamo-amygdala and thalamo-cortico-amygdala circuits in auditory fear conditioning. *J Neurosci* 12:4501–4509.
- Romanski LM, Averbeck BB (2009) The primate cortical auditory system and neural representation of conspecific vocalizations. *Annu Rev Neurosci* 32:315–346.
- Schmolek MT, Wang Y, Hanes DP, Thompson KG, Leutgeb S, Schall JD, Leventhal AG (1998) Signal timing across the macaque visual system. *J Neurophysiol* 79:3272–3278.
- Schwartz ZP, David SV (2018) Focal suppression of distractor sounds by selective attention in auditory cortex. *Cereb Cortex* 28:323–339.
- Scott BH, Mishkin M (2016) Auditory short-term memory in the primate auditory cortex. *Brain Res* 1640:264–277.
- Spence CJ, Driver J (1994) Covert spatial orienting in audition: exogenous and endogenous mechanisms. *J Exp Psychol Human* 20:555–574.
- Stefanacci L, Amaral DG (2002) Some observations on cortical inputs to the macaque monkey amygdala: an anterograde tracing study. *J Comp Neurol* 451:301–323.
- Teshiba TM, Ling J, Ruhl DA, Bedrick BS, Peña A, Mayer AR (2013) Evoked and intrinsic asymmetries during auditory attention: implications for the contralateral and neglect models of functioning. *Cereb Cortex* 23:560–569.
- Tsunada J, Liu AS, Gold JI, Cohen YE (2016) Causal contribution of primate auditory cortex to auditory perceptual decision-making. *Nat Neurosci* 19:135–142.
- Yin P, Strait DL, Radtke-Schuller S, Fritz JB, Shamma SA (2020) Dynamics and hierarchical encoding of non-compact acoustic categories in auditory and frontal cortex. *Curr Biol* 30:1649–1663.e5.
- Yukie M (2002) Connections between the amygdala and auditory cortical areas in the macaque monkey. *Neurosci Res* 42:219–229.
- Zaksas D, Pasternak T (2006) Directional signals in the prefrontal cortex and in area MT during a working memory for visual motion task. *J Neurosci* 26:11726–11742.