Hierarchical effects of choice related activity and neural encoding during feature selective attention

Jennifer L. Mohn[1,2], Joshua D. Downer[1,3], Kevin N. O'Connor[1,2], Jeffrey S. Johnson[1,2], and Mitchell L. Sutter[1,2]

[1] Center for Neuroscience; University of California, Davis; Davis, CA 95618; USA

[2] Department of Neurobiology, Physiology and Behavior; University of California, Davis; Davis, CA 95618; USA

[3] Current address: Dept. of Otolaryngology – Head and Neck Surgery; University of California, San Francisco; San Francisco, CA, 94158; USA

Corresponding Author:

Mitchell L. Sutter

Center for Neuroscience

University of California, Davis

1544 Newton Court Davis CA 95618

mlsutter@ucdavis.edu

No. of pages: 28

No. of Figures (Main): 6

No. of Tables: 1

No. of words: Abstract (214); Introduction (797); Results (1034) Discussion (1017)

Conflict of Interest: The authors declare no competing financial interest

Acknowledgements: This work was funded by NIH NIDCD grant DC02514 (MLS)

Abstract

Selective attention is necessary to sift through, form a coherent percept and make behavioral decisions on the vast amount of information present in most realistic sensory environments. How and where selective attention is employed in cortex and how this perceptual information then informs the relevant behavioral decisions is still not well understood. Studies probing selective attention and decision making in visual cortex have been enlightening as to how sensory attention might work in that modality, whether or not similar mechanisms are employed in auditory attention is not yet clear. Therefore, we trained rhesus macaques on a feature selective attention task, where they switched between reporting changes in temporal and spectral dimensions to a broadband noise stimulus. We investigated how the encoding of the temporal information by single neurons in primary and secondary (lateral belt, ML) auditory cortex were affected by the different attention conditions. We found a large population of neurons that showed mixed-selectivity to these features. We found no difference in temporal encoding between the attentional conditions. We found an interesting shift in choice-related activity values between attentional conditions, which suggests this - and likely other - activity correlated to the choice is not simply motor preparation or action, and supports the relationship between reported choice-related activity and the decision and perceptual process.

Introduction

The auditory system is often faced with the difficult challenge of attending to a specific sound in a noisy environment, such as following a conversation in a loud room. The neural mechanisms by which the auditory system attends to one sound source and ignores distracting sounds, however, are not yet understood. Studies probing the mechanisms underlying auditory attention in cortex have been largely concerned with task engagement, wherein the effects of active performance on neural activity is compared to those of passive listening. Studies utilizing this paradigm have shown that task engagement can improve neural sound discrimination (Niwa et al. 2012a, 2015a), modulate neuronal tuning in primary auditory cortex (A1) as well as higher-order areas (Fritz et al. 2003; Fritz 2005; Fritz et al. 2007; Lee and Middlebrooks 2011a), and alter the structure of correlated variability within neural populations (Downer et al. 2015). Though informative, this active/passive paradigm makes it difficult to disentangle arousal and motor effects from the mechanisms more specifically employed in selectively attending to a single sound source amidst auditory 'clutter'.

The neural basis of auditory feature-selective attention has been relatively unexplored, however. This form of attention, which segregates particular sound features, such as intensity or fundamental frequency, is essential for tasks such as discriminating between talkers in a noisy environment (Bregman 1994; McDermott 2009; Bizley and Cohen 2013; Shinn-Cunningham 2008; Woods and McDermott 2015). Furthermore, it can prove useful for listeners to switch between attended sound features because the most distinctive feature dimensions may vary across sources (Woods and McDermott 2015; Bregman 1994).

In visual cortex, feature-selective attention has been suggested to follow a gain model similar to spatial attention, where responsivity to the attended feature increases in cells tuned to the attended feature and decreases in cells tuned to orthogonal features (Martinez-Trujillo and Treue 2004; Maunsell and Treue 2006). Work on selective attention in auditory cortex (AC) has largely focused on spatial

attention(Engle and Recanzone 2013; Lee and Middlebrooks 2011b; Scott et al. 2007). These studies have found that AC employs a mechanism similar to that reported in visual cortex where a gain in neural activity occurs when attention is directed into the receptive field of a neuron and, conversely, activity is suppressed when attention is directed outside the receptive field (Maunsell and Treue 2006). We endeavored to see if feature selective attention in AC is also facilitated by a gain in activity in neurons tuned to an attended feature.

How and where task relevant sensory information is transformed into a decision in the brain is still largely unclear. There have been mixed reports of activity correlated to the reported decision in AC (Christison-Lagay et al. 2017; Niwa et al. 2012b; Tsunada et al. 2016; Tsunada and Cohen 2014). This choice-related activity has been reported in some studies as early as primary auditory cortex (A1) (Christison-Lagay et al. 2017; Niwa et al. 2012b), with an increasingly larger proportion of neurons showing activity correlated to the decision as one progresses further along the auditory hierarchy to middle lateral (ML) and anterolateral (AL) belt areas (Niwa et al. 2013; Tsunada et al. 2016). This suggests that sensory evidence becomes increasingly transformed to inform a task-relevant decision as the information is passed along the processing stream (Bizley and Cohen 2013; Hackett 2011; Romanski et al. 1999).

There has also been uncertainty as to whether the reported choice activity in AC could be more reflective of motor activity, rather than more directly decision-related activity, as a typical Go/No-Go task requires movement for report of one choice, but not the other (Brosch 2005; Niwa et al. 2013). Here, we employ a Yes/No task format in which a movement is required for both responses in order to disentangle motor-related and choice-related activity in AC.

Here we investigated whether a mechanism for feature-selective attention similar to that in visual cortex is employed in primary (A1) and secondary (middle lateral belt, ML) auditory cortex using noise that was either amplitude modulated (AM) or bandwidth restricted (ΔBW). Monkeys were

presented sounds that varied either in spectral or temporal dimensions, or both, and performed a detection task in which they reported change along one of these feature dimensions. Amplitude modulation (AM) is a salient communicative sound feature for humans and other animals (Schnupp 2006; Shannon et al. 1995; Van Tasell et al. 1987; Wang et al. 2007) and can be helpful in sound source segregation (Bregman 1994; Grimault et al. 2002). Spectral content changes were used as a difficulty-matched attentional control. We expected to see a gain in AM encoding when animals were cued to attend to that feature, compared to when they were cued to attend ΔBW changes. We also examined choice-related activity in AC, expecting to find a larger proportion of neurons with significant choice-related activity in ML than A1.

Materials and Methods

Subjects.

Subjects were two adult rhesus macaques, (one male, one female 7-13 kg). All procedures were approved by the University of California–Davis Animal Care and Use Committee and met the requirements of the United States Public Health Service policy on experimental animal care.

Stimuli.

Stimuli were constructed from broadband Gaussian (white) noise bursts (400 ms; 5 ms cosine ramped), 9 octaves in width (40 to 20480 Hz). Four different seeds were used to create the carrier noise, which was frozen on a trial. To introduce variance along spectral and temporal dimensions, the spectral bandwidth of the noise was narrowed (ΔBW) or the noise envelope was sinusoidally amplitude modulated (AM). The extent of variation in each dimension was manipulated to measure behavioral and neural responses above and below threshold for detecting each feature.

Sound generation methods have been previously reported (O'Connor et al., 2011). Briefly, sound signals were produced using an in-house MATLAB program and a digital-to-analog converter (Cambridge Electronic Design [CED] model 1401). Signals were attenuated (TDT Systems PA5, Leader LAT-45), amplified (RadioShack MPA-200), and presented from a single speaker (RadioShack PA-110) positioned approximately 1.5 m in front of the subject centered at the interaural midpoint. Sounds were generated at a 100 kHz sampling rate. Intensity was calibrated across all sounds (Bruel & Kjaer model 2231) to 65 dB at the outer ear. It is important to note that some methods of generating ΔBW introduce variation in that sound's envelope, however we implemented a synthesis method that constructs noise using a single-frequency additive technique and thus avoids introducing envelope variations that could serve as cues for ΔBW (Strickland and Viemeister 1997).

Recording procedures.

Each animal was implanted with a head post centrally behind the brow ridge and a recording cylinder over an 18 mm craniotomy over the parietal lobe using aseptic surgical techniques (O'Connor et al. 2005). Placement of the craniotomy was based on stereotactic coordinates of auditory cortex to allow vertical access through parietal cortex to the superior temporal plane (Saleem and Logothetis 2007).

All recordings took place in a sound attenuating, foam-lined booth (IAC: 9.5x10.5x6.5 feet) while subjects sat in an acoustically transparent chair (Crist Instruments). Three quartz-coated tungsten microelectrodes (Thomas Recording, 1–2 MΩ; 0.35 mm horizontal spacing; variable, independently manipulated vertical spacing) were advanced vertically to the superior surface of the temporal lobe. Extracellular signals were amplified (AM Systems model 1800), bandpass filtered between 0.3 Hz and 10 kHz (Krohn-Hite 3382), and then converted to a digital signal at a 50 kHz sampling rate (CED model 1401). During electrode advancement, auditory responsive neurons were isolated by presenting various

sounds while the subject sat passively. When at least one auditory responsive single unit was well isolated, we measured neural responses to the two features while the subjects sat passively awake. At least 10 repetitions of each of the following stimuli were presented: the unmodulated noise, each level of bandwidth restriction, and each of the possible AM modulation test frequencies (described below). We also measured pure tone and band passed noise tuning to aid in distinguishing area boundaries. After completing these tuning measures, experimental behavioral testing and recording began. When possible, tuning responses to the tested stimuli were again measured after task performance, to ensure stability of electrodes throughout the recording. Contributions of single units (SUs) to the signal were determined offline using principal components analysis-based spike sorting tools from Spike2 (CED). Spiking activity was at least 4–5 times the background noise level. Fewer than 0.1% of spike events assigned to single unit clusters fell within a 1 ms refractory period window. Only recordings in which neurons were well isolated for at least 180 trials within each condition (excluding instruction trials) were included in analysis here.

Cortical field assessment.

Recording locations were determined using both stereotactic coordinates (Martin and Bowden 1996) and established physiological measures (Merzenich and Brugge 1973; Rauschecker and Tian 2000; Tian and Rauschecker 2004). In each animal, we mapped characteristic frequency (CF) and sharpness of band-passed noise tuning to establish a topographic distribution of each. Tonotopic gradient reversal, BW distribution, spike latency and response robustness to pure tones was used to determine the boundary between A1 and ML and assign single units to an area (Downer et al. 2017; Niwa et al. 2015a). Recordings were assigned to their putative cortical fields *post hoc* using recording location, tuning preferences, and latencies.

Feature attention task.

This feature attention task has been previously described in detail (Downer et al. 2017). Briefly, subjects moved a joystick laterally to initiate a trial, wherein an initial sound (the S1, always the 9-octave wide broadband noise stimulus) was presented, followed by a second sound (S2) after a 400ms inter-stimulus interval (ISI). The S2 could be the unmodulated broadband noise stimulus, or AM modulated, BW narrowed or both.

Only three values of each feature (AM, ΔBW) were presented, limiting the size of the stimulus set in order to obtain reasonable power for neural data analysis. The stimulus space was further reduced by presenting only a subset of the possible co-varying stimuli. Within each recording session, we presented 13 total stimuli. To equilibrate difficulty between the two features, we presented values of each feature so that one was near threshold, one was slightly above, and one far above threshold. Thresholds were determined for each feature independently for each subject before the co-varying feature attention task began. For Monkey U, the ΔBW values were 0.375, 0.5, and 1 octave and the AM depth values were 28%, 40%, and 100%. For Monkey W, the ΔBW values were 0.5, 0.75, and 1.5 octaves and the AM depth values were 40%, 60%, and 100%.

For all analyses in which data are collapsed across subjects, ΔBW values and AM values are presented as ranks (ΔBW 0-3 and AM 0-3) (e.g., Fig. 1). Within a given session, AM sounds were presented at only a single frequency. Across sessions, a small set of frequencies was used (15, 22, 30, 48, and 60 Hz). The AM frequency was selected randomly each day. Subjects were cued visually via an LED above the speaker as to which feature to attend (green or red light, counterbalanced between subjects). Additionally, each block began with 60 “instruction” trials in which the only targets presented were along the attended feature dimension (i.e., sounds containing the distractor feature were not presented). Subjects were to respond with a “yes” (up or down joystick movement, counterbalanced across subjects) on any trial in which the attended feature was presented, otherwise, the correct

response was “no” (opposite joystick movement). Hits and correct rejections were rewarded with a drop of water or juice and misses and false alarms resulted in a penalty (3–5 s timeout).

During the test conditions, the S2 was unmodulated broadband noise on 25% of the trials, co-varying on 25% of the trials, and contained only ΔBW or AM on 25% of the trials respectively. Sounds in the set were presented pseudo-randomly such that, over sets of 96 trials, the entire stimulus set was presented exhaustively (including all four random noise seeds). Block length was variable, based in part on subjects’ performance, to ensure a sufficient number of correct trials for each stimulus. Not including instruction trials, block length was at least 180 trials and at most 360 trials, to ensure that subjects performed in each attention condition at least once during the experiment. Subjects could perform each attention condition multiple times within a session. Only sessions in which subjects completed at least 180 trials per condition (excluding instruction trials) were considered for analysis in this study.

Analysis of single neuron feature selectivity

Neurons were categorized as ‘increasing’ or ‘decreasing’ their responsiveness to each feature by comparing their response to stimuli with each feature present to their responsiveness over all trials. Spike counts (SC) were calculated over the entire 400ms stimulus window. SCs in response to feature present stimuli were normalized over the entire spike count distribution across both features, including unmodulated noise, for that cell. To characterize this response function, we calculated a feature-selectivity index (FSI) for each feature as follows:

[1] $$FSI_{AM} = \frac{SC_{AM_{>0},\Delta BW_0} - SC_{AM_0,\Delta BW_0}}{SC_{AM_{>0},\Delta BW_0} + SC_{AM_0,\Delta BW_0}}$$

[2] $$FSI_{BW} = \frac{SC_{\Delta BW_{>0},\mathrm{AM}_0} - SC_{\Delta BW_0,AM_0}}{SC_{\Delta BW_{>0},AM_0} + SC_{\Delta BW_0,AM_0}}$$

Where SC_x is the mean SC in response to the given set of stimuli designated by the subscript. A Kruskal – Wallis rank-sum test was used to determine the significance of the FSI for each neuron.

Phase projected vector strength

Vector strength (VS) is a metric that describes the degree to which the neural response is phase-locked to the stimulus (Goldberg and Brown 1969; Mardia and Jupp 2000). VS is defined as:

[3] $$VS = \frac{\sqrt{(\sum_{i=1}^{n} \cos\theta_i)^2 + (\sum_{i=1}^{n} \sin\theta_i)^2}}{n}$$

Where *n* is the number of spikes over all trials and θ_i is the phase of each spike, in radians, calculated by:

[4] $$\theta_i = 2\pi \frac{t_{i\ mod\ p}}{p}$$

Where t_i is the time of the spike (in ms) relative to the onset of the stimulus and *p* is the modulation period of the stimulus (in ms). When spike count is low, VS has a tendency to report as spuriously high. Phase projected Vector Strength (VSpp), is a variation on VS developed to help mitigate issues with low SC trials (Yin et al. 2011). VSpp is calculated by first calculating VS for each trial, then the mean phase angle of each trial is compared to the mean phase angle of all trials, and the trial VS value is penalized if out of phase with the global mean response. VS_{pp} is defined as:

[5] $$VS_{pp} = VS_t \cos(\phi_t - \phi_c)$$

Where VS_{pp} is the phase-projected vector strength per trial, VS_t is the vector strength per trial, as calculated in [1], an ϕ_t and ϕ_c are the trial-by-trial and mean phase angle in radians, respectively, calculated for each stimulus by:

[6] $$\phi = arctan2\left(\frac{\sum_{i=1}^{n} \sin\theta_i}{\sum_{i=1}^{n} \cos\theta_i}\right)$$

Where *n* is the number of spikes per trial (for ϕ_t) or across all trials (for ϕ_c). In this report, we report VSpp exclusively to measure phase-locking, as SC tended to be relatively low and VS and VSpp tend to be in good agreement with the exception of low SCs where VSpp tends to be more accurate than VS (Yin et al. 2011).

Analysis of neural discriminability

We applied the signal detection theory-based metric area under the receiver operating characteristic (ROCa) (Green and Swets 1974) to measure how well neurons could discriminate between feature present and feature absent stimuli. ROCa represents the probability an ideal observer can detect the presence of the target feature given only the neural responses (either firing rate or VSpp). To calculate ROCa, we partition the trial-by-trial neural responses into two distributions: those when the target feature was present in the stimulus and trials where it was absent. Then we determine the proportion of trials in each group where the neural response exceeds a criterion value. We repeat the measure using 100 criterion values, covering the whole range of responses. The plot of each pair of probabilities is the ROC, and the area under this plot is the ROCa. ROCa is bounded by 0 and 1, where both extremes indicate perfect discrimination between target feature present and absent stimuli, and 0.5 indicates a chance level of discrimination between the two distributions.

Analysis of choice-related activity

Choice probability (CP) is an application of ROC analysis used to measure the difference between neural responses contingent on what the animal reports, for example, reporting whether a stimulus feature is present or absent (Britten et al. 1992, 1996). Similar to ROCa described above, CP values are bounded by 0 and 1, and a CP value of 0.5 indicates no difference (or perfect overlap) in the neural responses between 'target present' and 'target absent' responses. A CP value of 1 means for every trial that the animal reports a feature, the neuron fired more than on trials where the animal did not report the feature. A CP value of 0.5 indicates no difference in the neural responses between 'report feature present' and 'report feature absent' trials. Stimuli that did not have at least 5 yes and 5 no responses were excluded from analyses. CP was calculated based on both firing rate and on VSpp. For rate-based CP, we calculated CP both for each stimulus separately, and pooled across stimuli. We

calculated this stimulus-pooled CP by first separating the 'yes' and 'no' response trials within stimulus, then converting these rates into z-scores within a stimulus, then combined these z-scored responses across stimuli. CP was calculated during both the 400ms stimulus presentation (S2) and during the response window (RW), the time after stimulus offset and prior to the response (0.2 – 3s).

Results

We recorded activity from 92 single units in A1 (57 from Monkey W, 35 from Monkey U) from 33 recording sessions and 122 single units in ML (49 from Monkey W, 73 from Monkey U) over 39 recording sessions as animals performed a feature-selective attention task.

Feature tuning

There was no significant difference in the proportions of neurons in A1 and ML that encoded AM (47.8% A1, 38.5% ML; $p = 0.08$, $\chi2$ test). We found a large proportion of neurons in both A1 and ML that were sensitive to the relatively small changes in ΔBW (maximum ΔBW = 2 octaves, see Methods) from the 9-octave wide unmodulated noise, though there was no difference in the proportion of ΔBW encoding neurons between areas (32.6% A1, 29.5% ML; $p = 0.18$, $\chi2$ test) (Table 1).

A large population of neurons decreased firing rate for increasing AM depth ('decreasing cells') in both A1 and ML (Table 1). We also found that nearly half of the neurons in both A1 and ML decreased firing rate for increasing ΔBW (Table 1). However, the population of neurons that *significantly* encoded AM was largely dominated by cells that increased firing rate for increasing AM depth in both A1 and ML, with only 13.6% of AM encoders classified as 'decreasing' units in A1, and 10.6% of AM encoders 'decreasing' in ML. For significant ΔBW encoders, the population was more evenly split between 'increasing' and 'decreasing' units in both A1 and ML, with 43.3% of BW encoders 'decreasing' units in A1, and 30.6% in ML. In both A1 and ML, there was a significant positive correlation between

AM and BW selectivity, so cells that had a positive FSI_{AM} also tended to have a positive FSI_{BW} (A1 Pearson's rho = 0.3143, p = 0.002; ML Pearson's rho = 0.3109, p = 5.3 e-4) (Figure 1). We found no significant difference between A1 and ML in the proportions of 'increasing' and 'decreasing' encoding cells for either AM ('Increasing' p = 0.21 χ^2 test; 'Decreasing' p = 0.11 χ^2 test) or BW ('Increasing' p = 0.22 χ^2 test; 'Decreasing' p = 0.52 χ^2 test).

Vector strength encoding

We found a similar proportion of cells in A1 and ML with significant phase-locked encoding of AM (p = 0.77, χ^2 test), as measured by phase-projected vector strength (VSpp) (Table 1). As in previous reports (Niwa et al. 2013), we found VSpp to be weaker in ML than A1 (Figure 2). In both A1 and ML, there was no significant difference in phase-locking (VSpp) between the attend AM and attend ΔBW conditions (Figure 2).

Feature discriminability and context effects

We used the signal detection theory-based area under the Receiver Operating Characteristic (ROCa) to measure how well an ideal observer could discriminate 'feature present' stimuli from 'feature absent' stimuli based on the neural responses (either firing rate or VSpp). For both AM and ΔBW features, as the modulation level increased neural discriminability also increased. However, there was no effect of attentional context on either feature in A1 at any level of feature modulation for either A1 (Figure 3a) or ML (Figure 3b).

VSpp-based discrimination (ROCa) of AM from unmodulated sounds was similar in A1 and ML at high modulation depths, with the A1 population getting increasingly better than ML as the modulation depth (and hence detectability of AM) decreased. However, there was no difference in VSpp discriminability between the attend AM and attend ΔBW conditions for any modulation depth (Figure 4).

Choice related activity

We found a similar proportion of neurons in A1 (19.5%) and ML (26.2%) with significant choice-related activity during the stimulus window ($p = 0.31$, $\chi2$ test) In both areas, the population of neurons with significant choice related activity during the response window (from stimulus end to joystick movement) was larger than during the stimulus window, and the proportions of neurons were again similar between the two areas (41.3% A1, 34.4% ML, $p = 0.41$, $\chi2$ test).

In A1, during the attend AM condition, CP values were evenly distributed about 0.5 during both the stimulus presentation (S2 mean CP = 0.5134, $p = 0.87$ Wilcoxon signed rank) and the response window (RW mean CP = 0.4943, $p = 0.43$ signed rank) (Figure 5). In contrast, during the attend ΔBW context, the CP values tended to be lower than 0.5 during both the stimulus (S2 mean CP = 0.4860, $p = 0.02$ Wilcoxon signed rank test) and the response window (RW mean CP =0.4525, $p = 4.2\ e\text{-}8$ signed-rank) (Figure 5). That is, during the attend ΔBW condition, the population of neurons tended to increase firing rate for 'target feature absent' reports, whereas during the attend AM condition, it was equally likely that a neuron increased firing rate for 'target feature absent' as decrease. There was a significant difference in the population CP distributions between attention conditions during the RW ($p = 0.004$, Wilcoxon signed rank), though not during the S2 ($p = 0.06$, Wilcoxon signed rank).

The choice-related activity in ML was similar to that in A1 during the stimulus presentation. During the attend AM condition, activity was evenly distributed about 0.5 (S2 mean CP = 0.4997, $p = 0.94$, Wilcoxon signed-rank) (Figure 6). During the attend ΔBW condition, CP tended to be less than 0.5 (S2 mean CP = 0.4860, $p = 0.043$ signed-rank) (Figure 6). However, during the response window, CP values tended to be less than 0.5 in both the attend AM condition (mean CP = 0.4850, $p = 0.004$ Wilcoxon signed-rank) and the attend ΔBW condition (mean CP = 0.4641, $p = 2.7\ e\text{-}5$ Wilcoxon signed-rank) (Figure 5). This is in contrast to A1 where CP values tended to be lower than 0.5 only in the attend

BW condition. There was no significant difference in the distribution of CP values in ML neurons between the attend AM and attend ΔBW conditions during the stimulus presentation ($p = 0.15$, Wilcoxon signed rank). However, there was a significant difference in the CP distribution between the attend AM and attend ΔBW conditions in ML during the response window ($p = 0.033$ Wilcoxon signed-rank), reflecting the population shift to CP values less than 0.5 in the attend ΔBW condition compared to the attend AM condition.

Discussion

We found a large proportion of cells in ML that decreased firing rate with increasing AM detectability, similar to previous findings in ML (Niwa et al. 2013). However, unlike this previous study where ML had a significantly larger population of cells with decreasing AM depth functions than A1, we found a similarly large proportion of A1 neurons with decreasing AM depth functions. Further, the majority of neurons that significantly encoded AM depth had increasing AM depth functions in both A1 and ML. This suggests that the encoding of amplitude modulation can be flexible dependent upon the behavioral and sensory demands of the task. In essence, with increased perceptual difficulty it may be necessary to engage the encoding capabilities of hierarchically 'higher' areas in order to solve the task more than is required in simpler tasks, and thus the encoding schemes look more similar between these two areas. We also found a large population of cells in both A1 and ML that were sensitive to changes in bandwidth. This was particularly surprising as the changes in bandwidth were relatively small compared to the 9-octave wide unmodulated noise.

Attending to the target-feature did not significantly improve single neuron amplitude modulation or bandwidth restriction detection in A1 or ML. This seems surprising considering the wide array of effects that have been previously reported in auditory cortex during different tasks, such as in an active engagement versus passive listening task (Niwa et al. 2013). In that study, an improvement in

both rate-based and temporal AM encoding was observed in A1 and ML neurons when animals performed a single-feature AM detection task compared to when animals passively listened to the same stimuli (Niwa et al. 2013, 2015b). We did not see a similar level of encoding improvement, likely due to the more fine-tuned form of attention needed to perform this task. One could expect to observe smaller effects from this more selective form of attention than in a passive versus active listening task, as the difference between attending to one feature of a sound compared to another is much smaller than switching between paying attention to a sound and passively having sounds presented. Furthermore, arousal, as measured with pupillometry, has recently been shown to correlate with increases in activity, gain and trial-to-trial reliability of A1 neurons (Schwartz et al. 2019), which could account for some of the effects seen in task engagement paradigms.

Feature-selective attention has been shown to have gain effects on neurons tuned to the attended feature in visual cortex (Ni and Maunsell 2019; Treue and Trujillo 1999). It is possible that we did not see a similar gain effect of feature-selective attention in AC due to the mixed-selectivity we found in the encoding of these features (i.e. most neurons were sensitive to both AM and ΔBW). Perhaps segregation of features to which cells have mixed selectivity requires a different mechanism relative to separate, independent feature encoding (as is typically the case in studies similarly aiming to examine feature selective attention in visual cortex). One such possible mechanism might be the structure of correlated variability within the population, which has been shown to be modulated by feature attention (Downer et al. 2017).

However, it is likely that mixed-selectivity is not the only reason we did not see a visual cortex-like gain effect. In a recent study where rats performed a frequency categorization task with shifting boundaries, Jaramillo and colleagues similarly found that neurons in AC did not improve their discriminability with attentional context (Jaramillo et al. 2014). This similar lack of enhancement seen in a task where only a single feature is modulated, suggests that the mechanism for feature-selective

attention in auditory cortex could be enacted via a mechanism different than that reported in visual cortex.

We saw an interesting difference in the distribution of choice related activity between the attentional conditions, where the correlation between firing rate and choice report shifted direction between conditions. During the attend AM context, CP was evenly distributed about 0.5 with some neurons with significant choice activity at either extreme. In contrast, during the attend BW context, CP values were shifted towards 0, with very few neurons having significant choice-related activity greater than 0.5 (increasing FR for 'feature present' response). Neurons in auditory cortical areas may also respond to motor activity (Brosch 2005). Some previous reports on choice related activity have been difficult to interpret, as they employed a Go/No-Go task format in which one perceptual choice required a movement and the other choice did not (Brosch 2005; Niwa et al. 2013). Therefore, the choice-related activity observed was difficult to disentangle from a general preparation to move.

The task reported here was a Yes/No task, requiring a motor response to each decision (target present versus target absent). The shift in choice-related activity between attention conditions observed here, where neurons that increased their firing rate for 'target feature present' reports shifted to decreasing firing rate for the same response between attention conditions shows that this choice-related activity cannot simply reflect motor preparation or action. This then strengthens the relationship between this activity and the decision or attention process.

The lack of clear attentional improvement of single neuron encoding found in this study suggests one or more of the following: (1) feature attention in the auditory cortex is not implemented by single neurons; (2) the feature-selective attention necessary for this particular task occurs at a later stage in auditory processing; (3) the mixed selectivity of single neurons in A1 and ML for these features make them insensitive to selective attention effects, in contrast to feature-selective attention neurons studied found in visual cortex (Martinez-Trujillo and Treue 2004; Maunsell 2015; Maunsell and Treue

2006), Probing A1 and ML with a similar task using features for which neurons show clear selectivity, might further reveal effects of feature attention in individual neurons. While we did not see robust differences in encoding between attentional conditions, the difference in attentional choice-related activity reveals that it is not simply reflective of motor preparation, and suggests that activity correlated to reported choice as early as A1 could be informing perceptual and decision processes.

	AM Rate Coder	Decreasing AM	Increasing AM +AM coder	Decreasing AM + AM coder	BW Coder	Decreasing BW	Increasing BW +BW Coder	Decreasing BW + BW Coder	VS Coder
A1 (n=92)	47.8%	32.6%	41.3%	6.5%	32.6%	42.5%	18.5%	14.1%	32.6%
ML (n=122)	38.5%	27.1%	34.4%	4.1%	29.5%	41.8%	20.5%	9.0%	30.33%

Table 1: Distribution of cells in A1 and ML with increasing response functions vs. decreasing response functions to both features. The majority of cells that significantly encoded AM (AM coder) had increasing rate depth functions. Cells that significantly encoded ΔBW (BW Coder) in A1 were about equally likely to be increasing as decreasing. BW coders in ML were more likely to have increasing functions than decreasing.

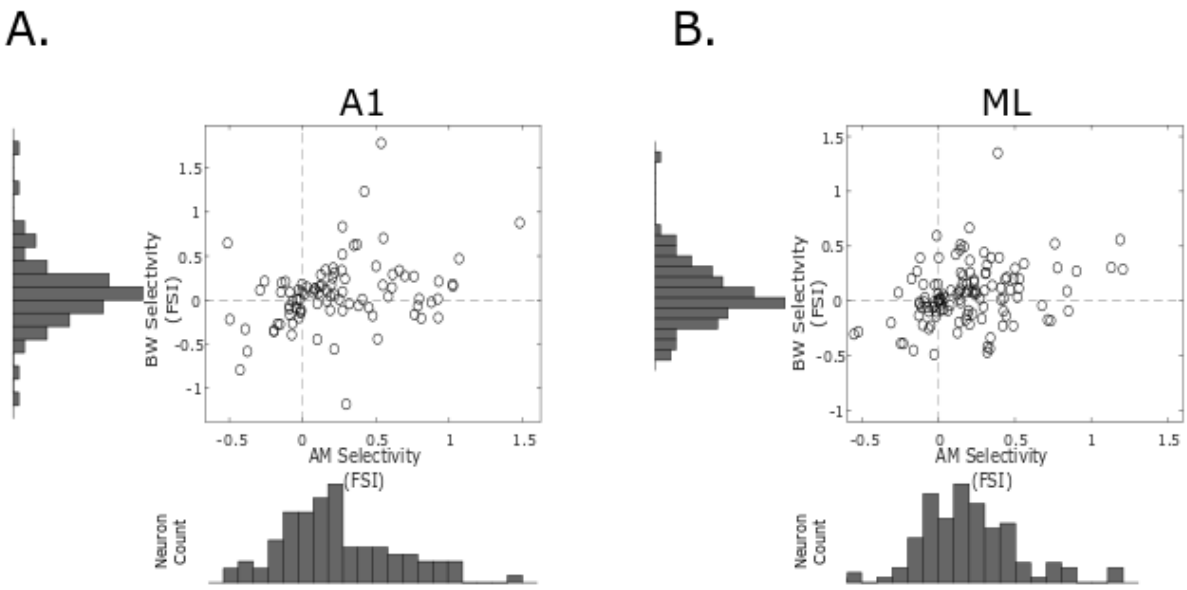

Figure 1: Single neuron feature selectivity index (FSI), a measure of how sensitive a neuron is to changes in each feature value separately. **A.** A1: a positive correlation between AM and ΔBW selectivity (Pierson rho = 0.3143, p = 0.002) **B.** ML: positive correlation between AM and BW selectivity (Pierson's rho = 0.3109, p = 5.32 e-4)

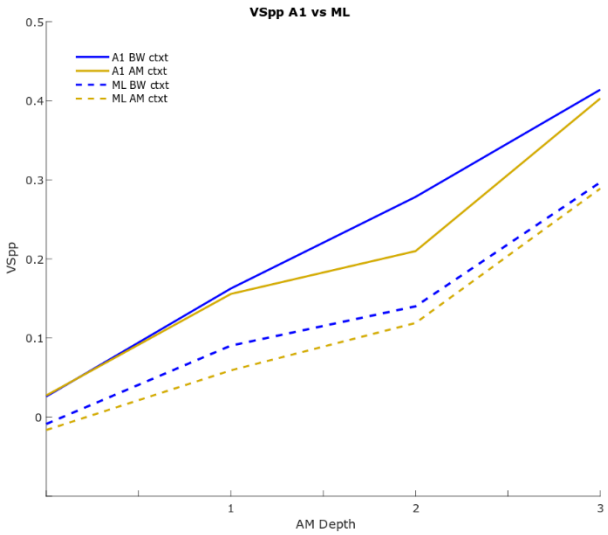

Figure 2. Average phase locking ability of single units in A1 (solid lines) and ML (dashed lines), as measured by phase projected vector strength (VSpp). There was no significant difference in either area between attend AM (gold) attend ΔBW (blue) conditions.

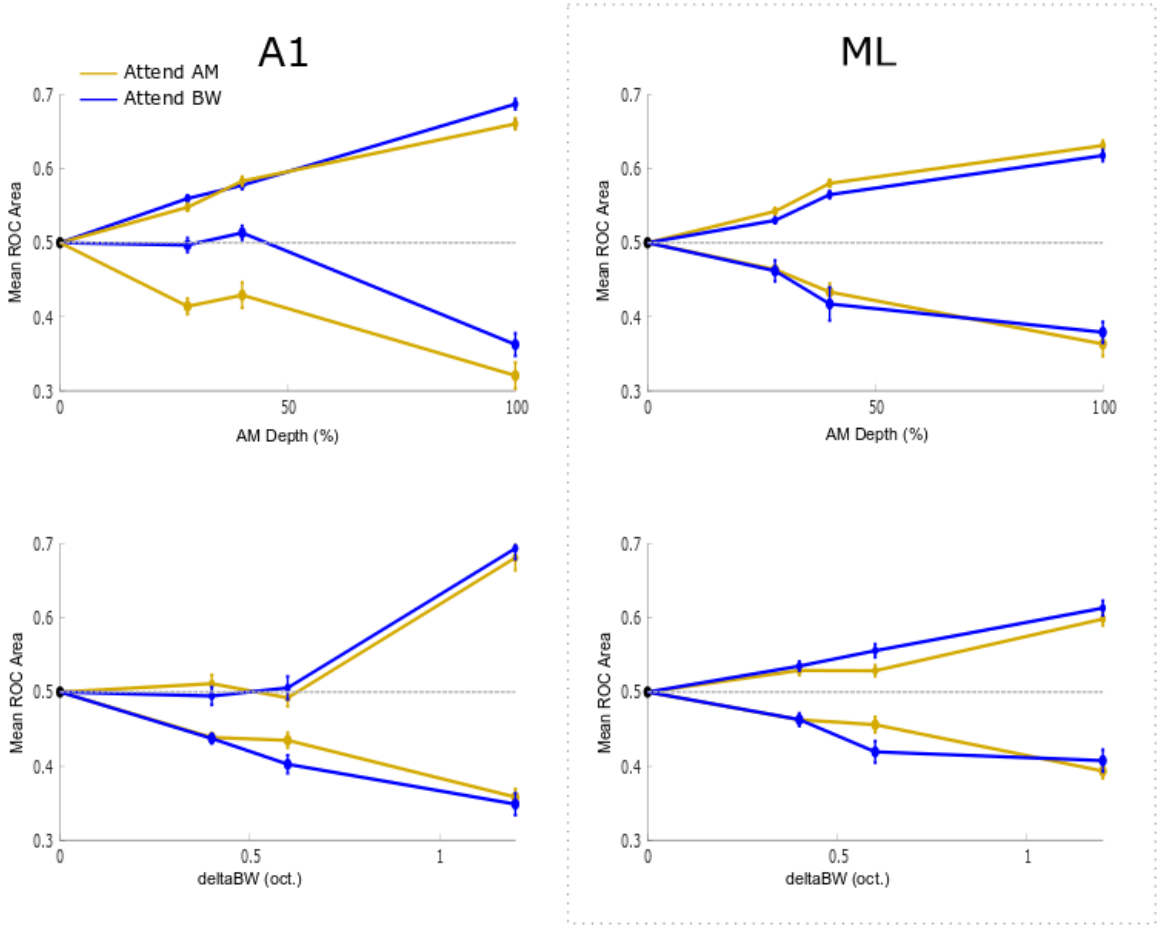

Figure 3: Firing rate based ROCa for each feature by attention condition. Blue lines indicate attend BW condition, yellow lines indicate attend AM condition. A. AM encoding in A1 (38 increasing, 6 decreasing cells) B. BW encoding in A1 (17 increasing, 13 decreasing cells) C. AM encoding in ML (42 increasing, 5 decreasing cells) D. BW encoding in ML (25 increasing, 11 decreasing cells)

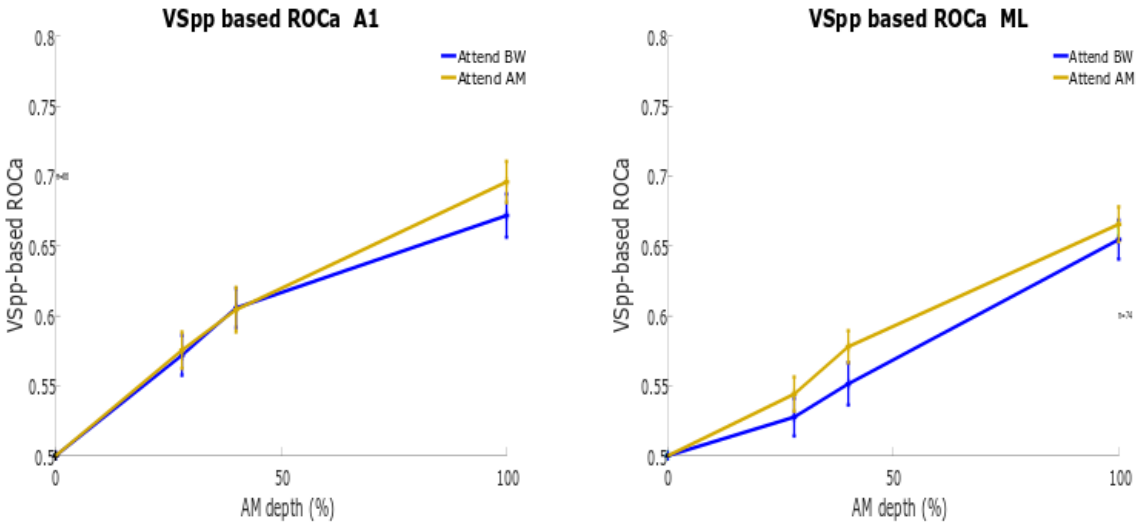

Figure 4: AM synchronizing responses in A1 and ML for attend AM (yellow) and attend BW (blue) conditions. **A.** Phase-projected vector strength is greater in A1 (solid) than ML (dashed) at all AM depths **B.** VS-based discriminability (ROCa) in A1 is not significantly different between attentional contexts. **C.** VS-based ROCa in ML is similar to A1. Discriminability is not different between attentional contexts.

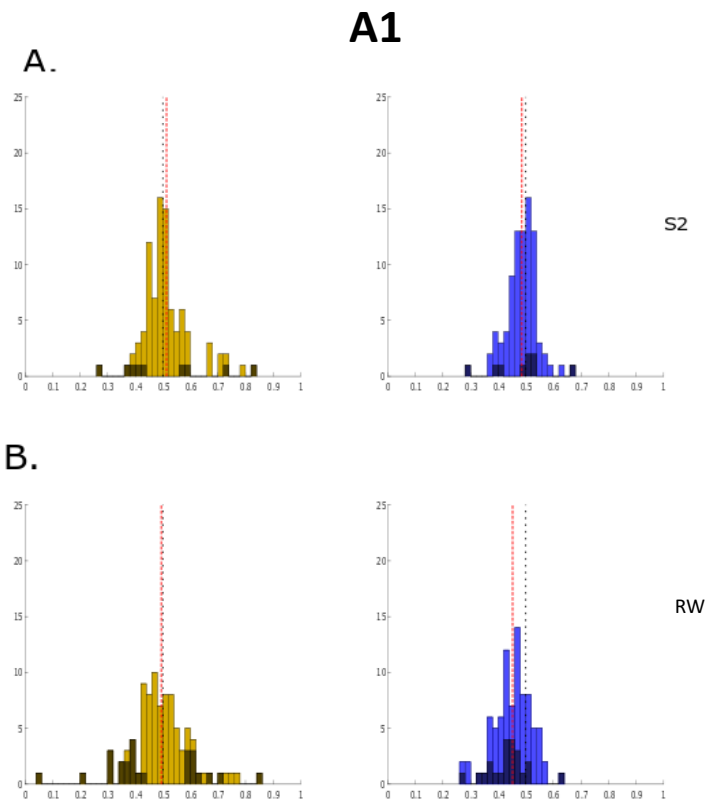

Figure 5: Choice probability in A1. Values closer to 0 indicate increased activity for 'target feature not present' response, whereas 1 indicates increased activity for 'target feature present' response. Darker colored bars indicate cells with significant choice activity. Black dotted line indicates 0.5, red dashed line denotes the population mean. **A.** CP during the test stimulus (S2) presentation is evenly distributed about 0.5 in both attention contexts. **B.** There are more neurons with significant CP during the response window than the stimulus in both attentional contexts. In the attend AM context (yellow), cells with significant CP are equally distributed above and below 0.5. In the attend ΔBW context (blue), cells with significant CP are almost only seen below 0.5 and population mean CP is significantly less than 0.5.

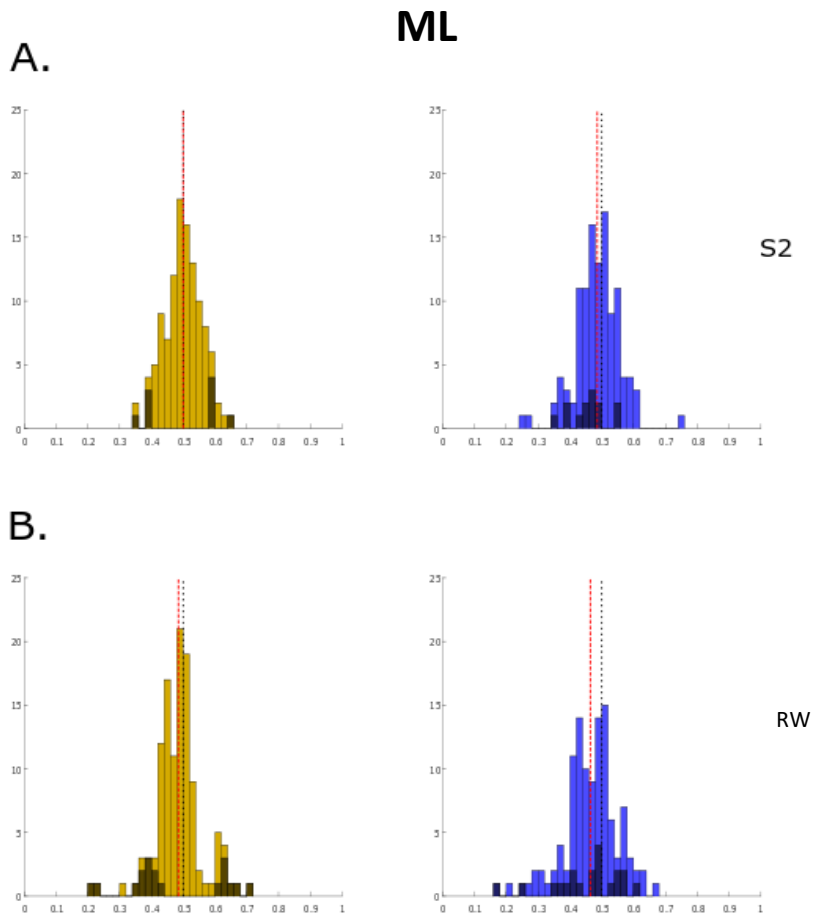

Figure 6: Choice probability in ML, as in Figure 5. **A.** CP during stimulus presentation **B.** CP during the response window. More cells with significant CP than during the S2 in both attentional contexts. In the attend AM context (yellow), cells with significant CP are found both above and below 0.5. In the attend ΔBW context (blue), the population mean is significantly less than 0.5 and most cells with significant CP have values less than 0.5.

References

Bizley JK, **Cohen YE**. The what, where and how of auditory-object perception. *Nature Reviews Neuroscience* 14: 693–707, 2013.

Bregman AS. *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press, 1994.

Britten K, **Shadlen M**, **Newsome W**, **Movshon J**. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *The Journal of Neuroscience* 12: 4745–4765, 1992.

Britten KH, **Newsome WT**, **Shadlen MN**, **Celebrini S**, **Movshon JA**. A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience* 13: 87–100, 1996.

Brosch M. Nonauditory Events of a Behavioral Procedure Activate Auditory Cortex of Highly Trained Monkeys. *Journal of Neuroscience* 25: 6797–6806, 2005.

Christison-Lagay KL, **Bennur S**, **Cohen YE**. Contribution of spiking activity in the primary auditory cortex to detection in noise. *Journal of Neurophysiology* 118: 3118–3131, 2017.

Downer JD, **Niwa M**, **Sutter ML**. Task Engagement Selectively Modulates Neural Correlations in Primary Auditory Cortex. *Journal of Neuroscience* 35: 7565–7574, 2015.

Downer JD, **Rapone B**, **Verhein J**, **O'Connor KN**, **Sutter ML**. Feature-Selective Attention Adaptively Shifts Noise Correlations in Primary Auditory Cortex. *The Journal of Neuroscience* 37: 5378–5392, 2017.

Engle JR, **Recanzone GH**. Characterizing spatial tuning functions of neurons in the auditory cortex of young and aged monkeys: a new perspective on old data. *Frontiers in Aging Neuroscience* 4, 2013.

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

Fritz JB. Differential Dynamic Plasticity of A1 Receptive Fields during Multiple Spectral Tasks. *Journal of Neuroscience* 25: 7623–7635, 2005.

Fritz JB, **Elhilali M**, **Shamma SA**. Adaptive Changes in Cortical Receptive Fields Induced by Attention to Complex Sounds. *Journal of Neurophysiology* 98: 2337–2346, 2007.

Goldberg JM, **Brown PB**. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *Journal of Neurophysiology* 32: 613–636, 1969.

Green DM, **Swets JA**. *Signal detection theory and psychophysics*. Huntington, NY: Krieger Publishing Company, 1974.

Grimault N, **Bacon SP**, **Micheyl C**. Auditory stream segregation on the basis of amplitude-modulation rate. *The Journal of the Acoustical Society of America* 111: 1340–1348, 2002.

Hackett TA. Information flow in the auditory cortical network. *Hearing Research* 271: 133–146, 2011.

Jaramillo S, **Borges K**, **Zador AM**. Auditory Thalamus and Auditory Cortex Are Equally Modulated by Context during Flexible Categorization of Sounds. *Journal of Neuroscience* 34: 5291–5301, 2014.

Lee C-C, **Middlebrooks JC**. Auditory cortex spatial sensitivity sharpens during task performance. *Nature Neuroscience* 14: 108–114, 2011a.

Lee C-C, **Middlebrooks JC**. Auditory cortex spatial sensitivity sharpens during task performance. *Nature Neuroscience* 14: 108–114, 2011b.

Mardia KV, **Jupp PE**. *Directional statistics*. Chichester ; New York: J. Wiley, 2000.

Martinez-Trujillo JC, **Treue S**. Feature-Based Attention Increases the Selectivity of Population Responses in Primate Visual Cortex. *Current Biology* 14: 744–751, 2004.

Maunsell JHR. Neuronal Mechanisms of Visual Attention. *Annual Review of Vision Science* 1: 373–391, 2015.

Maunsell JHR, **Treue S**. Feature-based attention in visual cortex. *Trends in Neurosciences* 29: 317–322, 2006.

McDermott JH. The cocktail party problem. *Curr Biol* 19: R1024-1027, 2009.

Merzenich MM, **Brugge JF**. Representation of the cochlear partition on the superior temporal plane of the macaque monkey. *Brain Research* 50: 275–296, 1973.

Ni AM, **Maunsell JHR**. Neuronal Effects of Spatial and Feature Attention Differ Due to Normalization. *J Neurosci* 39: 5493–5505, 2019.

Niwa M, **Johnson JS**, **O'Connor KN**, **Sutter ML**. Active Engagement Improves Primary Auditory Cortical Neurons' Ability to Discriminate Temporal Modulation. *Journal of Neuroscience* 32: 9323–9334, 2012a.

Niwa M, **Johnson JS**, **O'Connor KN**, **Sutter ML**. Activity Related to Perceptual Judgment and Action in Primary Auditory Cortex. *Journal of Neuroscience* 32: 3193–3210, 2012b.

Niwa M, **Johnson JS**, **O'Connor KN**, **Sutter ML**. Differences between Primary Auditory Cortex and Auditory Belt Related to Encoding and Choice for AM Sounds. *Journal of Neuroscience* 33: 8378–8395, 2013.

Niwa M, **O'Connor KN**, **Engall E**, **Johnson JS**, **Sutter ML**. Hierarchical effects of task engagement on amplitude modulation encoding in auditory cortex. *Journal of Neurophysiology* 113: 307–327, 2015a.

O'Connor KN, **Petkov CI**, **Sutter ML**. Adaptive Stimulus Optimization for Auditory Cortical Neurons. *Journal of Neurophysiology* 94: 4051–4067, 2005.

Rauschecker JP, **Tian B**. Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceedings of the National Academy of Sciences* 97: 11800–11806, 2000.

Romanski LM, **Tian B**, **Fritz J**, **Mishkin M**, **Goldman-Rakic PS**, **Rauschecker JP**. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature Neuroscience* 2: 1131–1136, 1999.

Saleem K, **Logothetis Nikos K**. *A combined MRI and histology atlas of the rhesus monkey brain in sterotaxic coordinates*. Burlington, MA: Academic Press, 2007.

Schnupp JWH. Plasticity of Temporal Pattern Codes for Vocalization Stimuli in Primary Auditory Cortex. *Journal of Neuroscience* 26: 4785–4795, 2006.

Schwartz ZP, **Buran BN**, **David SV**. Pupil-associated states modulate excitability but not stimulus selectivity in primary auditory cortex. Molecular Biology.

Scott BH, **Malone BJ**, **Semple MN**. Effect of Behavioral Context on Representation of a Spatial Cue in Core Auditory Cortex of Awake Macaques. *Journal of Neuroscience* 27: 6489–6499, 2007.

Shannon RV, **Zeng F-G**, **Kamath V**, **Wygonski J**, **Ekelid M**. Speech Recognition with Primarily Temporal Cues. *Science* 270: 303–304, 1995.

Shinn-Cunningham BG. Object-based auditory and visual attention. *Trends in Cognitive Sciences* 12: 182–186, 2008.

Strickland EA, **Viemeister NF**. The effects of frequency region and bandwidth on the temporal modulation transfer function. *The Journal of the Acoustical Society of America* 102: 1799–1810, 1997.

Tian B, **Rauschecker JP**. Processing of Frequency-Modulated Sounds in the Lateral Auditory Belt Cortex of the Rhesus Monkey. *Journal of Neurophysiology* 92: 2993–3013, 2004.

Treue S, **Trujillo JCM**. Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* 399: 575–579, 1999.

Tsunada J, **Cohen YE**. Neural mechanisms of auditory categorization: from across brain areas to within local microcircuits. *Frontiers in Neuroscience* 8, 2014.

Tsunada J, **Liu ASK**, **Gold JI**, **Cohen YE**. Causal contribution of primate auditory cortex to auditory perceptual decision-making. *Nature Neuroscience* 19: 135–142, 2016.

Van Tasell DJ, **Soli SD**, **Kirby VM**, **Widin GP**. Speech waveform envelope cues for consonant recognition. *The Journal of the Acoustical Society of America* 82: 1152–1161, 1987.

Wang L, **Narayan R**, **Grana G**, **Shamir M**, **Sen K**. Cortical Discrimination of Complex Natural Stimuli: Can Single Neurons Match Behavior? *Journal of Neuroscience* 27: 582–589, 2007.

Woods KJP, **McDermott JH**. Attentive Tracking of Sound Sources. *Current Biology* 25: 2238–2246, 2015.

Yin P, **Johnson JS**, **O'Connor KN**, **Sutter ML**. Coding of Amplitude Modulation in Primary Auditory Cortex. *Journal of Neurophysiology* 105: 582–600, 2011.