Multiple sounds induce frequency tuning modulations that degrade stimulus information in monkey inferior colliculus

Authors: Shawn M. Willett and Jennifer M. Groh

Affiliations: Duke University

Abstract: How the brain encodes information about simultaneously occurring stimuli is largely unknown, especially if those stimuli recruit largely overlapping populations of neurons. One hypothesis is that tuning curves might sharpen or shift to limit the number of stimuli driving any given neuron. To investigate this possibility, we recorded the activity of neurons in the inferior colliculus while monkeys localized either one or two simultaneous sounds that differed in frequency. Although monkeys were quite capable of performing this task (~90% correct), comparison of frequency tuning in single versus dual sound trials revealed frequency selectivity was degraded in the dual-sound case. Frequency tuning curves broadened, and frequency modulation of firing rate was attenuated in dual sound conditions. These tuning curve changes led to a slightly worse performance on dual sound trials by a maximum-likelihood decoder trained to predict the presented condition only from held-out spike counts. These results do not support that changes in frequency response functions could be used to overcome the overlap problem in the representation of simultaneous sounds, and suggest recent evidence of alternations in firing rate between rates corresponding to each of the two stimuli may be a more promising line of research.

Introduction:

Natural environments contain numerous stimuli. A fundamental question is how the brain simultaneously encodes multiple items, particularly when more than one item falls within a given neuron's receptive field. Two possible solutions to this multiplicity conundrum have been proposed: that receptive fields effectively shift or shrink to limit the number of stimuli within them, and/or that fluctuating activity patterns allow individual neurons to contribute to encoding only one particular stimulus at any given time. We previously reported evidence of fluctuating activity (Caruso et al., 2018), but how the underlying tuning curves may change and their potential contribution to encoding multiple stimuli remains unknown. In particular, we evaluated whether and how the presence of more than one sound alters the frequency tuning properties of neurons in the inferior colliculus (IC).

The monkey inferior colliculus (IC) is an ideal structure to probe how neurons code concurrent stimuli for two reasons. First, nearly every auditory signal converges to the IC prior to ascending to the thalamus(Aitkin & Phillips, 1984), so information must be preserved at this stage in order to be available to higher areas. How exactly the IC preserves information about sounds remains ambiguous (for review Winer & Schreiner, 2005), but the IC clearly contributes to the perception of sound frequency (Pages et al., 2016) and to sound localization (Jenkins & Masterton, 1982; Kelly & Kavanagh, 1994). Second, the representations of sound frequency and location in the monkey IC are coarse. Neurons in the monkey IC can have quite broad and complex frequency receptive fields (Ryan & Miller, 1978; Versnel et al., 2009; Zwiers et al., 2004). Indeed, a systematic mapping of multiunit responses throughout the monkey IC revealed that 40-80% of frequency modulated sites responded to a pure tone stimulus of approximately 800 Hz (Bulkin & Groh, 2011), falling off to a low of about 20% of sites responding to higher frequency sounds. Given that these findings involved pure tones, it can be inferred that many different natural broadband stimuli will be capable of activating each individual IC neuron. Spatial tuning is also broad, involving a monotonic rate code in which IC neurons typically increase their firing rates as a sound is positioned more contralaterally (Groh et al., 2003; Zwiers et al., 2004).

In short, the almost complete convergence of ascending auditory signals in the IC, together with broad frequency and spatial tuning in this structure, suggests that any two sounds potentially recruit similar populations of IC neurons. Nevertheless, humans (Best et al., 2003; Perrott, 1984a, 1984b; Zhong & Yost, 2017) and monkeys (Caruso et al., 2018) are capable of localizing 2 concurrent sounds, implying the IC population somehow overcomes these coding constraints.

One possibility is that IC neurons overcome this coding constraint by sharpening or shifting their tuning curves limiting the degree to which any two stimuli fall within their receptive field. There is strong precedence that tuning curves are not an immutable property of neurons (Weinberger, 1995). For example, spatial tuning in cat auditory cortex is modulated by behavioral state, units show broad selectivity under anesthesia (Imig et al., 1990; Middlebrooks et al., 1994; Middlebrooks & Pettigrew, 1981) but become more selective in the awake cat (Mickey & Middlebrooks, 2003), and even more selective in cats performing a spatial auditory task (Lee & Middlebrooks, 2011). Moreover, there is ample evidence of surround suppression in the visual system (Chettih & Harvey, 2019; Hartline & Ratliff, 1957) as well as response profile

modulations associated with experience (Freedman et al., 2005) or auditory attention (Connell et al., 2014).

Could the monkey IC also sharpen tuning curves to facilitate encoding of multiple stimuli? Previous work in the rabbit IC showed concurrent sound presentation broadens spatial tuning (Day et al., 2012). However, it remains to be seen if frequency tuning is affected by the presence of a concurrent sound. The present study investigates if and how frequency tuning curves are modulated by the presence of an additional sound and if these changes are beneficial to decoding stimulus information from the population.

Single neuron activity was recorded from the IC while monkeys made saccades or sequences of saccades to single and dual sounds (bandpass sounds of different frequencies). Monkeys were able to perform this task greater than 80% correct and their performance slightly improved with a larger frequency separation between the two simultaneously presented sounds. Comparison of the frequency tuning curves in the absence vs. presence of a second sound of fixed frequency revealed a heterogenous pattern of changes, with most neurons actually broadening their tuning curves in the presence of an additional sound. Although this effect was present with either a high or low frequency paired sound, the effect was larger when the concurrent sound was low frequency. A maximum-likelihood decoder used to decode the stimulus on a held-out trial performed slightly worse when the held-out trial was a dual-sound trial. Thus, the broadening of frequency tuning curves due to presence of an additional sound seems to limit the available stimulus information encoded by the population of IC neurons. We conclude that changes to the frequency tuning properties of IC neurons in the presence of multiple sounds are unlikely to solve the multiplicity problem.

Methods:

General: All procedures were approved by the Duke Institutional Animal Care and Use Committee. Two adult, female, rhesus monkeys participated in this study. Monkey Y and monkey N participated in previous studies (Monkey Y: Caruso et al., 2018; Monkey N: Caruso et al., 2017). General procedures have been previously described in greater detail (Caruso et al., 2018; Groh et al., 2001). Briefly, under aseptic surgery conditions a head post, to restrain the head during experiments, and a scleral search coil, to allow for eye tracking (Judge et al., 1980; Robinson, 1963), were implanted. After recovery monkeys were trained to perform the localization task. Once the task was learned, a recording chamber (Crist Instruments) was implanted over the left IC in both monkeys. The location and angle of the chamber were aimed at the left IC based on stereotaxic coordinates and was confirmed after implantation with structural MRI (Groh et al., 2001).

Training and Behavior: The sound localization task used in this study was designed and implemented with proprietary software (Beethoven, RYKLIN Software Inc) and was adapted from a previous study (Caruso et al., 2018). There were two major adaptations. First, additional sound frequencies were used. Second, sounds in this study are only located at ±12° meaning on dual sound trials the sounds are 24° apart compared to 30° in Caruso et al., (2018).

Initially, monkeys were trained to saccade to the origin of a sound (e.g. to the location of the speaker). Monkey Y was then trained to saccade to the source of sounds presented in a sequence. Over time the inter-sound-interval was decreased until eventually the sounds were presented simultaneously (Caruso et al., 2018). Monkey Y was trained on all conditions except

the control conditions (52 of 56 conditions). Monkey N was trained on the final task, including all conditions.

During behavioral training and recording sessions monkeys sat with head movements restrained in a dark, sound attenuated room. A fixation light (LED) at 0° in both the horizontal and vertical dimensions signified the start of the trial. Monkeys would begin the trial by holding fixation on the light. After a variable duration of fixation (600-700 ms) either a single sound or two sounds would be presented while the monkey held fixation. After a variable overlap of fixation and sound presentation (1000-1100 ms) the fixation light was turned off cueing the monkey to saccade to the origin of the sound or sounds. The monkey had a brief temporal window for their eye position to leave the fixation window and for their eye position to enter the first target window. The monkey then had to hold fixation in the target window briefly before initiating a second saccade to the second target window or until the single sound trial ended. If the monkeys performed the trial correctly, they were rewarded with a grape juice solution (30% Welches grape juice (either red or white), 70% water). Reward was delivered through a drinking tube and controlled via a solenoid.

Eye position was acquired through an implanted scleral coil and Riverbend field system. Eye coil outputs were calibrated each day before behavioral and recording sessions. All behavioral data, including eye position data, were saved in Beethoven files for offline analysis.

Sound Stimuli: All sounds were generated using proprietary software (Beethoven, RYKLIN Software Inc.) and then passed through appropriate amplifiers to two speakers matched in their frequency response characteristics. Speakers were positioned at ±12° in the horizontal azimuth and at 0° in elevation at a distance approximately 1 meter from the monkey. Sounds were calibrated using a microphone at the position where the monkey's head would be during the experiment. Sounds were calibrated to be 55 dB SPL when presented in isolation; these stimuli resulted in 60 dB SPL sound reaching the monkey on dual sound trials. The intensity difference is not expected to have substantive effects (Caruso et al., 2018). We kept the sound source signals constant across conditions because that is the real-world case; simultaneously occurring sounds are louder than when the component sounds are presented in isolation. All experiments were conducted in an anechoic chamber; the wall and ceilings were lined with echo attenuating foam (Sonex) and the floor was carpeted.

All sounds were bandpass filtered sounds with pass bands of ±20% of the average center frequency. Sounds were frozen for each experimental session; that is for any given experimental session every instance of the bandpass noise stimulus of a given frequency involved the same waveform. The center frequencies were equally spaced in 0.19 octave intervals: 420, 500, 595, 707, 841, 1000, 1189, 1414, 1682 and 2000 Hz. These sounds were chosen due to the large proportion of IC neurons responsive at these frequencies (Bulkin & Groh, 2011), increasing the chance that more than one sound would significantly drive any recorded neuron.

Across all experimental sessions there were 56 conditions, involving either one or two sounds in combination. For simplicity, we will refer to the sounds in two subsets, A sounds and B sounds. A sounds consisted of the 8 middle sound frequencies (500 to 1682 Hz) while the B sounds were either 420 or 2000 Hz. On single sound trials any one of the 10 (8 A + 2 B) sounds could be presented in isolation at either ± 12°. On dual sound trials, an A sound was presented at either ± 12° and paired with a B sound at the opposite location. The dual and single sound trials accounted for 52 of the 56 conditions (8 A sounds x 3 paired conditions (none, 420 or 2000 Hz) x 2 locations = 48 + 2 B sounds presented in isolation x 2 locations = 56 conditions). To ensure

monkeys were actively perceiving both sounds and not just making 2 saccades to remembered locations in response to more complex spectral content (e.g. the dual sound trials have power in more bands than single sound trials), a control trial-type was added, and accounted for the last 4 conditions. The control trial presented a 1189 Hz sound with either B sound at either ±12°, that is to say the 1189 Hz and B sound were targeted to the same speaker location (2 sounds (1189+420 Hz or 1189+2000 Hz) x 2 locations = 4). Monkeys typically made single saccades on these control trials and reached high performance levels.

To acquire enough trials per condition and to accommodate cell holding times (typically 1-2 hours), the number of conditions per experimental recording session were limited by presenting the A sounds at only one location (either ±12°) each day. Therefore, each recording session consisted of 32 conditions. The eight A sounds presented alone or in combination with either B sound accounted for 24 of the 32 conditions (8 sounds x 3 paired tones x 1 location = 24). The 2 B sounds were still presented at both locations, this allowed for counterbalancing of single trials appearing at both locations (2 sounds x 2 locations = 4). The control conditions were not changed across days and accounted for the last 4 conditions. (2 sounds x 2 locations = 4). However, the control condition was implemented after monkey Y recordings were completed, therefore monkey Y recording sessions only consist of 28 conditions.

Conditions were randomly interleaved: any given trial had a 45% chance to be a single-sound trial, a 45% chance to be a dual-sound trial, or a 10% chance to be a control trial (or for monkey Y a 50% chance to be either a single- or dual-sound trial). Within these trial types the likelihood of any condition was equal. Therefore, any given trial had a 3.75% (45% / 12 single conditions) chance to be any single condition, a 2.8125% (45% / 16 dual conditions) chance to be any dual condition, or a 2.5% (10 % / 4 control conditions) chance to be any control condition (or for monkey Y a 4.167% chance to single-sound or a 3.125% chance to be dual-sound).

Neural Recordings: Neural activity was acquired through a Multichannel Acquisition Processor (MAP system, Plexon Inc). Single unit spike times were sent to Beethoven and saved for offline analysis. Additionally, single unit and multiunit signals were sent to an external speaker for auditory detection of neural activity fluctuations.

Prior to each recording session the chamber was cleaned and a grid and turret (Crist Instrument, 1mm spaced grid) were inserted inside the recording chamber. Prior to recordings, grid positions likely to intercept the IC were localized using structural MRI (Groh et al., 2001). A single tungsten micro-electrode (FHC, ~1-5 MΩ) was backloaded into a stainless-steel guide tube. The guide tube was then manually inserted through the dura. To prevent large deflections of the electrodes, when advancing to deep structures like the IC, the guide tube extended into the brain roughly 1cm. The electrode was then advanced with a microdrive (NAN Instruments) at 100-200 µm/s until the electrode tip was just outside of the guide tube. This was indicated by a mark on the electrode placed when the electrode was backloaded into the guide tube and confirmed by a change in the hash typical upon electrode exit of the guide tube and entrance into the brain. The depth on the microdive was then zeroed. The electrode was then advanced at roughly 0.8 µm/s while monkeys sat passively listening to band pass filtered sounds of variable frequencies presented every few seconds. Once the expected depth and/or sound related activity (determined by ear) was encountered in the hash, movement was stopped, and the electrode was allowed to settle for approximately 1 hour. The settling depth typically occurred 22 mm after exiting the guide tube. The electrode was then advanced in 1 µm increments until a cell was well isolated. Single units were isolated on-line using the box method (Plexon SortClient software).

If no cells were well isolated the electrode was retracted at roughly 8 µm/s. It was common to retract due to isolation failure.

One hundred and five neurons were recorded. Forty-five of these neurons came from the monkey Y's right IC. Fifty-nine of these neurons came from monkey N's right IC and 1 neuron was recorded from monkey N's left IC.

Analysis: *Behavior:* All behavioral analysis was done on attempted trials. Attempted trials were defined as: the monkey fixated during the entire fixation and left the fixation window after the go cue. Proportion correct was measured as the proportion of attempted trials performed correctly: Proportion correct = # of Correct Trials / # of Attempted Trials

Neural Activity: All analysis of neural activity was done on correct trials and during the first 500 ms after sound onset. In this bin the monkeys were fixating straight ahead which should limit the confounds by eye position (Bulkin & Groh, 2012; Groh et al., 2001; Porter et al., 2006) or eye movements (Bulkin & Groh, 2012; Porter et al., 2007) known to occur in the macaque IC. Neurons were classified as responsive if their spike count 500 ms after sound onset across all tested frequencies was significantly different from their spike count during baseline, the 500 ms prior to sound onset (t-test, $p < 0.05$). Ninety-three of the 105 neurons were classified as responsive. Neurons were classified as frequency selective, or tuned, if their spike counts 500 ms after sound onset were significantly modulated by frequency (one-way ANOVA, $p < 0.05$). Sixty-three of 105 neurons were classified as frequency selective.

Point Image: The point image refers to the proportion of recorded units that are responsive to a given condition (Capuano & McIlwain, 1981). Neurons were classified as responsive if, for a given condition, their spike count 500 ms after sound onset was significantly different than their spike count 500 ms prior to sound onset (t-test, $p < 0.05$).

d' (d-prime): To investigate how differently a cell responded for A and B sounds we calculated a sensitivity index or d':

$$d' = \frac{\mu_A - \mu_B}{\sqrt{\frac{1}{2}(\sigma_A^2 + \sigma_B^2)}}$$

Where μ_A and σ_A^2 corresponds to the mean spike count and spike count variance for sound A, respectively, and μ_B and σ_B^2 corresponds to the mean spike count and spike count variance for sound B, respectively. This metric quantifies the separability of the sound A and sound B spike count distributions. We took the magnitude of the d' value (|d'|) since the order in the numerator is arbitrary (the A or B sound could be the first term). A |d'| value of 0 indicates the sound A and sound B spike count distributions are similar and higher values of |d'| indicate the sound A and sound B spike count distributions are more distinguishable.

Equivalent Rectangular Receptive Fields (ERRFs): ERRFs were calculated in a similar fashion to previous work (Lee & Middlebrooks, 2011). EERFs were calculated by using a non-baseline subtracted tuning curve as follows: The area under the tuning curve was measured with the *trapz()* function in MATLAB, which integrates the area using the trapezoid method. We used linear interpolation between data points so there is no ambiguity in the area under the curve using this method. This area was then divided by the peak firing rate of the tuning curve, creating a rectangle with a width dependent upon tuning curve breadth and how frequency modulates the peak firing rate and the firing rate across all sounds.

Maximum-Likelihood Decoding: The decoding algorithm used here was adapted from previous work (Wallisch, 2014) and is similar to other implementations of maximum-likelihood decoding (Day & Delgutte, 2013; Jazayeri & Movshon, 2006). The goal of the decoder was to infer the stimulus that was presented on a set of held-out trials (1 held-out trial of a given condition per cell). The predicted condition was the condition (1 out of 26 possible conditions) that had highest population-log-likelihood. This computation was repeated 1000 times per condition. Since there were not 1000 unique trials per condition for each cell the particular held-out trial on a given repeat was sampled from all possible trials with replacement.

To determine the population-log-likelihood first the likelihood for each cell was computed given the held-out trials spike count. First, the weights were determined by finding the average spike count for each condition without the held-out data. Then likelihood for all 26 conditions given the spike count was computed:

$$\ln\,(Condition\ Likelihood) = \ln\,(P(spike\ count \mid frequency))$$

This results in a vector with 26 elements for each cell. Each element corresponds to the log-likelihood of that condition for that held-out spike count. Essentially the log-likelihood is given by the difference between the held-out spike count and the mean firing rate across all conditions. These log-likelihood vectors are then summed across all cells to generate a population-log-likelihood:

$$\ln(Population\ Condition\ Likelihood) = \sum_{i=1}^{N} \ln\,(P(spike\ count_i \mid stimulus)$$

Where spike $count_i$ corresponds to the *ith* neuron spike count, and N is the number of neurons in the population. Therefore, for each repeat of the decoding computation, the result is a 26-element vector where each element corresponds to the population-log-likelihood for each condition. The predicted condition for that repeat is the element with the largest population-log-likelihood. Importantly, for any repeat the inferred condition could have been one of the 26 conditions (8 A sounds x 3 sound pairs = 24 conditions + 2 B sounds = 26 conditions). We used all responsive units (N = 93) for the decoding analysis.

Results:

Two monkeys were trained on a sound localization task (see Methods) adapted from a previous study (Caruso et al., 2018). In this task monkeys reported the location of a sound presented alone or the locations of two sound presented simultaneously (Fig. 1A). Monkeys perform this task quite well. Of all attempted dual-sound trials monkeys perform roughly 90% correct (Fig. 1B), slightly better than the 80% correct observed on single sound trials. This high degree of performance on the dual-sound trials was only modestly affected by the frequency separation between the two sounds: monkeys performed at about 87% correct even for the smallest frequency separation of 0.19 octaves (Figure 1C), rising slightly to 91% correct for the largest frequency separation of 2 octaves.

This good performance does not appear to be the result of recruiting largely unique sets of neurons to encode each of the two stimuli. One hundred five neurons were recorded across two monkeys while they performed the sound localization task, 88% (N = 93) of these neurons were responsive to sounds. The majority of sound responsive neurons significantly responded to the majority of the stimuli when presented individually (Fig. 2, red curve). The lowest responsive condition across all conditions, still, significantly modulated the firing rate of approximately 44% (N = 41) of sound responsive neurons, and the most responsive condition evoked activity in

about 70% (N = 67) of IC neurons. The large proportion of responsive units suggests that sounds presented in this experiment recruit overlapping populations of neurons; and, along with the low frequency bias, is similar to previous macaque IC findings (Bulkin & Groh, 2011). Importantly, for nearly every dual-sound condition the proportion of responsive units increases compared to the single sound condition (Fig. 2, blue and black curves vs. the red curve). This increase in responsiveness is the opposite of what is expected if units were assumed to sharpen their tuning.

So, how sensitive are individual IC neurons to sound frequency? Roughly 67% (N = 63) of responsive neurons had spike counts that were significantly modulated by frequency across all 10 sounds frequencies presented at the same location (one-way ANOVA, $p < 0.05$). To see if this frequency selectivity may underlie the slight improvement in behavioral performance as a function of frequency separation, we implemented a d' analysis (see Methods). This analysis quantified the separability of the spike count distributions corresponding to the component sounds of the dual-sound conditions (e.g. 500 (A-sound) vs. 420 (B-sound) when either were presented in isolation). Figure 3 shows a population of neurons that increase their |d'| as the A sound is more separate from the B sound (darker colors on the right side of panel 3A and the left side of 3B). Note that the unit order in two panels are different; they're sorted so the unit with the highest summed |d'| across the panel corresponds to unit 1. Importantly, figure 3 also shows a large minority of cells do not differentiate between any of the A and B sounds, evidenced by the white bands at the bottom of the figure 3 (roughly unit 70 to 93). Overall, this analysis suggests that largely overlapping populations of neurons are recruited for any given sound, and that only a fraction of the recorded populations (roughly the top 20 units in both panels) possess the proper sensitivity to underlie the improvement in performance due to separation of the frequencies in the dual-sound conditions.

Could, then, neural responses change when two sounds are presented in some way to facilitate the encoding of both sounds? Figure 4 shows 6 example cells, comparing their single-sound and dual-sound tuning profiles. On single-sound trials (red traces), neurons could prefer low frequencies (Fig. 4A, E, F), intermediate frequencies (Fig. 4B, D), or high frequencies (Fig. 4C). Concurrent sounds could change tuning in multiple ways. The neurons in (Fig. 4A-C) all become less responsive to sounds overall: the responses in the dual sound cases were generally lower than the responses to the "better" of the two single sounds (e.g. black curves fall between the red and blue curves). Some neurons appear to shift their preferred sound (Fig. 4D), and some neurons appear unaffected by the presence of an additional sound (Fig. 4E, F). Interestingly, some neurons displayed multiple peaked tuning curves that were robust across single and paired-sound conditions (Fig. 4F). Qualitatively, across the population it seems that the majority of neurons become less responsive in the presence of two simultaneous sounds, but whether this suppression leads to receptive fields that sharpen or broaden requires quantification at the population level.

We first statistically assessed how neural firing was modulated by frequency on dual-sound trials compared to single sound trials using one-way ANOVAs calculated separately for dual and single sound conditions. Approximately 61% of neurons (N = 57) had spike counts that were significantly modulated by sound frequency when the A sounds were presented in isolation. This number dropped to roughly 52% (n = 47) when the A sounds were paired with either 420 or 2000 Hz. To evaluate effect size, we compared the F-statistic (a metric of how much frequency modulates spike count) across these conditions. Figure 5A and 5C plot the F-statistic when sounds are presented in isolation against when they are paired with 420 or 2000 Hz, respectively. The F-statistic is larger when sounds are presented in isolation compared to when sounds are

presented with 420 Hz (Fig. 5A). The difference distribution (420 paired F-Statistic – sound alone F-Statistic per cell) is significantly less than zero (mean = -7.989 ±1.219 S.E.M, $p < 0.05$; Fig. 5B); which indicates that the majority of neural firing rates are less modulated by frequency in the 420 Hz paired condition. However, when sounds are presented with 2000 Hz IC neuron spike counts are equally modulated by sound frequency (Fig. 5C). The difference distribution (2000 paired F-Statistic vs. sound alone F-statistic per cell) is not significantly different than zero (mean = -1.142±0.899 S.E.M, $p = 0.207$; Fig. 5D). This analysis suggests that neurons become less selective for sound frequency when sounds are paired with low frequency sounds; however, the F-statistic does not provide information concerning if tuning curves narrow or broaden.

To quantify the change in tuning curve width across each unit we implemented the equivalent rectangular receptive field analysis (ERRF) from (Lee & Middlebrooks, 2011). Briefly, the ERRF takes the area under the tuning curve and divides the area by the peak firing rate, creating a rectangular receptive field with a width that is dependent upon both frequency selectivity and how frequency modulates firing rate. For example, many of the cells display a suppression in the dual sound condition (Fig. 4 A-C), yet these suppressions seem to be of different kinds. Generally, suppressions that compress the overall range in firing rate (Fig. 4B & C) tend to increase ERRF width while suppressions that suppress selective conditions but allow for the same range of firing rate (Fig. 4A) tend to decrease the ERRF width. To compare the single-sound tuning curve width to the dual-sound tuning curve width, the distribution of each was plotted against each other (Fig. 6B & D). If there ERRF widths were the same they would fall upon the unity line (black line in Fig. 6B, D). If the ERRF widths were narrower in the dual-sound condition the distribution would be shifted below unity and if the ERRF widths were narrower in the single sound condition the distribution would be shifted above unity. It is quite apparent in both, but especially in the 420 Hz paired condition that the ERRF widths are less narrow than the single-sound ERRF widths. To quantify if these width changes are significant, the difference between each neuron's dual-sound ERRF width and single-sound ERRF width was calculated. The 420 Hz paired ERRF widths were significantly broader than the single-sound ERRF widths (Fig. 6C); the difference distribution was significantly greater than 0 (mean = 0.75±0.11 S.E.M, $p < 0.05$). Although the effect was smaller the 2000 Hz paired ERRF widths were also significantly broader than the single-sound EERF widths (Fig. 6D); the difference distribution was significantly greater than 0 (mean = 0.26 ± 0.11 S.E.M, $p < 0.05$). In summary, across the population the majority of neurons actually broaden their tuning curve in response to presentation of two simultaneous sounds, especially if the simultaneously presented sound is low frequency.

So far there is evidence that IC neuron activity is less modulated by frequency and their tuning curves broaden in the dual sound condition. However, could some neurons be shifting their tuning curves (e.g. Fig 4D)? Shifting of tuning curves may be a useful way of assigning neurons with preferences intermediate to either sound into subpopulations that encode one of the presented stimuli. To investigate if neurons shifted their tuning curve, we selected neurons with firing rates significantly modulated by frequency and with single-sound best frequencies that fell within the middle six A sounds (595-1414 Hz, leaving out the edge cases for which a shift outside the tested range could not be observed). We then compared their single-sound best frequency to their best frequency in the dual sound conditions. Seventeen neurons met these criteria; although this population is small, if a shift in receptive field is present it should be observable in this population. Figure 7 shows the best frequency for the 17 cells and how they change in the 420 Hz paired condition (Fig. 7A) or the 2000 Hz paired condition (Fig. 7B).

Although, there is a slight shift of the mean BF away from the paired sound, it is not significant for either the 420 Hz paired sound (single mean = 759.3 ± 76.48 S.E.M vs. dual mean = 937.6 ± 113.6 S.E.M, $p = 0.072$) or the 2000 Hz paired condition (single mean = 759.3 ± 76.48 S.E.M vs. dual mean = 740.9 ± 75.67 S.E.M, $p = 0.56$). These findings suggest that shifting of receptive fields is not a ubiquitous computation within the population and is unlikely to substantially contribute to the encoding of simultaneous sounds.

In short, the modulations of frequency tuning on dual sound trials do not appear to be optimized for the preservation of information about either sound. To determine the impact of multiple sounds on the efficacy of coding stimulus related information, a population-maximum-likelihood decoder was implemented to infer the condition from only the spike count on a set of held-out trials (see Methods). We assumed Poisson firing distributions and independent neurons. So, for each held-out trial, the log-likelihood for each condition was summed across the population and the condition with the maximum-population-log-likelihood was predicted to be the held-out trials condition. There were 1000 held-out trials per condition. Overall, the decoder was able to decode every condition above chance (1/26 = ~ 3.846%). Figure 8A, C, and E shows the proportion of correctly predicted conditions across the 3 paired sound conditions, sounds presented in isolation, sounds paired with 420 Hz, or sound paired with 2000 Hz, respectively. Qualitatively, the mean across all sounds presented in isolation (~66.7% correct) was slightly higher than the mean across all 420 Hz paired conditions (~47.1% correct) or the mean across all 2000 Hz paired conditions(~51.8% correct). The likelihood functions show that the confusion in predicting the correct condition tends to come from nearby conditions (Fig. 8B, D, F). This can be further seen in the confusion matrix (Fig. 9A) where the off diagonals are slightly darker in the dual conditions relative to the single-sound conditions. This suggests that the decoder is largely confusing proximal conditions (e.g. 420+500 vs 420+595). This is likely due to the likelihood functions, in general, being much higher when there is a low frequency sound present (420 Hz paired and when 2000 Hz is paired with a low frequency: Fig. 9B). The changes in the frequency response functions ultimately seem to degrade the amount of information concerning each sound resulting in slightly worse performance in the decoding in the dual-sound conditions.

Until now all the analyses have investigated how the monkey IC encoding of sound is altered by the presence of an additional sound. With the implementation of the maximum-likelihood decoder, we can ask if information about the sounds is encoded similarly when sounds are presented in isolation verses when sounds are presented simultaneously. To investigate if the monkey IC uses the similar codes for the single and dual-sound conditions, we attempted to decode the dual sound condition using the decoding weights learned from the single sound trials. Each held-out dual sound trial could, therefore, be labeled as one of the 8 paired sounds. Performance was markedly worse in the 420 Hz condition when decoded with the single-sound weights (accuracy = 16.5%; chance 12.5%) compared to the dual-sound weights (accuracy ~ 47.1%; chance ~ 3.846%). This is evident in the confusion matrix (Fig. 10A) which shows 595 Hz is the most predicted stimulus for seven of the eight conditions. Although to a lesser degree, the performance in the 2000 Hz condition when decoded with the single-sound weights (accuracy = 40.65%; chance = 12.5%) is also worse compared to the dual-sound weights (accuracy ~ 51.8% correct; chance ~3.846%). Unlike the 420 Hz paired condition, the 2000 Hz paired conditions seem to confuse proximal conditions (Fig. 10B). These results show that the IC may use different coding strategies in the single-sound and paired-sound conditions, which implies that the IC's readout might have to use different weights depending on the number of sounds

Discussion

How the brain simultaneously encodes multiple items, particularly when more than one item falls within a given neuron's receptive field, is an obvious yet often overlooked problem in neuroscience. It is often assumed the population representation solves this, but the problem remains if the items recruit largely overlapping populations. To address this question, we first verified that monkeys can relatively easily report the occurrence of two bandpass noise stimuli separated by as little as 0.19 octaves. We then investigated how frequency response functions of monkey IC neurons are affected by the presence of that second, concurrently presented sound. We found it likely that the sounds we presented recruited largely overlapping populations (Fig. 2 & 3), and that the majority of these cells' firing rates were less modulated by sound frequency (Fig. 5) and had their tuning curves broadened (Fig. 6). There was little evidence of systematic shifts in best frequency (Fig. 7).

Overall, these changes in frequency response functions reduce the information available to the maximum-likelihood decoder, decreasing the decoding accuracy in the dual-sound conditions (Fig. 8 & 9). The especially poor decoding performance when using the single sound weights to decode out dual sound conditions (Fig. 10) suggests that a different read-out is needed depending on the number of sounds – a startling possibility given the brain cannot have prior knowledge of how many sounds are in the environment except by virtue of what it detects via sensory input.

Our results seem to be in-line with prior work showing degradation of information when there are multiple stimuli presented (Day et al., 2012; Day & Delgutte, 2013; Henry & Kohn, 2020), and if this reduction in information scales with number of stimuli it could be potentially explain the finding that number of identifiable sounds saturates at three (Zhong & Yost, 2017). Generally, these results refute the notion that the changes in frequency response functions could be used to overcome the multiplicity problem and suggest recent evidence of alternations in firing rates between those evoked by individual stimuli (Caruso et al., 2018) may be a more fruitful research direction.

In the present study, we made the simplifying assumption that the time-and-trial-pooled average activity of a neuron to a particular combination of stimuli is reflective of the information that neuron encodes. If the underlying activity on dual sound trials actually fluctuates between the A-like and B-like response patterns, the time-and-trial-pooled average will be a poor measure of the information present in neural signals. The fact that average responses to AB sounds were generally between the average responses to A and B sounds presented alone supports the possibility that fluctuations may underlie at least a portion of the results observed here. Future work will test this possibility.

In principle, changes in frequency tuning as investigated here and fluctuating activity patterns as investigated previously (Caruso et al., 2018) both have the potential to limit the degree to which a given neuron is faced with the task of encoding more than one stimulus at a time. It remains possible that information-preserving changes in frequency tuning might be more evident when tested with a wider range of sound frequencies, allowing a fuller exploration of frequency tuning curves. However, the granularity of our testing, 0.19 octave spacing, was clearly coarser than monkeys' perceptual abilities, and the overall range, 2 octaves, should have been adequate to demonstrate an effect if changes in frequency tuning were a major contributor to these perceptual abilities.

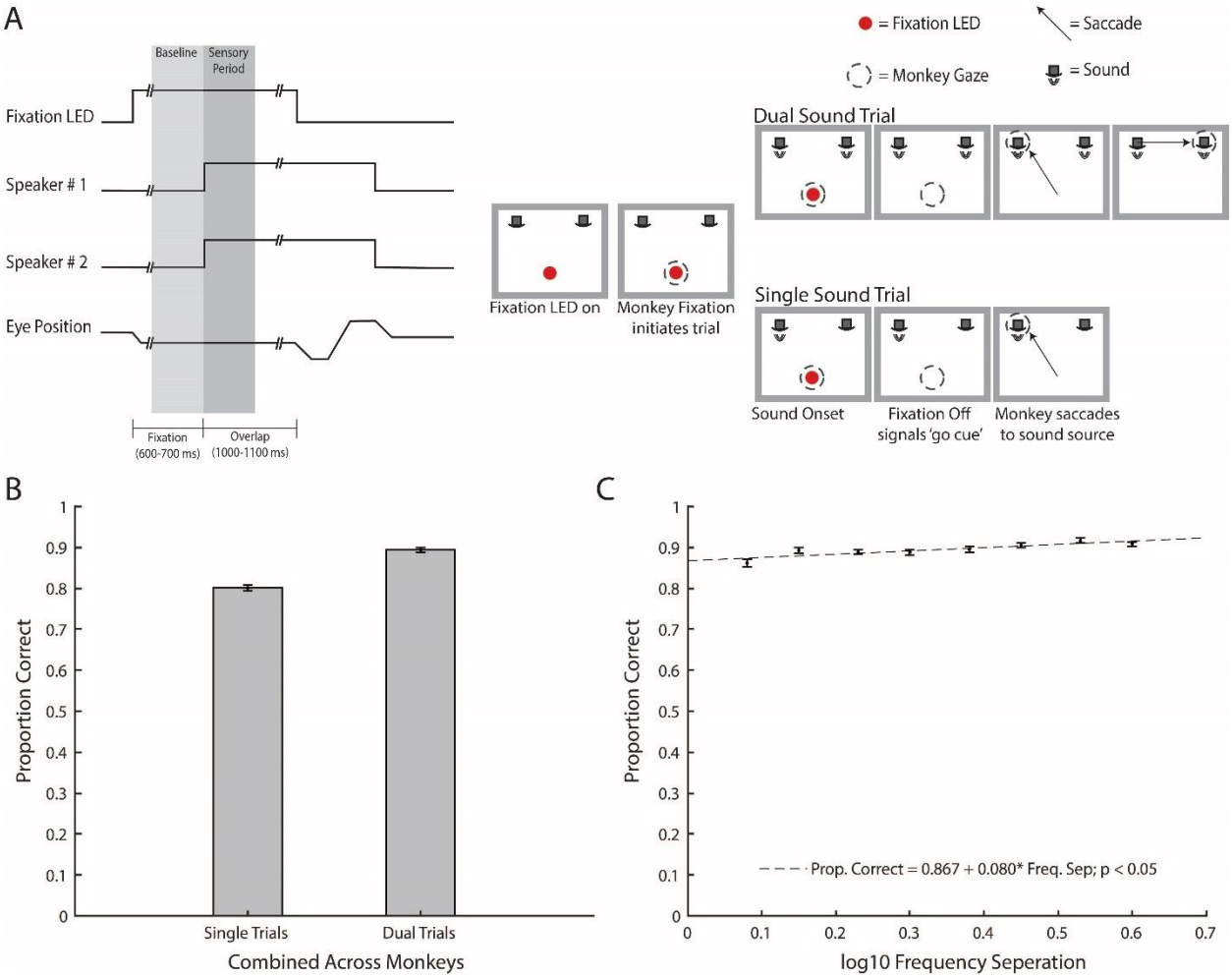

Figure 1 Task schematic and behavioral performance. (**A**) Monkeys were trained on a sound localization task. Left panel: A time course plot of the task. Speaker 2 onset and the second saccade and fixation would only occur if it was a dual-sound trial. Right Panel: In the sound localization task the monkey would initiate a trial by fixating an LED. After variable interval fixation (600-700 ms), either 1, or 2 simultaneous, sound(s) were presented from one or two locations. After a variable interval of sound/fixation overlap (1000-1100 ms) the fixation light is turned off, which cued the monkey to report to the origin(s) of the sound(s) via saccade(s). On single sound trials monkeys made a single saccade. On dual-sound trials monkeys made a sequence of saccades to the origin of one sound and then the other. (**B**) The combined proportion of correct single-sound trials (mean = 0.8011 ± 0.0069) and dual-sound trials (mean = 0.8941 ± 0.0060) across both monkey Y and monkey N for 105 behavioral sessions collected during cell recordings. (**C**) Change in performance across both monkeys as a function of the frequency separation between the two sounds. Performance increases as distance between sound frequencies increase (significantly fit by a line with a slope > 0; slope 0.867, $p < 0.05$). Error bars represented standard error of the mean (**B & C**).

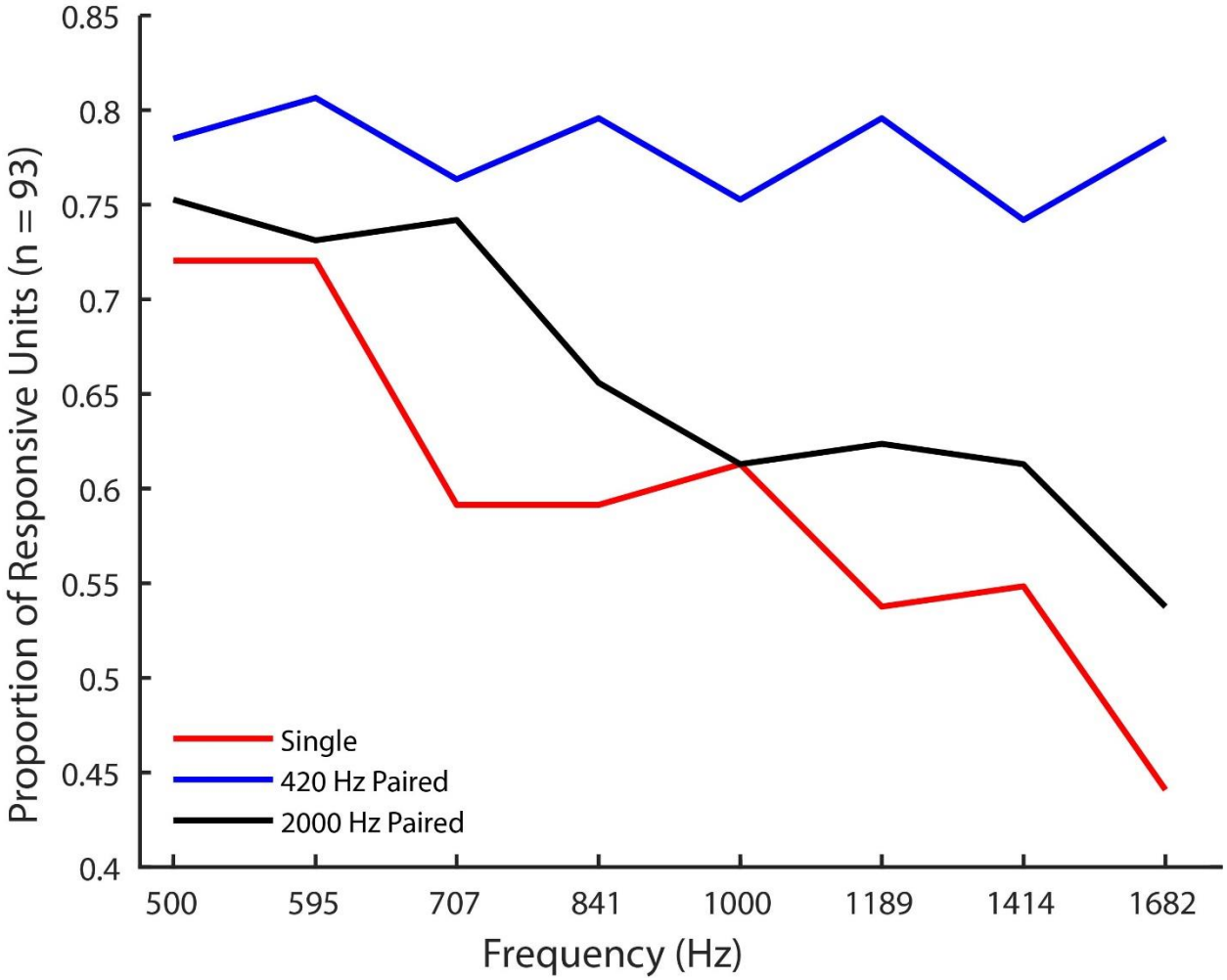

Figure 2 A point image plot, or the proportion of the responsive population that responds to a given frequency. Each point corresponds to the proportion of the responsive population that is significantly responsive compared to baseline for the given condition. A unit was counted as responsive if its firing rate in 0-500 ms after sound onset was significantly different than its firing rate from -500-0 ms before sound onset (t-test, $p<0.05$). Red corresponds to sounds presented in isolation, while blue and black correspond to dual-sound trials paired with 420 Hz or 2000 Hz, respectively.

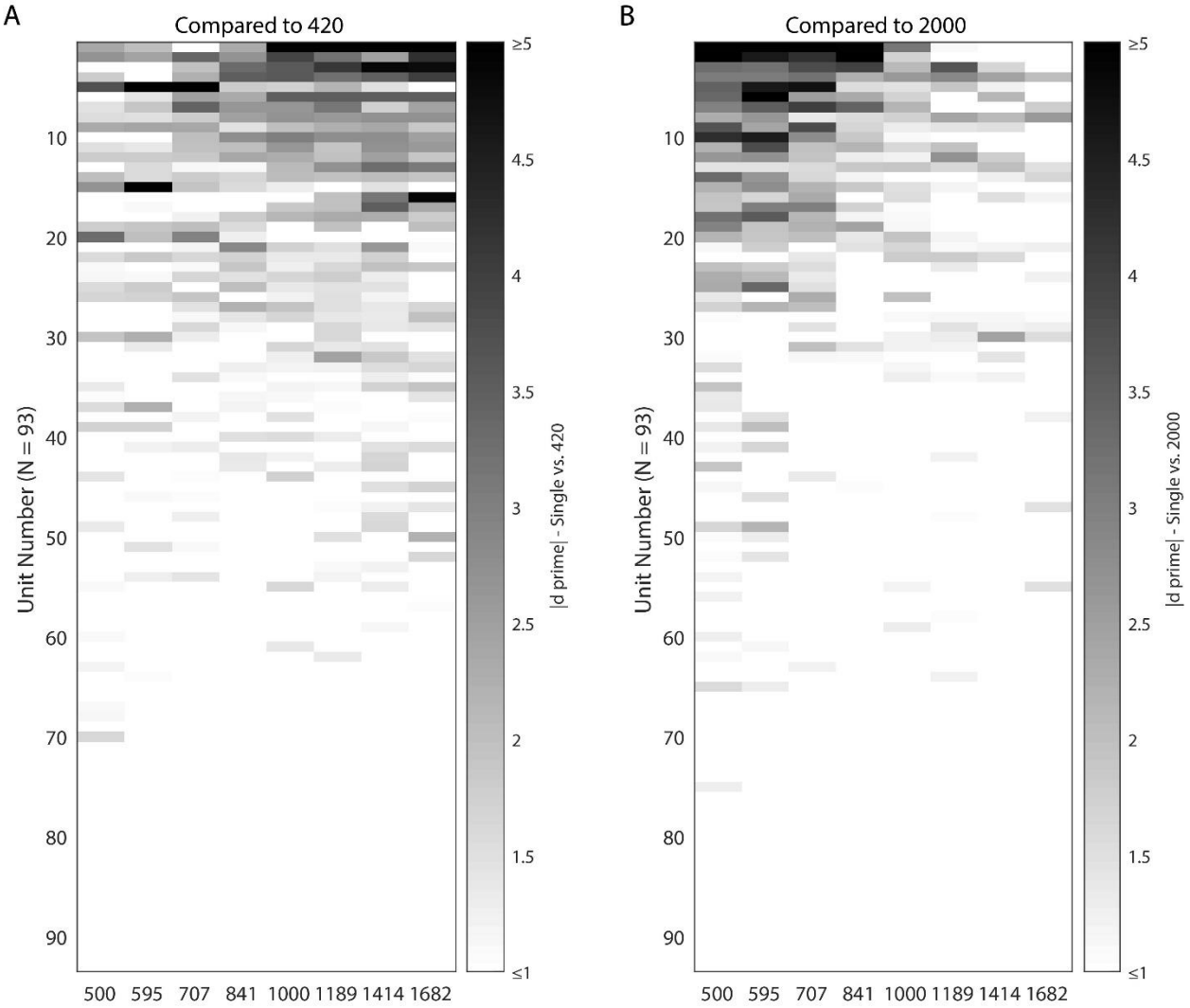

Figure 3 A sensitivity index, or d', between the A and B sounds when they are presented in isolation. (A) The magnitude of d' between each of the 8 A sounds and 420 Hz. A higher |d'| corresponds to a larger difference between the spike count distributions of the two sounds. Each row corresponds to a unit. Each column corresponds to a particular sound frequency. Units are sorted by their total |d'| across all 8 A sounds compared to 420 Hz. (**B**) The same as panel (**A**) but compared to 2000 Hz. Note these units are sorted by their total |d'| across all 8 A sounds compared to 2000 Hz. Therefore, unit n in **A** does not necessarily correspond to unit n in panel **B**.

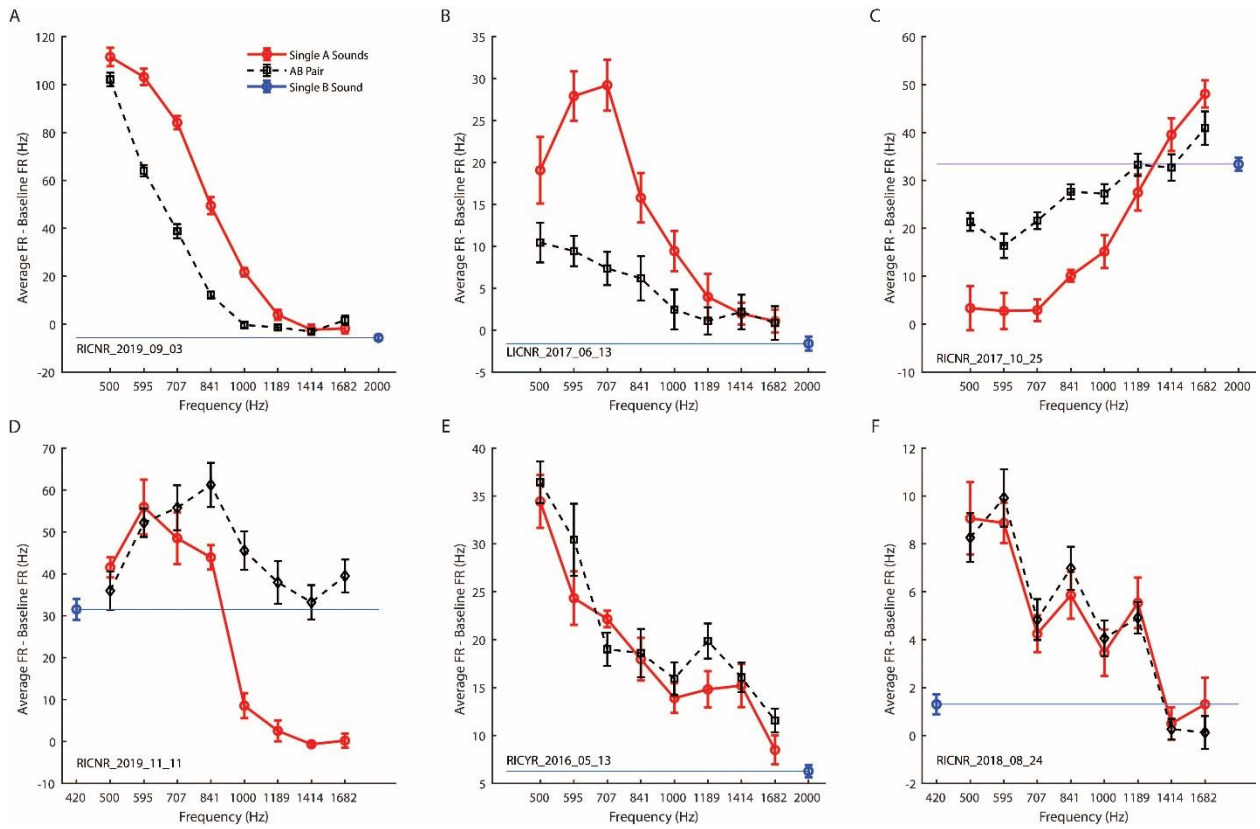

Figure 4 Example modulations of frequency tuning curves due to the presence of an additional sound in 6 different cells. (**A**) The tuning curve across conditions for a cell that displays sharpening in the dual-sound condition. Each point corresponds to the baseline (-200 ms bin w.r.t sound onset) subtracted mean firing rate (500 ms bin w.r.t sound onset). Red points correspond to sounds presented in isolation. The blue point corresponds to the paired sound when presented in isolation. The blue line is a reference line, so the firing rate of the paired sound is visible across the panel. The black points correspond to the dual-sound condition. (**B and C**) The same as panel (**A**) but for two different cells that show a decrease in response and selectivity. (**D**) The same as panel (**A**) but for a cell that shifts its preferred frequency away from the paired sound. (**E & F**) The same as panel (**A**) but for two cells that seem to be unaffected by the presence of an additional sound. Error bars represent standard error of the mean(**A-F**).

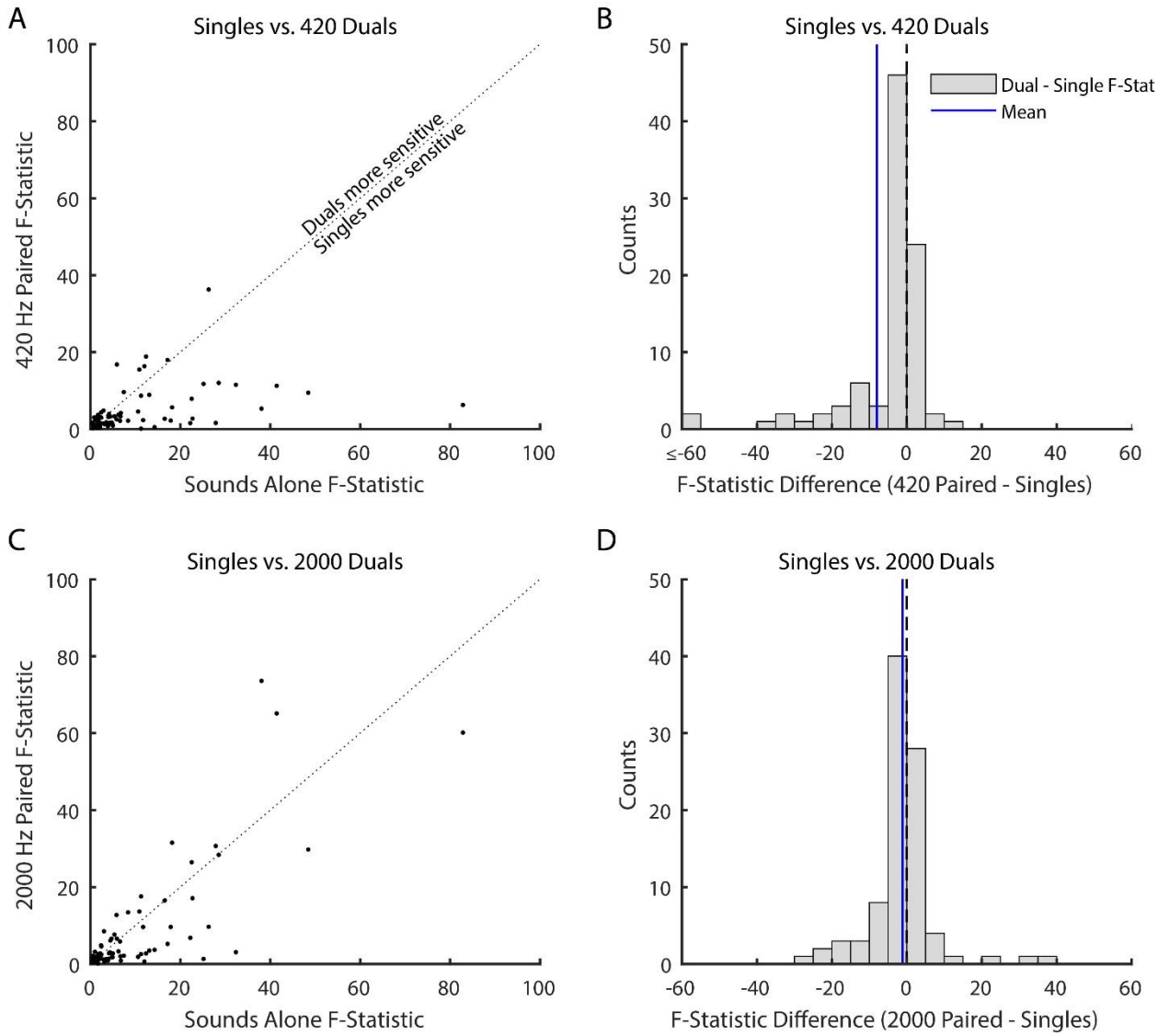

Figure 5 Frequency modulation of firing rate is affected by the presence of an additional sound. (**A**) The F-statistic measured from a one-way ANOVA investigating firing rate across different frequencies (all 8 A sounds). A high F-statistics indicates that there are large differences in firing rate across the different frequencies. The single-sound F-statistics are plotted against the 420 Hz paired F-statistics. Each point corresponds to one unit, out of 93 sound responsive units. Points above the unity line indicate that frequency modulates firing rate more in the 420 Hz paired condition. Points on the unity line indicate there is no change in the modulation of firing rate across the two conditions. Points below the unity line indicate that frequency modulates firing more in the single-sound condition. (**B**) The distribution of change in F-statistic per cell ($\text{F-Statistic}_{\text{Single}} - \text{F-statistic}_{\text{Dual}}$). A value of 0 indicates no change across the two conditions, a value greater than 0 indicates the 420 Hz paired condition was more sensitive to frequency and a value less than 0 indicates the single-sound condition was more sensitive to frequency. The single-sound condition was significantly more sensitive to frequency (mean = -7.99 ± 1.22 S.E.M, $p < 0.05$). The blue vertical line indicates the mean of the distribution. The black vertical line is at 0. (**C**) The same as (**A**) but for compared to the 2000 Hz paired condition. (**D**) The same as (**B**) but compared to the 200 Hz paired condition. There was no difference in the modulation of firing rate across the two conditions (mean = -1.14 ± 0.9 S.E.M, $p = 0.21$). Note that in (**A & C**) an outlier data point was removed for visualization purposes. In panel (**A**) its position is x =364.5, y = 31.2. In panel (**C**) its position is x = 364.5, y = 396.3.

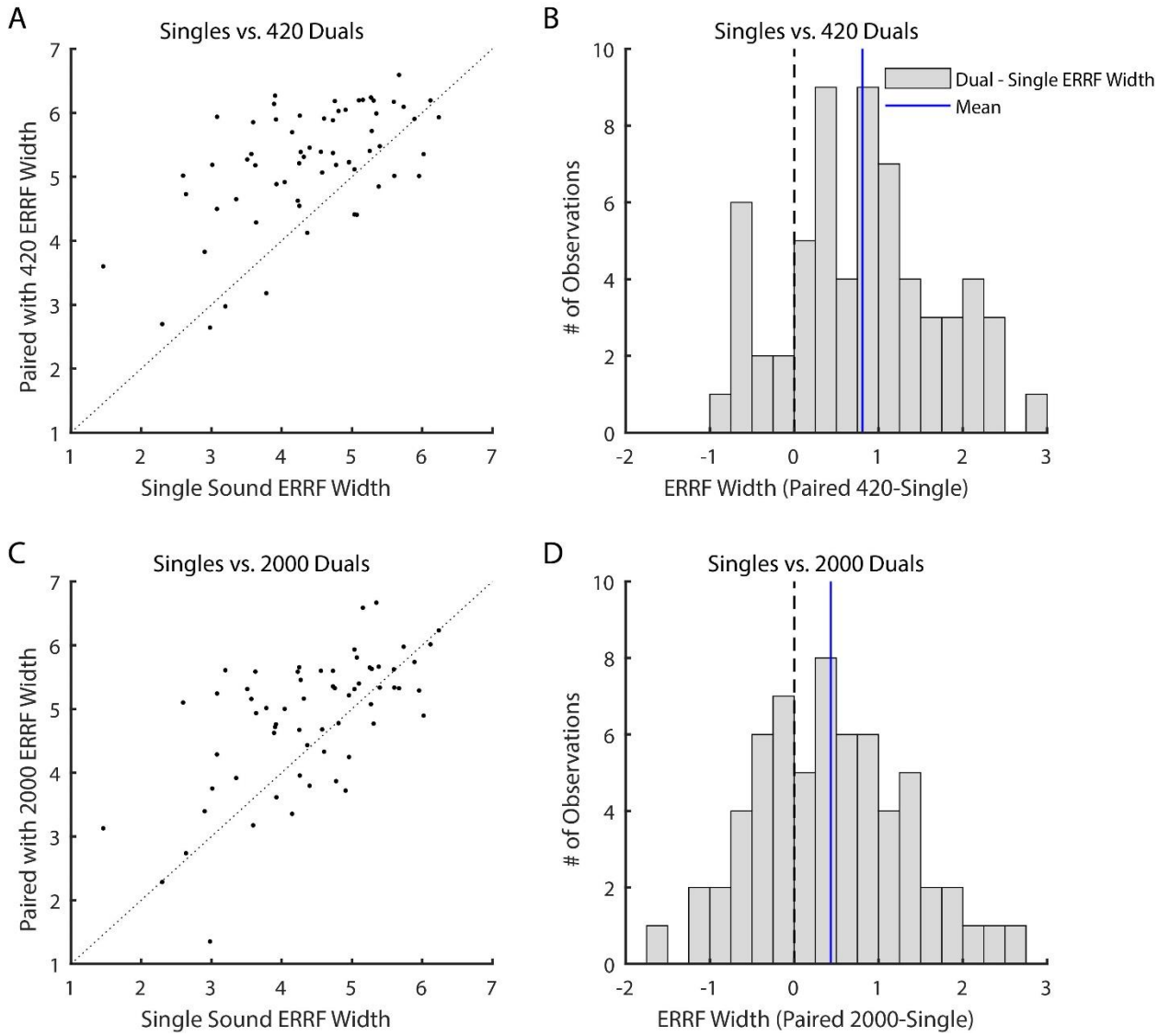

Figure 6 ERRF width is larger in dual-sound conditions. (**A**) The width of the ERRF for each unit in the single sound condition plotted against the width of the ERRF for each unit plotted in the 420 Hz paired condition. Each point represents a unit, out of 63 frequency modulated cells. The dashed line indicates unity, or ERRFs of equal width in the single sound and 420 Hz paired condition. Points above unity indicate ERRFs of greater width in the dual-sound condition. Points below unity indicate ERRFs of greater width in the single-sound condition. (**B**) The distribution of change in ERRF width per cell ($ERRF_{420} – ERRF_{Single}$). A value of 0 indicates the ERRF width did not change across conditions. Values greater than 0 indicate the ERRF is broader, while values less than 0 indicated the ERRF is narrower. The ERRF width in the 420 Hz dual sound condition is significantly broader (mean = 0.75±0.11 S.E.M, $p < 0.05$). The blue vertical line indicates the mean of the distribution. The black vertical line is at 0. (**C**) The same as (**A**) except for the 2000 Hz paired condition. (**D**) The same as (**C**) except for the 2000 Hz paired condition. The mean in the 2000 Hz paired condition is significantly broader (mean = 0.26 ± 0.11 S.E.M, $p < 0.05$).

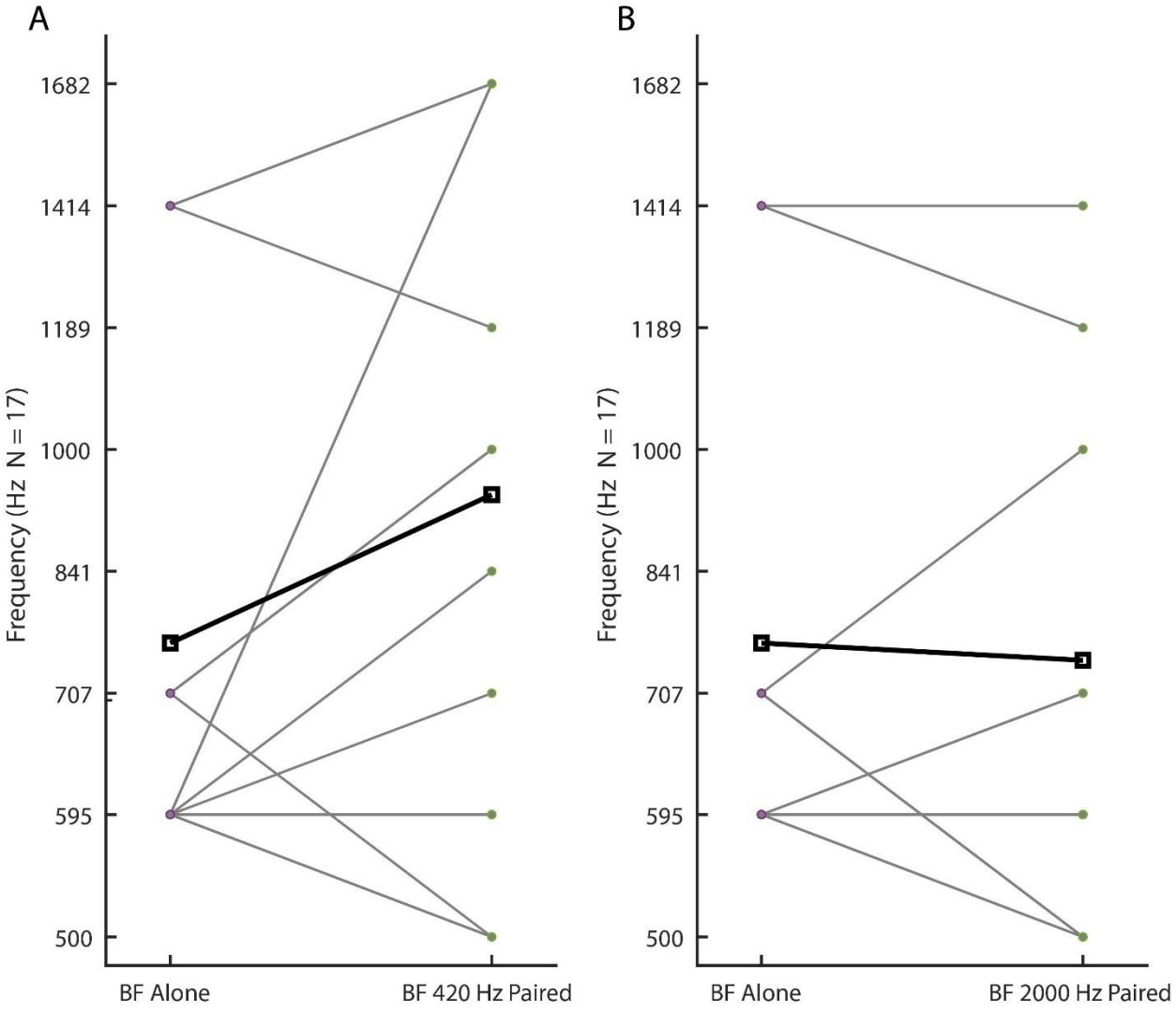

Figure 7 The presence of an additional sound does not cause a large shift in best frequency across the population of recorded cells. (**A**) The single-sound best frequency plotted against the 420 Hz paired best frequency in the 17 cells that have best frequencies between 595-1414 Hz. There is no significant shift in best frequency across the two conditions (single mean = 759.3 ± 76.48 S.E.M vs. dual mean = 937.6 ± 113.6 S.E.M, p = 0.072). (**B**) The same as (**A**) but for the 2000 Hz paired condition. There is no significant shift in best frequency across the two conditions (single mean = 759.3 ± 76.48 S.E.M vs. dual mean = 740.9 ± 75.67 S.E.M, p = 0.56).

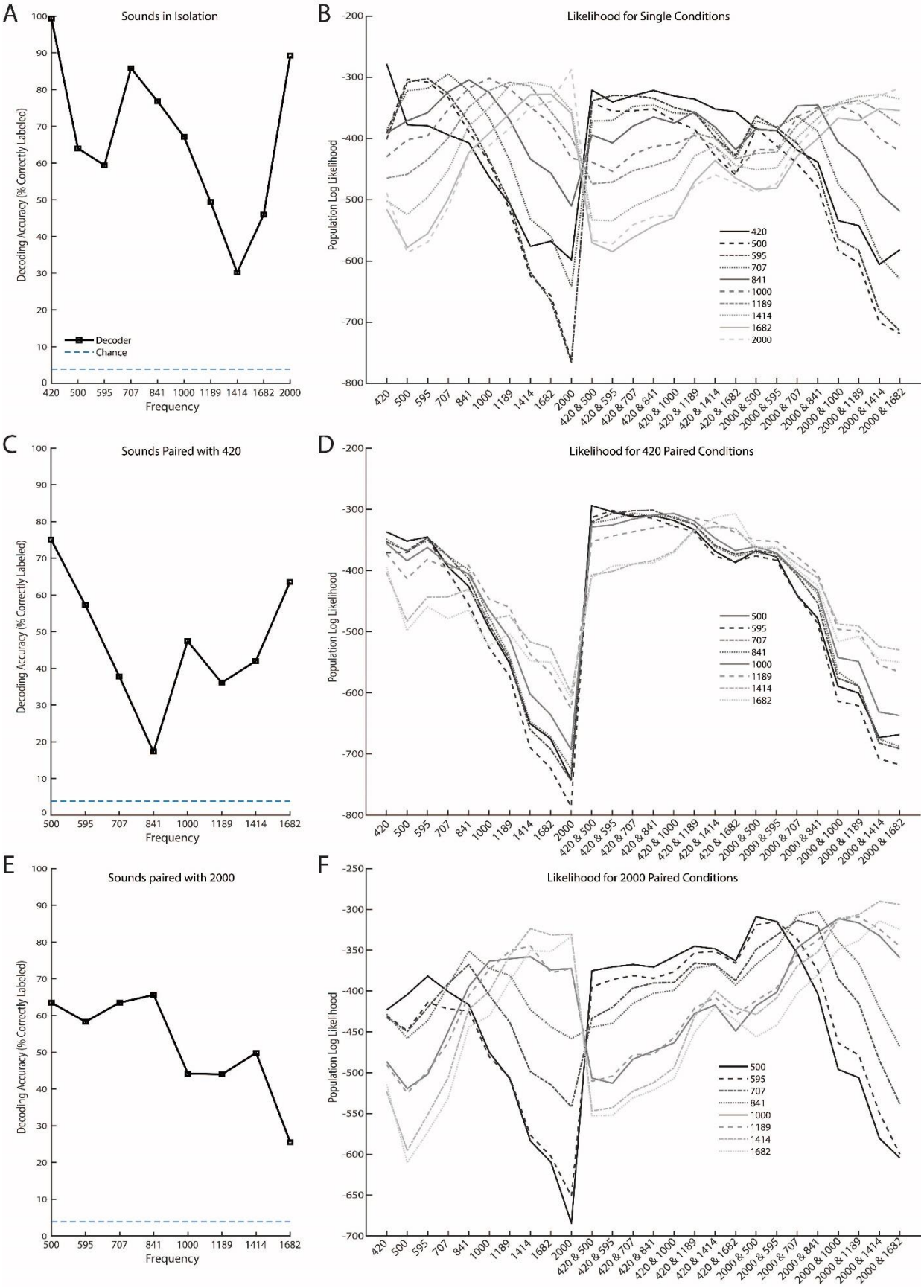
A
Sounds in Isolation
Decoding Accuracy (% Correctly Labeled)
Frequency
Decoder
Chance
B
Likelihood for Single Conditions
Population Log Likelihood
420
500
595
707
841
1000
1189
1414
1682
2000
C
Sounds Paired with 420
D
Likelihood for 420 Paired Conditions
E
Sounds paired with 2000
F
Likelihood for 2000 Paired Conditions

Figure 8 Accuracy of population maximum-log-likelihood decoding and mean log-likelihood functions across conditions. (**A**) The accuracy of the population (N = 93) maximum-log-likelihood decoder across each single sound condition. Each point corresponds to the percent of correctly labeled repeats (out of 1000) for each condition. Chance level prediction is 1/26 or ~ 3.85% (**B**) The mean population log-likelihood function for each condition across all potential predicted conditions. Each condition had the chance of being labeled as any one of the 26 single or dual conditions. (**C**) The same as panel (**A**) but for 420 Hz paired conditions. (**D**) The same as panel (**B**) but for 420 Hz paired conditions. (**E**) The same as panel (**A**) but for 2000 Hz paired conditions. (**F**) The same as panel (**B**) but for 2000 Hz paired conditions.

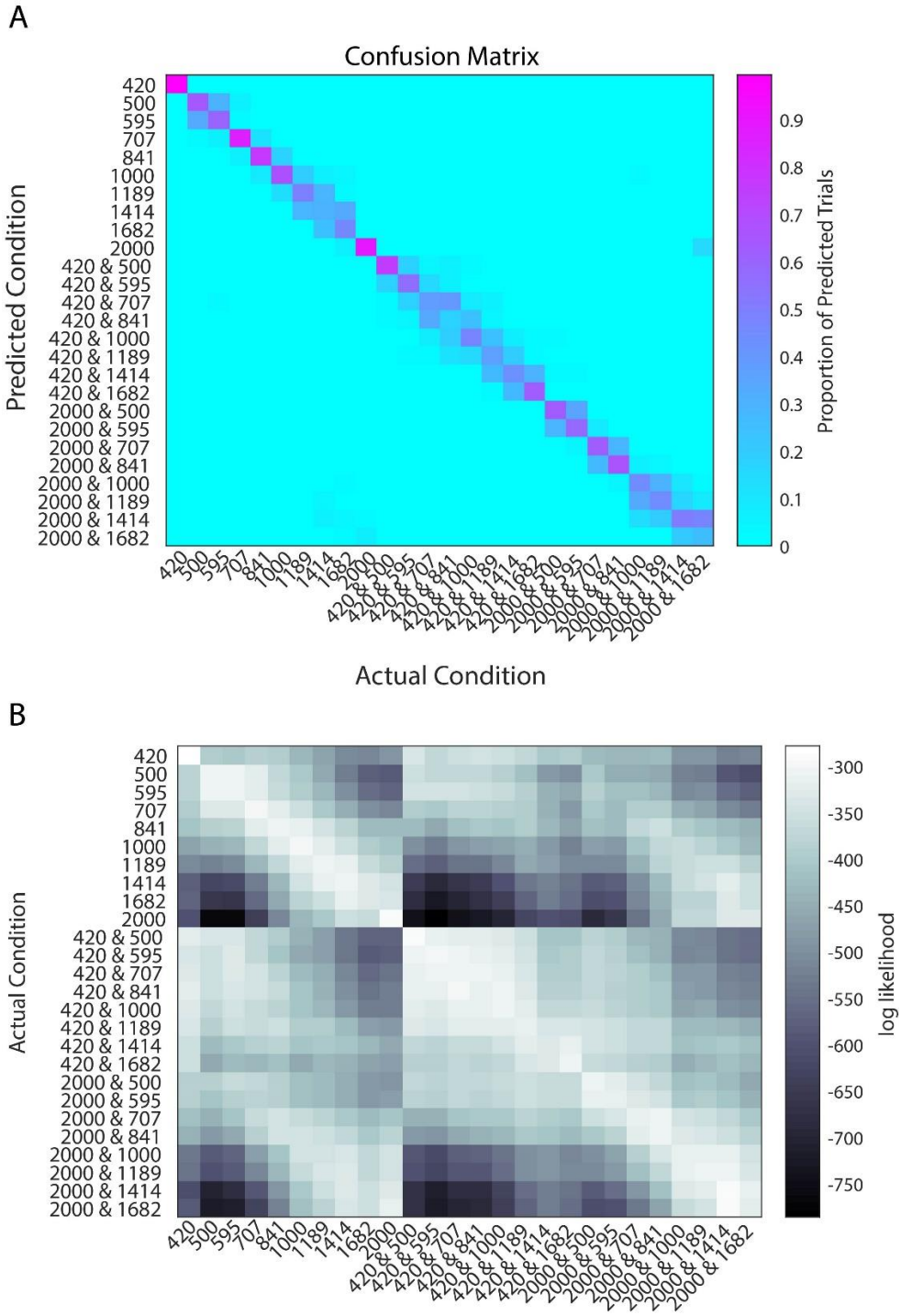

Figure 9 Confusion and mean-likelihood matrices for the decoder. (**A**) The confusion matrix for the decoder. Each column corresponds to the actual condition and each row corresponds to the predicted condition. The color of each square is the proportion of predicted trials (out of 1000) that fell into that bin. The diagonal corresponds to correct predictions. (**B**) A matrix of mean log-likelihood across conditions for the decoder. Each row corresponds to the likelihood function for that condition while each column corresponds to the mean log-likelihood for that condition across all actual conditions (e.g. leftmost column corresponds to the mean log-likelihood for 420 Hz across all conditions).

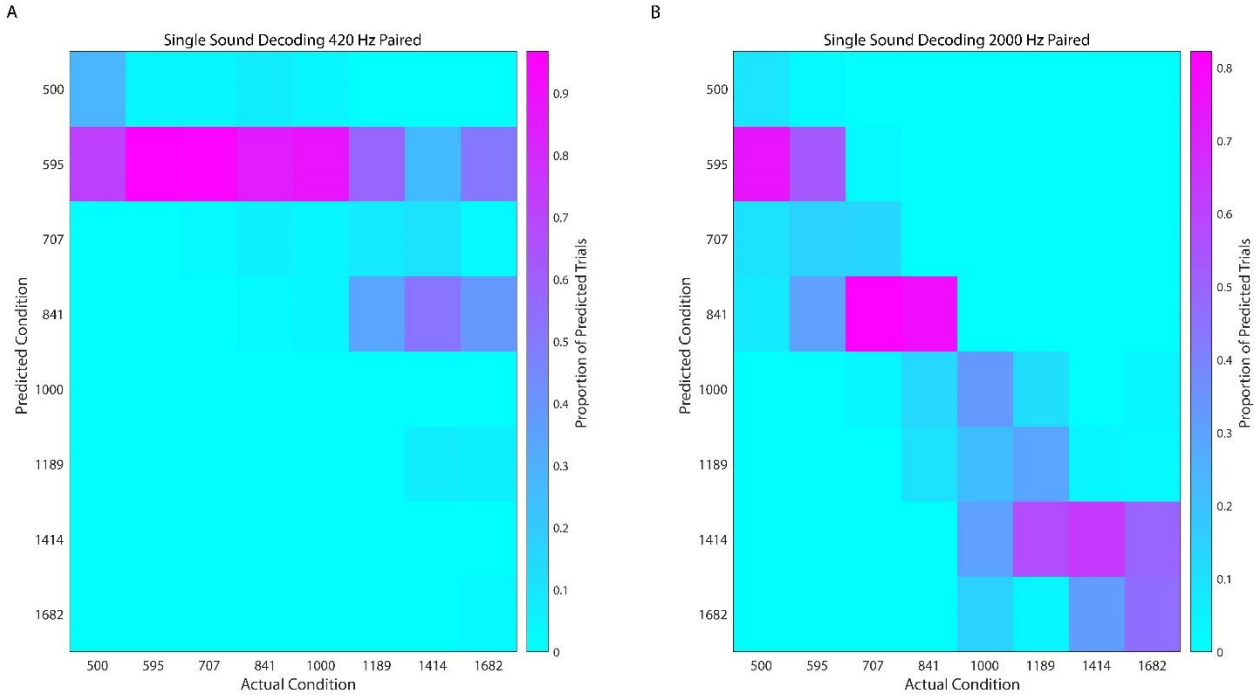

Figure 10 Decoding the paired condition with the single-sound weights impairs performance. (**A**) The confusion matrix for decoding 420 Hz paired conditions using the weights learned from single-sound trials. Each column corresponds to the actual condition and each row corresponds to the predictable condition. The color of each square is the proportion of predicted trials (out of 1000) that fell into that bin. The diagonal corresponds to correct predictions. (**B**) The same as (**A**) but for decoding the 2000 Hz paired conditions.

References:

Aitkin, L. M., & Phillips, S. C. (1984). Is the inferior colliculus and obligatory relay in the cat auditory system? *Neuroscience Letters*, *44*(3), 259–264. https://doi.org/https://doi.org/10.1016/0304-3940(84)90032-6

Best, V., van Schaik, A., & Carlile, S. (2003). Separation of concurrent broadband sound sources by human listeners. *The Journal of the Acoustical Society of America*, *115*(1), 324–336. https://doi.org/10.1121/1.1632484

Bulkin, D. a, & Groh, J. M. (2011). Systematic mapping of the monkey inferior colliculus reveals enhanced low frequency sound representation. *Journal of Neurophysiology*, *105*(February 2011), 1785–1797. https://doi.org/10.1152/jn.00857.2010

Bulkin, D. A., & Groh, J. M. (2012). Distribution of eye position information in the monkey inferior colliculus. *Journal of Neurophysiology*, *107*(3), 785–795. https://doi.org/10.1152/jn.00662.2011

Bulkin, D. a., & Groh, J. M. (2012). Distribution of visual and saccade related information in the monkey inferior colliculus. *Frontiers in Neural Circuits*, *6*(September), 1–12. https://doi.org/10.3389/fncir.2012.00061

Capuano, U., & McIlwain, J. T. (1981). Reciprocity of receptive field images and point images in the superior colliculus of the cat. *Journal of Comparative Neurology*, *196*(1), 13–23. https://doi.org/10.1002/cne.901960103

Caruso, V. C., Mohl, J. T., Glynn, C., Lee, J., Willett, S. M., Zaman, A., Ebihara, A. F., Estrada, R., Freiwald, W. A., Tokdar, S. T., & Groh, J. M. (2018). Single neurons may encode simultaneous stimuli by switching between activity patterns. *Nature Communications*, *9*(1), 2715. https://doi.org/10.1038/s41467-018-05121-8

Caruso, V. C., Pages, D. S., Sommer, M. A., & Groh, J. M. (2017). Beyond the labeled line: variation in visual reference frames from intraparietal cortex to frontal eye fields and the superior colliculus. *Journal of Neurophysiology*, *119*(4), 1411–1421. https://doi.org/10.1152/jn.00584.2017

Chettih, S. N., & Harvey, C. D. (2019). Single-neuron perturbations reveal feature-specific competition in V1. *Nature*, *567*(7748), 334–340. https://doi.org/10.1038/s41586-019-0997-6

Connell, M. N., Barczak, A., Schroeder, C. E., & Lakatos, P. (2014). Layer Specific Sharpening of Frequency Tuning by Selective Attention in Primary Auditory Cortex. *The Journal of Neuroscience*, *34*(49), 16496 LP – 16508. https://doi.org/10.1523/JNEUROSCI.2055-14.2014

Day, M. L., & Delgutte, B. (2013). Decoding Sound Source Location and Separation Using Neural Population Activity Patterns. *The Journal of Neuroscience*, *33*(40), 15837 LP – 15847. https://doi.org/10.1523/JNEUROSCI.2034-13.2013

Day, M. L., Koka, K., & Delgutte, B. (2012). Neural encoding of sound source location in the presence of a concurrent, spatially separated source. *Journal of Neurophysiology*, *108*(9), 2612–2628. https://doi.org/10.1152/jn.00303.2012

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2005). Experience-Dependent Sharpening of Visual Shape Selectivity in Inferior Temporal Cortex. *Cerebral Cortex*, *16*(11), 1631–1644. https://doi.org/10.1093/cercor/bhj100

Groh, J. M., Kelly, K. A., & Underhill, A. M. (2003). A Monotonic Code for Sound Azimuth in Primate Inferior Colliculus. *Journal of Cognitive Neuroscience*, *15*(8), 1217–1231.

Groh, J. M., Trause, A. S., Underhill, A. M., Clark, K. R., Inati, S., & Hall, M. (2001). Eye Position Influences Auditory Responses in Primate Inferior Colliculus. *Neuron*, *29*(2), 509–518.

Hartline, H. K., & Ratliff, F. (1957). INHIBITORY INTERACTION OF RECEPTOR UNITS IN THE EYE OF LIMULUS . *Journal of General Physiology*, *40*(3), 357–376. https://doi.org/10.1085/jgp.40.3.357

Henry, C. A., & Kohn, A. (2020). Spatial contextual effects in primary visual cortex limit feature representation under crowding. *Nature Communications*, *11*(1), 1687. https://doi.org/10.1038/s41467-020-15386-7

Imig, T. J., Irons, W. A., & Samson, F. R. (1990). Single-unit selectivity to azimuthal direction and sound pressure level of noise bursts in cat high-frequency primary auditory cortex. *Journal of Neurophysiology*, *63*(6), 1448–1466. https://doi.org/10.1152/jn.1990.63.6.1448

Jazayeri, M., & Movshon, J. A. (2006). Optimal representation of sensory information by neural populations. *Nature Neuroscience*, *9*(5), 690–696. https://doi.org/10.1038/nn1691

Jenkins, W. M., & Masterton, R. B. (1982). Sound localization: effects of unilateral lesions in central auditory system. *Journal of Neurophysiology*, *47*(6), 987–1016. https://doi.org/10.1152/jn.1982.47.6.987

Judge, S. J., Richmond, B. J., & Chu, F. C. (1980). Implantation of magnetic search coils for measurement of eye position: An improved method. *Vision Research*, *20*(6), 535–538. https://doi.org/https://doi.org/10.1016/0042-6989(80)90128-5

Kelly, J. B., & Kavanagh, G. L. (1994). Sound localization after unilateral lesions of inferior colliculus in the ferret (Mustela putorius). *Journal of Neurophysiology*, *71*(3), 1078–1087. https://doi.org/10.1152/jn.1994.71.3.1078

Lee, C.-C., & Middlebrooks, J. C. (2011). Auditory cortex spatial sensitivity sharpens during task performance. *Nature Neuroscience*, *14*(1), 108–114. https://doi.org/10.1038/nn.2713

Mickey, B. J., & Middlebrooks, J. C. (2003). Representation of Auditory Space by Cortical Neurons in Awake Cats. *The Journal of Neuroscience*, *23*(25), 8649 LP – 8663. https://doi.org/10.1523/JNEUROSCI.23-25-08649.2003

Middlebrooks, J. C., Clock, A. E., Xu, L., & Green, D. M. (1994). A panoramic code for sound location by cortical neurons. *Science*, *264*(5160), 842 LP – 844. https://doi.org/10.1126/science.8171339

Middlebrooks, J. C., & Pettigrew, J. D. (1981). Functional classes of neurons in primary auditory cortex of the cat distinguished by sensitivity to sound location. *The Journal of Neuroscience*, *1*(1), 107 LP – 120. https://doi.org/10.1523/JNEUROSCI.01-01-00107.1981

Pages, D. S., Ross, D. A., Puñal, V. M., Agashe, S., Dweck, I., Mueller, J., Grill, W. M., Wilson, B. S., & Groh, J. M. (2016). Effects of Electrical Stimulation in the Inferior Colliculus on Frequency Discrimination by Rhesus Monkeys and Implications for the Auditory Midbrain Implant. *The Journal of Neuroscience*, *36*(18), 5071 LP – 5083. https://doi.org/10.1523/JNEUROSCI.3540-15.2016

Perrott, D. R. (1984a). Concurrent minimum audible angle: A re-examination of the concept of auditory spatial acuity. *The Journal of the Acoustical Society of America*, *75*(4), 1201–1206. https://doi.org/10.1121/1.390771

Perrott, D. R. (1984b). Discrimination of the spatial distribution of concurrently active sound sources: Some experiments with stereophonic arrays. *The Journal of the Acoustical Society of America*, *76*(6), 1704–1712. https://doi.org/10.1121/1.391617

Porter, K. K., Metzger, R. R., & Groh, J. M. (2007). Visual- and saccade-related signals in the primate inferior colliculus. *Proceedings of the National Academy of Sciences*, *104*(45), 17855–17860. http://www.pnas.org/content/104/45/17855.full

Porter, K. K., Metzger, R. R., Groh, J. M., Kelly, K., Metzger, R. R., & Repre-, J. M. G. (2006). Representation of Eye Position in Primate Inferior Colliculus. *Journal of Neurophysiology*, *95*(3), 1826–1842. https://doi.org/10.1152/jn.00857.2005.

Robinson, D. A. (1963). A Method of Measuring Eye Movemnent Using a Scieral Search Coil in a Magnetic Field. *IEEE Transactions on Bio-Medical Electronics*, *10*(4), 137–145. https://doi.org/10.1109/TBMEL.1963.4322822

Ryan, A., & Miller, J. (1978). Single Unit Responses in the Inferior Colliculus of the Awake and Performing Rhesus Monkey. *Experimental Brain Research*, *32*(3), 389–407. http://link.springer.com/article/10.1007/BF00238710?LI=true

Versnel, H., Zwiers, M. P., & van Opstal, A. J. (2009). Spectrotemporal Response Properties of Inferior Colliculus Neurons in Alert Monkey. *The Journal of Neuroscience*, *29*(31), 9725 LP – 9739. https://doi.org/10.1523/JNEUROSCI.5459-08.2009

Wallisch, P. (2014). *Chapter 21 - Neural Decoding I: Discrete Variables* (P. Wallisch, M. E. Lusignan, M. D. Benayoun, T. I. Baker, A. S. Dickey, & N. G. B. T.-M. for N. (Second E. Hatsopoulos, Eds.; pp. 329–336). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-383836-0.00021-7

Weinberger, N. M. (1995). Dynamic Regulation of Receptive Fields and Maps in the Adult Sensory Cortex. *Annual Review of Neuroscience*, *18*(1), 129–158. https://doi.org/10.1146/annurev.ne.18.030195.001021

Winer, J. A., & Schreiner, C. E. (Eds.). (2005). *The Inferior Colliculus* (1st ed.). Springer-Verlag New York.

Zhong, X., & Yost, W. A. (2017). How many images are in an auditory scene? *The Journal of the Acoustical Society of America*, *141*(4), 2882–2892. https://doi.org/10.1121/1.4981118

Zwiers, M. P., Versnel, H., & Opstal, A. J. van. (2004). Involvement of Monkey Inferior Colliculus in Spatial Hearing. *Journal of Neurosciece*, *24*(17), 4145–4156. https://doi.org/10.1523/JNEUROSCI.0199-04.2004