Interareal and interlaminar differences in sound envelope encoding in core and parabelt auditory cortex

Chase A. Mackey[1], and Yoshinao Kajikawa[1,2]

1. Nathan Kline Institute for Psychiatric Research, Center for Biomedical Imaging and Neuromodulation, Orangeburg, NY 10962

2. New York University School of Medicine, Department of Psychiatry, New York, NY 10016

Corresponding authors: chase.mackey@nki.rfmh.org; yoshinao.kajikawa@nki.rfmh.org

Address: 140 Old Orangeburg Rd., Orangeburg NY, 10962

1-865-684-5345

The authors declare no conflicts of interest

Acknowledgements

Funding sources: R01DC015780, R01DC012947 and R01DC019979.

The authors thank Dr. Monica N. O'Connell for helpful comments on an earlier version of the manuscript.

Abstract

Amplitude-modulation (AM) plays an important role in the perception of complex sounds, and transformations in AM encoding may underlie aspects of complex sound perception. Previous studies have described a hierarchical progression across the auditory pathway, characterized by a decrease in the temporal precision of AM encoding. In human and nonhuman primates (NHP), the left hemisphere exhibits enhanced temporal encoding relative to the right hemisphere. The NHP model provides an opportunity to understand what circuit mechanisms generate these transformations in encoding by characterizing AM encoding in different intracortical circuits, and across the cortical hierarchy. To address this, here we report the

encoding of AM signals as a function of cortical layer and hemisphere in NHP core and parabelt auditory cortex (AC). We recorded electrophysiological activity using linear array multielectrodes positioned across cortical layers while AM noise and click trains were presented to awake NHPs. Core AC typically encoded all AM frequencies (1.6-200 Hz) with high (> 90%) classification accuracy, while sites in the parabelt encoded a subset of lower (~1.6-25 Hz) frequencies. Across both areas, the granular and infragranular layers displayed enhanced AM encoding relative to the supragranular layers. Both areas displayed enhanced AM encoding in the left hemisphere, restricted to the supragranular layers. These results represent the first analysis of AM encoding in the parabelt, indicating that significant temporal encoding of AM is still present in tertiary auditory cortex, and the layer-specific hemispheric differences suggest a potential supragranular layer origin of previously documented left-hemisphere dominance in temporal encoding.

Significance Statement (120 words)

The loudness of natural sounds fluctuate rapidly – a feature known as amplitude modulation (AM). The rate of AM allows us to differentiate the pitch of different voices, discriminate speech from music, and supports other foundational aspects of perception. This study investigated AM encoding in specific intracortical circuits across levels of auditory cortex. We demonstrate here that each level of the hierarchy displays enhanced encoding in the deep cortical layers relative to the superficial layers. We also observed that the left hemisphere dominance in AM encoding, evident in humans and nonhuman primates, was localized mainly to the superficial layers. This provides new understanding of how the third level of the nonhuman primate cortical hierarchy, the parabelt, encodes complex stimuli.

Introduction

Sound envelope fluctuations are a key sound feature that drives the perception of complex stimuli such as species-specific vocalizations and music (Ding et al., 2017; Henry et al., 2017; Rosen, 1992). Many previous studies have gained insight into how sound envelope encoding differs across the auditory pathway by recording neuronal responses to click trains and sinusoidally amplitude modulated (AM) tones and noise in nonhuman animals (Henry et al.,

2016; Joris and Yin, 1992; Joris et al., 2004; Langner and Schreiner, 1988; Mackey et al., 2024a; Nelson and Carney, 2007; Rhode et al., 2010; Wang et al., 2008). These studies document a hierarchical progression across the pathway from a spike-timing based code, where spikes are aligned to individual cycles of the sound envelope (often measured using vector strength), to an average rate-code which is more prevalent at and above the inferior colliculus. By using measures of discrimination and classification performance derived from neuronal responses, studies of sound envelope encoding have explicitly addressed hypotheses about how information in different parts of the auditory pathway could contribute to sound envelope perception (Downer et al., 2017, 2021; Henry et al., 2016; Johnson et al., 2012; Mackey et al., 2024a; Nelson and Carney, 2007; Sayles et al., 2013; Yao and Sanes, 2021), with two of the critical findings being that 1) small populations of neurons in the brainstem, midbrain, and cortex encode information sufficient to account for detection and discrimination of amplitude modulation depth and frequency (Downer et al., 2021; Johnson et al., 2012; Mackey et al., 2024a; Nelson and Carney, 2007; Penikis and Sanes, 2023; Sayles et al., 2013; Shi et al., 2024), and 2) there are notable differences in encoding schemes in primary vs. secondary auditory cortex (Downer et al., 2017; Johnson et al., 2024). The use of small populations (empirical or simulated) has been a key aspect of these studies. In regions such as A1, single-neuron measures have suggested temporal information not to be present in sufficient quantities to account for the perception of AM based on, for instance, the population's average vector strength alone (Johnson et al., 2012). However, subsequent studies utilizing decoding methods, which are agnostic to the presence of a rate vs. temporal code, have revealed that temporal information is preserved in the population (Downer et al., 2021). These findings informed the present study's use of decoding methods on population-level activity measures.

In the study of AM encoding, the consideration of species is important for many reasons, and we focus on three here: 1) Primates (human and nonhuman) display greater perceptual sensitivity to AM than rodents and parakeets (Mackey et al., 2024a; Moody, 1994), 2) nonhuman primates (NHP) display more temporally precise AM encoding than rodents (Hoglen et al., 2018), and 3) primate auditory cortex exhibits a specific three-stage hierarchical anatomical organization, core-belt-parabelt, that provides the closest approximation of human auditory cortex available in animal models (Hackett et al., 2001, 2014). Previous studies of human auditory cortex report hierarchical progressions in speech encoding (Hamilton et al., 2021; Zion

Golumbic et al., 2013) (e.g. Heschl's gyrus vs. Superior Temporal Gyrus) and interhemispheric differences in the encoding of sound envelope (Boemio et al., 2005). Namely, the left hemisphere preferentially encodes higher frequency envelope fluctuations, while the right hemisphere preferentially encodes lower frequency envelope fluctuations; and Heschl's gyrus encodes spectrotemporal speech features more precisely than the Superior Temporal Gyrus, which encodes higher level (e.g. semantic) information. The NHP model provides an opportunity to understand what circuits generate these transformations in encoding by characterizing AM encoding in different cortical circuits within area, and across the cortical hierarchy. For instance, studies recording simultaneously across cortical layers in auditory cortex have provided insight into the function of different cortical circuits in encoding temporal sound patterns involved in streaming (Banno et al., 2023; Fishman et al., 2001, 2004; Lakatos et al., 2013), intermodal (audiovisual) attention (Lakatos et al., 2009; O'Connell et al., 2014, 2015; Schroeder and Lakatos, 2009), and pitch perception of click trains (Steinschneider et al., 1998). But there is a dearth of NHP studies of AM encoding in the superior temporal gyrus, containing the higher-order parabelt region of auditory cortex; moreover, there are no published studies of AM encoding as a function of physiologically identified cortical layers in NHPs. Thus it is likely that studies of AM encoding with laminar resolution across the cortical hierarchy could yield insights into the circuits utilized for AM encoding, and temporal sound pattern encoding more generally, in humans. To address this, here we report the encoding of AM signals as a function of cortical layer and hemisphere in NHP core and parabelt auditory cortex.

Materials and Methods

All experimental procedures were approved by the Institutional Animal Care and Use Committee of the Nathan Kline Institute (NKI).

Animals. Four nonhuman primates (*Macaca mulatta*, two female) H, Q, S, and W were implanted with headposts and recording chambers. Recordings from core auditory cortex were done in chambers positioned to aim penetrations perpendicular to the supra temporal plane (Male W in the right hemisphere, female H in the left hemisphere, male S in the left hemisphere). NHPs were also implanted with superior temporal gyrus chambers (female Q only on the left, female H only on the left, male W bilaterally). Surgical details for accessing core and parabelt auditory

cortex have been described in our previous publications (Kajikawa et al. 2017; Kajikawa et al. 2015). Specifically, based on presurgical MRI, the chambers were positioned to aim penetrations perpendicular to the lower bank of the lateral sulcus, in the case of core AC. For accessing the parabelt, presurgical MRI was used to target the superior temporal gyrus (STG) via multiple landmarks, which aided in positioning optimally over the STG: the superior temporal sulcus (~20 mm above the ear canal and ~10 mm above the zygomatic arch over AP 0-20 mm), the lateral sulcus, and the zygomatic arch (described extensively in Kajikawa et al. 2015).

Stimulus parameters. While NHPs sat awake, AM noise or click trains were presented with randomized modulation frequency (AM noise) or presentation rate (clicks). For brevity, the term modulation frequency will be used throughout the manuscript to describe the presentation rate of the click trains. Modulation frequencies utilized in this study were 1.6, 3.1, 6.3, 8.8, 12.5, 17.5, 25, 36, 50, 71, 100, and 200 Hz. Signals were 640 ms long, 70 dB PL, 100% modulated, and were ramped on/off using a linear ramp that was 5 ms in duration.

Electrophysiological recordings. Data collection was conducted as described previously (Kajikawa et al., 2017; Mackey et al., 2024b). U-probes (Plexon) with a linear array of 23 channels (0.3– 0.5 MΩ at 1.0 kHz) at intervals of 100 or 200 micrometers between channels were used. Signals from all channels were amplified 10x, followed by 500x amplification before multiplexing and filtering. Each channel recorded field potentials (FP) (0.1–500 Hz) and MUA (500–5000 Hz using zero phase shift digital filtering, 48 dB/octave and rectified) simultaneously. When compared directly, the envelope MUA utilized here provides similar responses, and estimates of tuning, to thresholded multiunit spiking activity (Kayser, Petkov, Logothetis 2007; Super and Roelfsema, 2005; Shaheen and Liberman 2018), which supports the idea that our results are comparable to previous multiunit studies of click train and amplitude-modulation encoding. Both MUA and FP were sampled at 2 kHz and stored for offline analyses.

Data Analysis

Current source density (CSD) was calculated as the second spatial, i.e. across-channel, derivative of the FP for each time point of the laminar CSD profile (Freeman and Nicholson, 1975; Kajikawa and Schroeder, 2011; Mitzdorf, 1985).

Phase-locking value. The phase locking value of CSD and MUA signals were computed at each modulation frequency (1.6-200 Hz) using complex demodulation, equivalent to a Discrete Fourier Transform coefficient at a given modulation frequency (Lu et al., 2022; Tallon-Baudry et al., 1996). For each trial *k*, a signal (duration T), windowed by Hann function $w(t) = \frac{1}{2}\left(1 - cos\left(\frac{2\pi t}{T}\right)\right)$, was convolved with a narrowband kernel, $f(t) = e^{-2\pi i f t}$ as $a_k = \frac{1}{T}\int_{t=0}^{T} x_k(t) \cdot w(t) \cdot f(t)\, dt$. The resultant coefficient, a_k, is a complex, and its argument, $\varphi_k = \arg(a_k)$, depends on the timing of MUA fluctuation relative to the cycle of the band. The phase-locking value was derived from all *N* trials as $PLV = \left|\frac{1}{N}\sum_k^N e^{i\varphi_k}\right|$. PLV is mathematically the same vector-length statistic used for vector strength of spike trains. In classical vector strength, the phases are taken at each spike time relative to the stimulus cycle. When computing PLV for continuous signals, the phase of the narrowband signal is calculated per trial by complex demodulation at the modulation frequency, yielding phases across trials. Both measures quantify the concentration of phase angles on the unit circle for each cycle of the stimulus from 0-1, with 1 being perfect synchrony, but PLV is better suited for our continuous CSD and MUA signals. And because we aggregate many cycles within each trial before taking the across-trial resultant, we avoid inflating sample size by counting cycles as independent; the degrees of freedom are the trials. This makes our PLV comparable to spike-based vector strength while remaining appropriate for continuous signals.

$$\text{[Equation 1]} \quad f(t) = e^{-2\pi i f t}$$

$$\text{[Equation 2]} \quad w(t) = \frac{1}{2}\left(1 - cos\left(\frac{2\pi t}{T}\right)\right)$$

$$\text{[Equation 3]} \quad a_k = \frac{1}{T}\int_{t=0}^{T} x_k(t) \cdot w(t) \cdot f(t)\, dt$$

$$\text{[Equation 4]} \quad PLV = \left|\frac{1}{N}\sum_k^N e^{i\varphi_k}\right|$$

To statistically evaluate phase-locking, we constructed a random-phase null distribution (*n* = 1,000) that preserved the trial's demodulated amplitude. We computed the observed amplitude-

weighted vector, and estimated a *p-value* as the proportion of surrogates with a weighted vector greater than our observed weighted vector. We corrected for multiple comparisons across electrode channels using the Benjamini Hochberg method (Benjamini and Hochberg, 1995). For reference, we also computed corrected *p*-values from the more conventional Rayleigh's test, which provided qualitatively similar significance across the dataset.

Decoding Analysis and Mutual Information Estimation. To assess the decoding of stimulus identity from multichannel neural activity, we employed a multivariate classification and mutual information (MI) analysis pipeline based on linear discriminant analysis ("fitdiscr" in MATLAB; Statistics and Machine Learning Toolbox). This method fits a multivariate Gaussian model to each stimulus condition. The algorithm estimates the class-conditional means and a pooled covariance matrix, and constructs linear discriminant functions for each class. Each trial is then classified by assigning it to the class with the highest posterior probability, under the assumption of normally distributed features with equal covariance structure. This approach is equivalent to finding the linear projection that maximizes separation between class means while minimizing within-class variance. The analysis was conducted separately for each recording session and signal type (MUA or CSD) using custom-written MATLAB scripts. For each dataset, electrophysiological recordings were imported and epoched from -50 to 700 ms relative to stimulus onset and baseline corrected (-50 to 0 ms). Data used for decoding consisted of the time-domain signal from each trial and channel. Only channels estimated to be within a cortical layer compartment (Supragranular, Granular, or Infragranular) were included for decoding. For each channel and decoding window, data were segmented into non-overlapping bins; unless otherwise specified, the decoding window was the entire baseline corrected 700 ms window after the stimulus onset, and we qualitatively inspected the pre-stimulus interval to ensure there was no difference in chance-level decoding; additionally, chance performance in the pre-stimulus interval is included in the figures. Linear discriminant analysis (LDA) classifiers were trained using fourfold cross-validation to predict stimulus condition labels from trial-wise data. A confusion matrix was computed for each fold, and the resulting matrices were averaged across folds to yield a single empirical confusion matrix per electrode channel, time window, and permutation. This calculation was repeated 1000 times to compute the 95% confidence intervals (CI) around "True MI" depicted in the results section. To compute chance performance,

decoding analyses were repeated 1000 times with trial labels randomly shuffled. For each permutation and each decoding window, mutual information was computed from the resulting confusion matrix using the following formulation:

$$MI = \Sigma \{i, j\} P(i, j) \cdot log_2 [P(i, j) / (P(i) \cdot P(j))]$$

where P(i,j) is the joint probability of predicted and true labels, and P(i), P(j) are the marginal probabilities. MI was computed for both true and shuffled label conditions for all decoding time-windows, channels, and permutations. The resulting distributions were used to compute mean MI and 95% confidence intervals under both empirical and null conditions.

Statistical analysis. In all cases, criterion for significance was 0.05, and results requiring correction for multiple comparisons (e.g. the phase-locking values) were corrected via the Benjamini-Hochberg correction (see section on phase locking value). Linear mixed-effects model (LME) analysis was conducted using “fitlme” in MATLAB. Only channels where the Mutual Information lower CI exceeded the shuffled (chance) Mutual Information upper CI were retained for statistical analysis. To account for the nested structure of the data (multiple channels per recording), we modeled MI as a function of three fixed effects: cortical layer (supragranular, granular, infragranular), area (core vs. parabelt), and hemisphere (left vs. right), including their two-way interactions. Individual channels were included as a random intercept term. Significant effects of layer, area, and hemisphere indicate a significant difference from the reference groups, which were the supragranular layers, core AC, and the left hemisphere. In many nonhuman primate (NHP) neurophysiology experiments, including the present study, neural data are collected from many electrode contacts distributed across cortical layers and penetrations, with recordings repeated over numerous sessions in the same animal. In this context, the relevant statistical question is whether a given factor—such as cortical layer, area, or hemisphere—has a systematic effect on the measured neural variable (e.g., mutual information) across the population of recorded channels. Accordingly, the degrees of freedom in our mixed-effects model reflect the number of channels positioned in grey matter (confirmed using the CSD), recorded across sessions, often totaling approximately 700-1000 channels across the dataset. Treating individual channels as measuring different local neuronal ensembles is supported by our

electrode spacing (100-200 micrometers), and the observation that neighboring channels displayed markedly different response characteristics and decoding accuracy.

Results

Field potential and multiunit spiking activity were measured using linear array "laminar" multielectrode probes (see Methods) in response to amplitude-modulated noise, and click trains of different frequencies (i.e. repetition rates) in core auditory cortex (AM noise, n = 42; Clicks, n = 63) and parabelt auditory cortex (AM noise, n = 88; Clicks, n = 84) across four rhesus macaques (*Macaca mulatta*). From the field potential, current source density was calculated as the second derivative across electrode channels to mitigate volume conduction, and to estimate local transmembrane currents in specific cortical layer compartments with specific spatial resolution given by the interelectrode distance (100-200 micrometers). Using the frequently encountered activation pattern, characterized by an early onset, sharp L4 sink (negative CSD) co-localized with multiunit activity, followed by L2/3 and L5/6 CSD activation, individual electrode channels were assigned to the supragranular (L2/3), granular (L4), and infragranular (L5/6) compartments (Kajikawa et al., 2015; Lakatos et al., 2005; Steinschneider et al., 1992). Figure 1 shows trial-averaged CSD and MUA responses to AM noise of different modulation frequencies. Figure 1A depicts the response to 3.1 Hz AM noise in a representative core auditory cortex site, exhibiting an early (~15 ms), sharp CSD sink (negativity, red in the color map) characteristic of the thalamorecipient granular layer. The multiunit activity derived from the same signal exhibits an early peak around in the same electrodes. The supragranular and infragranular CSD signals exhibited longer latencies, as expected, which resulted in relatively steady-state (as opposed to phase-locked) responses to the higher modulation frequency signals (Figure 1B-D). Figure 1E depicts an example parabelt site, where laminar response profiles displayed more diversity than core AC, which we have previously described (Kajikawa et al., 2015; Mackey et al., 2024b). A notable difference between the site in Figure 1G-H and Figure 1C-D is that the parabelt multiunit activity displays less phase-locking, though there is noticeable rhythmicity to the late MUA suppression that could contribute to discrimination between modulation frequencies (Figure 1G). Summary phase-locking values (see Methods) are depicted in Figure 1I-L. Across layers, similar

phase-locking was observed, on average, in both the CSD and MUA signal. The parabelt differed noticeably from core AC in that significant phase-locking was almost never observed above 10 Hz. We restricted calculation of the mean to frequencies at which more than 10 neuronal ensembles (i.e. channels) displayed significant PLV (Figure 1I-L). While neuronal responses can be summarized using average amplitude or phase-locking measures, the underlying waveform is complex, particularly for continuous signals such as the CSD and envelope MUA depicted here. This motivated our use of decoding methods to assess information present in the response waveforms across our dataset.

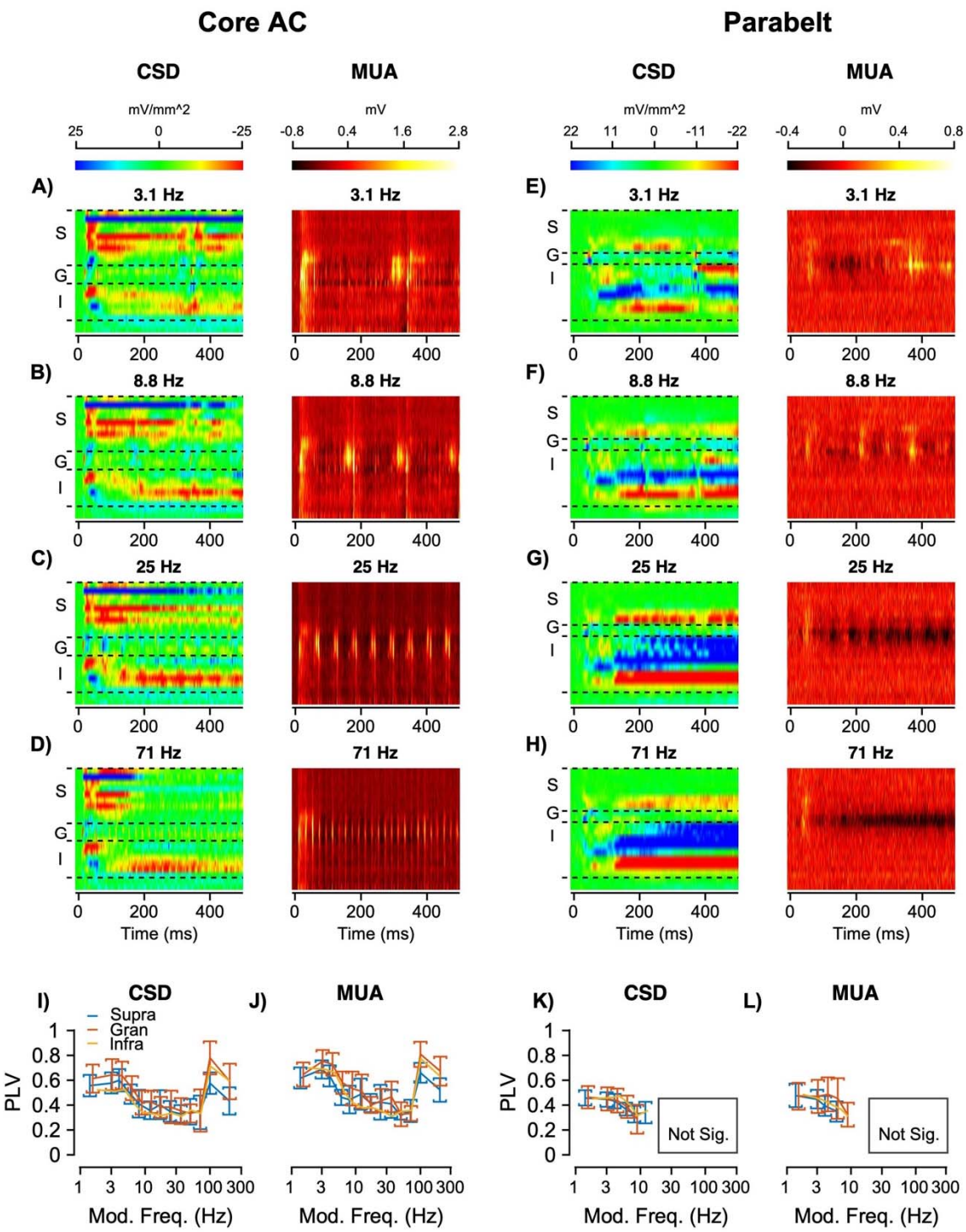


Figure 1 – Representative laminar activity profiles in response to sinusoidally amplitude modulated noise of different frequencies, and phase-locking values across the data set. The left column (**A-D**) show trial-averaged responses (n = 50) from a representative recording site in core auditory cortex to sinusoidally amplitude modulated (AM) noise of different frequencies.

The right column shows trial-averaged responses (n = 50) from a representative site in the parabelt (**E-H**) to SAM of different modulation frequencies. We computed a phase-locking value for all recordings, which are displayed for each layer and signal type (CSD and MUA) for core AC (**I** and **J**; mean and 95% CIs) and the parabelt (**K** and **L**; mean and 95% CIs). PLV was slightly jittered along the x-axis to enable visibility of the different data traces; in all cases, the modulation frequencies were the same (see Methods).

Classifying modulation frequency based on current source density and multiunit activity

To quantify how CSD and MUA encode information about AM stimuli, we utilized linear discriminant analysis (LDA; see Methods), and calculated Mutual Information (MI). MI is useful because it quantifies decoding accuracy across conditions, and the systematicity of decoding errors, as opposed to averaging accuracy across conditions. Figure 2 depicts the representative core and parabelt sites from Figure 1, with the LDA-derived confusion matrix for a channel in the granular layer for both sites (Figure 2A; Figure 2E), with MI values (right-hand panel) for all channels across layers (the y-axis). The color maps and MI plots were scaled the same across conditions and to emphasize the differences between core AC and parabelt. Noticeably, core AC exhibited nearly perfect decoding accuracy across conditions (Modulation frequencies). For example, the color map in Figure 1A indicates that most conditions were predicted correctly based on the granular channel CSD response in 100% of trials; and the granular channel MUA decoding was similar (Figure 1B). Across channels, it can be seen that MI, which provides an information theoretic metric for stimulus information across different stimulus conditions (i.e. across the entire confusion matrix). 12 modulation frequencies were utilized, yielding a maximum possible MI value of about 3.5 bits ($\log_2(12) = 3.5$), which further emphasizes how accurate the site in Figure 2A-D is. Interestingly, though the overall decoding accuracy was similar, the laminar profile of MI differed between AM noise and clicks: supragranular MUA, for example, exhibited MI values not significantly different than the infragranular channels in response to clicks, while AM noise did not elicit such a strong response in the supragranular MUA.

Relative to core AC, the parabelt demonstrated reduced decoding accuracy and MI, likely due to the parabelt's higher position in the processing hierarchy (Figure 2E-H). Though the granular channels exhibited the highest MI for this example, MI distributions were broader, i.e.

there was more variability in parabelt dataset than in core AC; thus, though the extragranular layers encoded very little information in this example, across the dataset they often displayed significant, above chance MI.

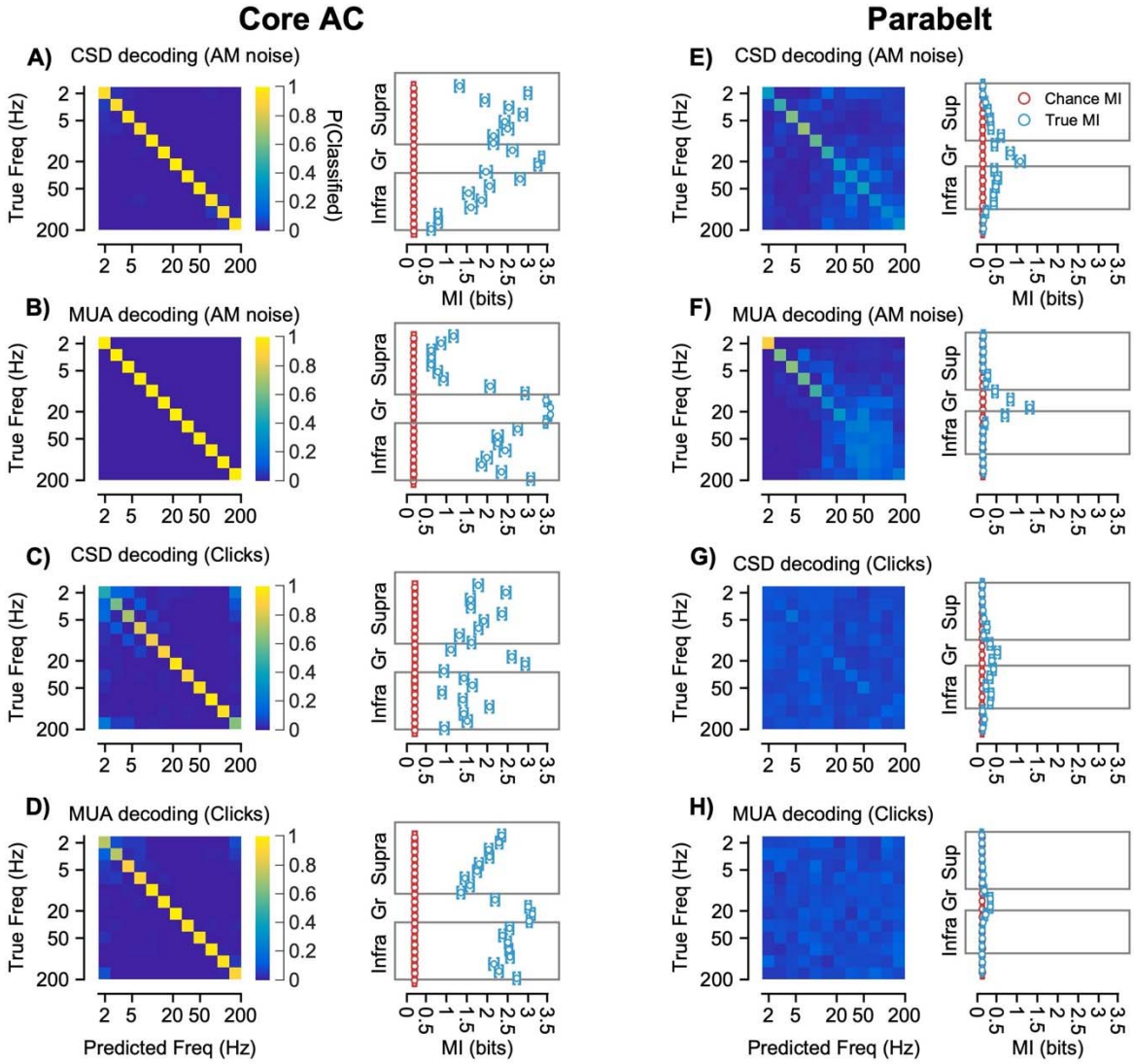


Figure 2 – Representative example decoding results from AM noise and click-train evoked responses from the recording sites in Figure 1. The left panels depict the confusion matrices which quantify the decoding accuracy (0-1) in each condition. The predicted condition is indicated on the x-axis, and the true condition indicated on the y-axis. The color axis represents the proportion of classifications. The mutual information (MI) values are plotted on the x-axis of each panel on the right. Observed, or "true" MI was computed as described in the Methods, and Chance MI was calculated after shuffling the condition labels. For the example in core auditory cortex, the confusion matrices for a channel in the granular layer (layer 4) are shown in response

to AM noise (**A** and **B**), and click trains (**C** and **D**), with corresponding mutual information values of all channels, which were calculated on the confusion matrices. Panels **E-H** depict the same analysis for the example in the parabelt.

In the MI plots, channels are labeled as being in the supragranular (Supra), granular (Gr), or infragranular (Infra) layer compartments. Color maps are scaled 0-1, and mutual information plots are scaled 0-3.5 bits (the maximum possible MI with 12 stimulus conditions). Error bars represent 95% confidence intervals calculated across 1,000 permutations.

Interhemispheric, interareal, and interlaminar differences in AM noise encoding

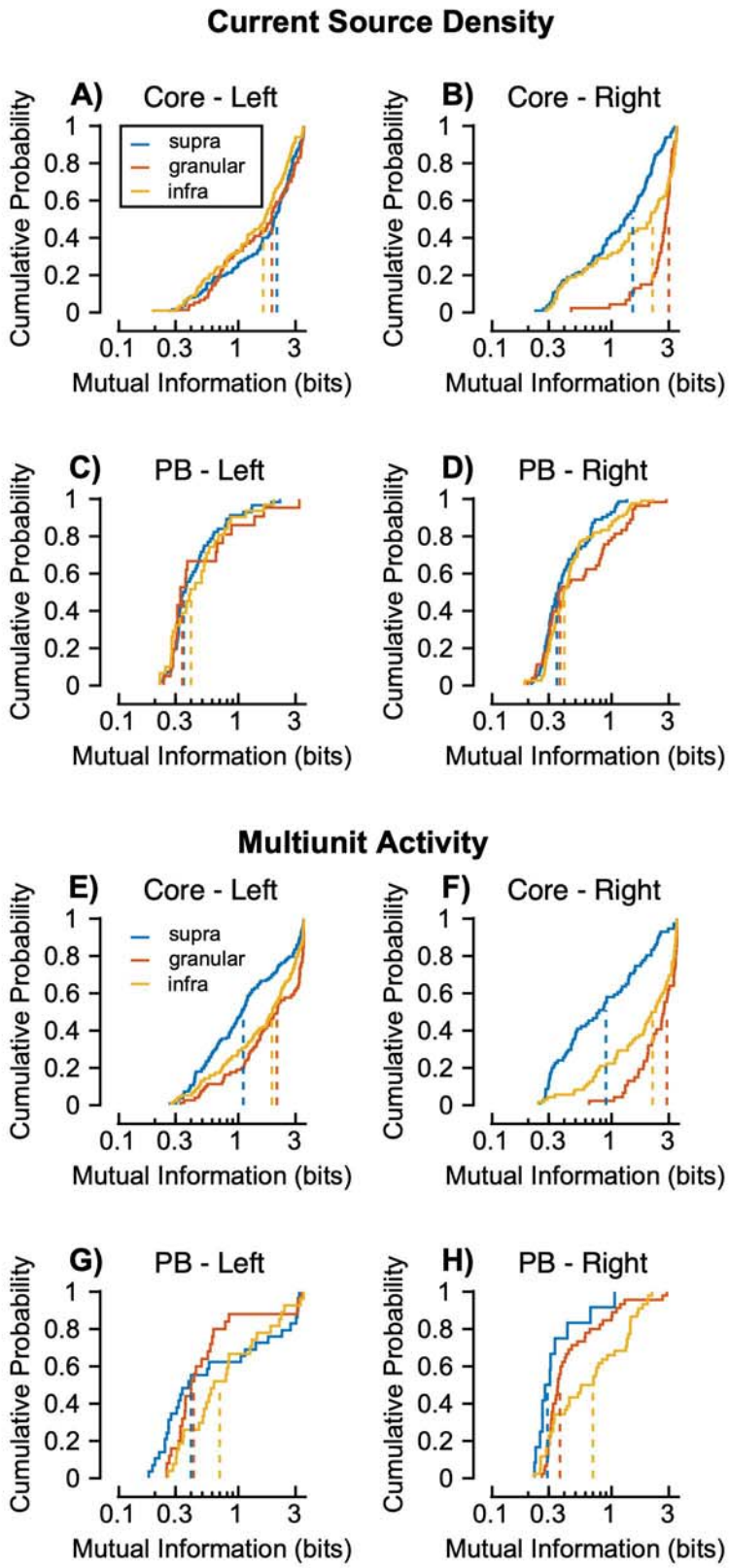
Across the dataset, we observed differences in AM encoding in different hemispheres, areas, and cortical layers. Figure 3 displays cumulative distributions of the MI values derived from CSD and MUA signals in the different hemispheres, layers, and cortical areas. A mixed effects model was used to quantify the effects of each factor (see Methods). MI was used as the dependent variable, while layer, area, and hemisphere were used as the independent variables (see Methods). Mixed-effects modeling revealed that mutual information (MI) calculated using current source density (CSD) was modulated by significant two-way interactions, which qualified the interpretation of main effects (Table 1). Consistent with previous work in human and nonhuman primate auditory cortex (Boemio et al., 2005; O'Connell et al., 2015) (also see Discussion section), the right hemisphere displayed reduced mutual information about AM relative to the left hemisphere (Table 1). The positive statistical interaction, *Layer*Hemisphere*, provides insight into how these hemispheric differences emerge: the granular and infragranular layer CSD signals did not display left hemisphere dominance (Figure 3 A-B; Figure 3 C-D; Table 1). This indicates that the enhanced encoding in the left hemisphere is localized mainly to the supragranular layers, and is not inherited from the medial geniculate body (as would be expected if the granular layer displayed left-hemisphere dominance); moreover, the difference is visible in the distributions of supragranular MI, which is visibly shifted to the left in core AC in Figure 3A vs. Figure 3B. MI derived from multiunit spiking activity (MUA) revealed a partially overlapping pattern (Table 2). Consistent with the CSD results, there was an overall reduction in MI in the right hemisphere, restricted to the supragranular layers (Figure 3 E-H). This was quantified with granular and infragranular layers exhibiting significant positive interactions with the right hemisphere (Table 2). Similar strength of the granular and infragraular interactions with

the right hemisphere were observed for the MI values calculated based on click train responses (Tables 3-4). Together, these results highlight the importance of cortical layers when considering interhemispheric effects, and may aid in interpreting previous human intracranial and noninvasive studies, which will be further elaborated upon in the Discussion section.

These results also indicated hierarchical differences between core AC and parabelt. The main effect of the parabelt was reduced MI, derived from the CSD and MUA, though this differed across layers. Specifically, enhanced MI observed in the granular and infragranular MUA across the dataset (Rightward shift in the orange and gold traces in Figure 3 E-F; Table 2), was attenuated in the parabelt relative to core AC, quantified by a significant interaction of granular layers with parabelt, and infragranular layers with parabelt in the AM noise and click datasets (Table 2-4). As with the hemispheric differences, these results highlight how interareal encoding differences vary across cortical layers, which will be considered in more depth in the Discussion section.

A notable difference in the AM noise and click train conditions was that, in the infragranular right hemisphere parabelt dataset, no channels exhibited above-chance mutual information values. While this reinforces the finding that the left hemisphere and core auditory cortex exhibit enhanced AM encoding, it suggests that click trains were not optimal for interareal comparisons. For this reason, the remaining analyses focus on the responses to AM noise.

A necessary caveat to the hemispheric differences we report is that only monkey W contributed to the right hemisphere dataset, while monkey W and the other 3 NHPs contributed to the left hemisphere dataset (see Methods). Thus, as is often the case with nonhuman primate studies often limited to 2-4 subjects, statistical generalization across animals is not warranted. Instead, these studies make statistical inferences across a neuronal population (e.g. the supragranular layers of primary auditory cortex). In this scenario, increasing the sample size from one NHP to two, does not formally enhance statistical power, as has been noted in previous work (Fries and Maris, 2022). Moreover, because we observe robust interhemispheric differences, it seems unwarranted to combine data across hemispheres.


Current Source Density
A) Core - Left
B) Core - Right
C) PB - Left
D) PB - Right
supra
granular
infra
Cumulative Probability
Mutual Information (bits)
0.1
0.3
1
3
0
0.2
0.4
0.6
0.8
1
Multiunit Activity
E) Core - Left
F) Core - Right
G) PB - Left
H) PB - Right

391

Figure 3 – Cumulative distributions of mutual information values extracted from responses to AM noise across areas, layers, and hemispheres. Distributions of MI values derived from the CSD signal are depicted for left and right hemispheres of core AC (**A-B**), and the parabelt (**C-D**). Distributions of MI values derived from the MUA signal are depicted for left and right hemispheres of core AC (**E-F**), and the parabelt (**G-H**). Only above-chance values are displayed (see Methods). Dashed lines demarcate the median value of each distribution.

Table 1 – Significant terms in a mixed effects model analysis of mutual information values extracted from CSD responses to AM noise (1.6-200 Hz)

Model: MI ~ Area + Layer + Hemisphere + Area*Layer + Layer*Hemisphere + Hemisphere*Area + (1|Channel)			
Term:	t-statistic	Degrees of freedom	P-value
Right hemi	-3.3	853	0.0009
Parabelt	-10.2	853	$4*10^{23}$
Granular layer*Right Hemisphere	5.7	853	$1*10^{-8}$
Infragranular layer*Right Hemisphere	4.1	853	$5*10^{-5}$
Granular layer*Parabelt	-3.2	853	0.0016

Table 2 – Significant terms in a mixed effects model analysis of mutual information values extracted from MUA responses to AM noise (1.6-200 Hz)

Model: MI ~ Area + Layer + Hemisphere + Area*Layer + Layer*Hemisphere + Hemisphere*Area + (1|Channel)			
Term:	t-statistic	Degrees of freedom	P-value
Right hemi	-2.2	689	0.029
Granular layer	5.4	689	$1*10^{-7}$

Infragranular layer	3.8	689	0.0001
Right Hemisphere	-2.2	689	0.029
Granular layer*Right Hemi	3.4	689	$8*10^{-4}$
Infragranular*Right Hemi	2.8	689	0.005
Granular layer*Parabelt	-4.9	689	$1*10^{-6}$
Infragranular layer*Parabelt	-2.2	689	0.03
Right Hemi*Parabelt	-2.7	689	0.007

Table 3 – Significant terms in a mixed effects model analysis of mutual information values extracted from CSD responses to click trains (1.6-200 Hz)

Model: MI ~ Area + Layer + Hemisphere + Area*Layer + Layer*Hemisphere + Hemisphere*Area + (1|Channel)			
Term:	t-statistic	Degrees of freedom	P-value
Granular layer	3.96	1180	$8*10^{-5}$
Parabelt	-10.4	1180	$1.7*10^{-24}$
Granular layer*Right Hemisphere	4.1	1180	$4*10^{-5}$
Infragranular layer*Right Hemisphere	4.2	1180	$2.6*10^{-5}$
Granular layer*Parabelt	-2.7	1180	0.0066
Right Hemi*Parabelt	-2.5	1180	0.0137

Table 4 – Significant terms in a mixed effects model analysis of mutual information values extracted from MUA responses to click trains (1.6-200 Hz)

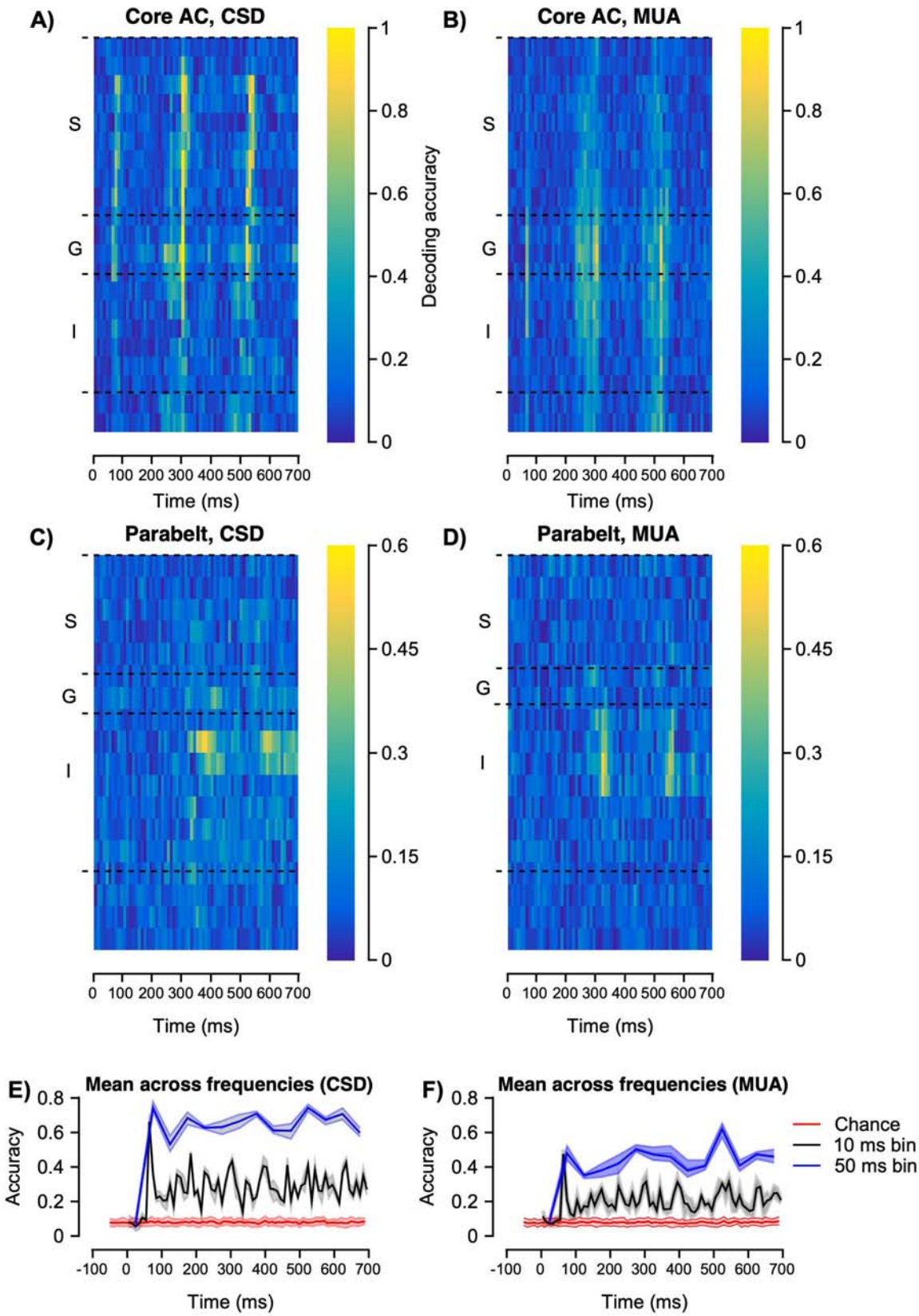
Model: MI ~ Area + Layer + Hemisphere + Area*Layer + Layer*Hemisphere + Hemisphere*Area + (1|Channel)			
Term:	t-statistic	Degrees of freedom	P-value
Granular layer	5.5	997	$3.7*10^{-8}$
Infragranular layer	5.1	997	$5*10^{-7}$
Parabelt	-6.8	997	$1*10^{-11}$
Infragranular layer*Right Hemisphere	4.0	997	$6*10^{-5}$
Granular layer*Right Hemisphere	3.45	997	0.0058
Granular layer*Parabelt	-2.7	997	0.0069
Infragranular*Parabelt	-2.4	997	0.0170
Right hemi*Parabelt	-2.7	997	0.0067

AM encoding over time

Having characterized how the encoding of AM differs across regions using the entire response waveform, it was then of interest to investigate how encoding unfolds across time, which can address questions, albeit indirectly, about spike-timing vs. spike rate codes in previous studies. Analyzing decoding results over time also addresses what time-points in the stimulus response may be driving the decoding accuracy in our previous results, which utilized the entire response waveform (Figures 1-3). To address these issues, responses were binned into 10 or 50 ms bins and the decoding process was repeated for each bin (see Methods). Figure 4 shows example decoding results over time, for the 6.3 Hz AM noise stimulus for core (Figure 4A-B) and parabelt (Figure 4C-D). These results complement the mutual information analyses, by showing that at a low modulation frequency (below ~25 Hz), sites in the parabelt often displayed far above-chance decoding accuracy. They also suggest that the highest magnitude response, in CSD or MUA, often encountered at the onset of the stimulus (Figure 1), does not always correspond to the peak decoding accuracy, which could be found throughout the duration of the

stimulus response. Consistent with this, decoding accuracy, when averaged across modulation frequencies, was relatively stable throughout the analysis window (Figure 4E-F) for both the CSD and MUA signals. These results also suggest that the time-varying aspect of the response to AM contributes heavily to its discriminability, inconsistent with a rate-code, because, if an average rate-code were at work, the peak decoding accuracy would be expected early in the stimulus response (see Discussion). This observation likely depends on our use of population-level activity, which has been found to exhibit robust temporal encoding in a previous single-unit study (see Introduction and Discussion; Downer et al. 2021).

To quantify these observations, the peak decoding time, across all modulation frequencies, was compared across layer, area, and hemisphere. A mixed effects model revealed similar results for the CSD and MUA signals. Namely, the right hemisphere had significantly earlier peak times (Tables 5-6), though this appeared to be restricted to core AC, indicated by the significant interaction of parabelt with the right hemisphere (Figure 5; Tables 5-6). The parabelt had significantly later peak times in both the CSD and MUA signals. Where CSD and MUA differed in their AM encoding was in the infragranular layers of the parabelt, which had earlier peak times than predicted by the infragranular layer term on its own (Tables 5-6). The same mixed effects model analysis revealed similar size and direction of effects for the 50 ms binned decoding results (Tables 7-8).


A)
Core AC, CSD
B)
Core AC, MUA
C)
Parabelt, CSD
D)
Parabelt, MUA
S
G
I
Decoding accuracy
Time (ms)
E)
Mean across frequencies (CSD)
F)
Mean across frequencies (MUA)
Accuracy
Chance
10 ms bin
50 ms bin

444

Figure 4 – Decoding AM noise over time. A) Representative example of decoding the CSD signal over time, for the 6.3 Hz AM noise condition in left hemisphere core auditory cortex. The decoding results for the same site's MUA are shown in (**B**). **C**) Representative example of decoding the CSD signal over time, for the 6.3 Hz AM noise condition in the right hemisphere parabelt auditory cortex. The decoding results for the same site's MUA are shown in (**D**). In all heat maps, the bin size was 10 ms. **E**) Mean decoding accuracy across all AM noise modulation frequencies, using the CSD signal from a granular layer channel (11) in panel A. **F**) Mean (across conditions) decoding accuracy using the MUA signal from a granular layer channel (11) in panel B. Data in panels **E** and **F** were decoded using either 10 or 50 ms time bins, indicated by the black and blue traces, which are plotted at the bin center. Chance level decoding was calculated using permutation analyses and is indicated by the red trace. Error bands represent 95% confidence intervals (1,000 permutations).

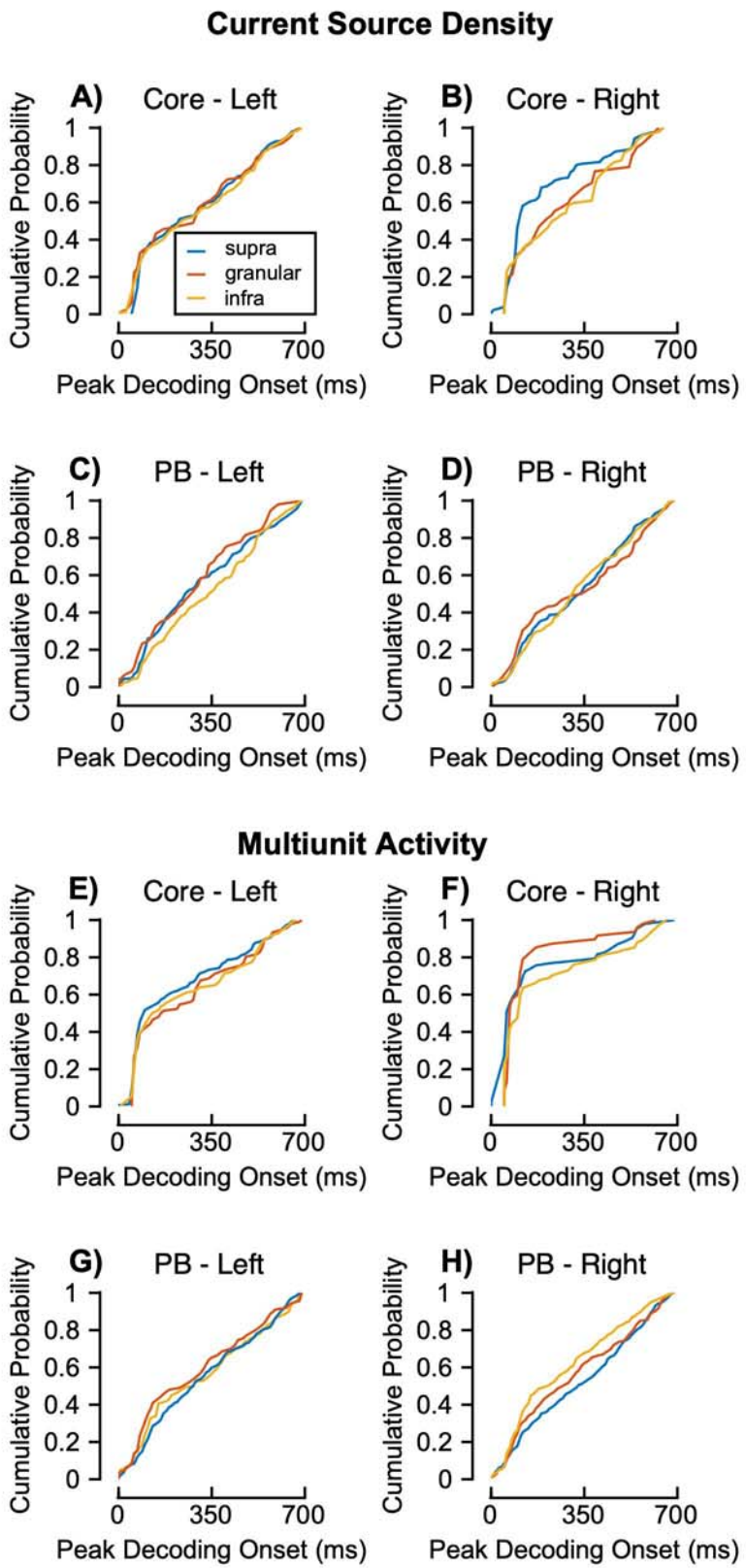


Figure 5 – Cumulative distributions of peak decoding onset in core and parabelt in response to amplitude modulated noise using a 10 ms bin size. Distributions of peak decoding accuracy values, derived from the CSD signal, are depicted for left and right hemispheres of core AC (**A-B**), and the parabelt (**C-D**). Distributions of peak values derived from the MUA signal are depicted for left and right hemispheres of core AC (**E-F**), and the parabelt (**G-H**). Only above-chance values are displayed (see Methods).

Table 5 – Mixed effects model of peak decoding onset time in the CSD signal, using a 10 ms bin (significant terms only)

Model: Peak Time ~ Area + Layer + Hemisphere + Area*Layer + Layer*Hemisphere + Hemisphere*Area + (1|Channel)			
Term:	t-statistic	Degrees of freedom	P-value
Right hemi	-3.0	1716	0.002
Parabelt	2.3	1716	0.02
Right Hemisphere*Parabelt	3.1	1716	0.0019

Table 6 – Mixed effects model of peak decoding onset time in the MUA signal, using a 10 ms bin (significant terms only)

Model: Peak Time ~ Area + Layer + Hemisphere + Area*Layer + Layer*Hemisphere + Hemisphere*Area + (1|Channel)			
Term:	t-statistic	Degrees of freedom	P-value
Right hemi	-2.3	1716	0.019
Parabelt	5.1	1716	$2*10^{-7}$
Infragranular	2.2	1716	0.03
Infragranular layer*Parabelt	-2.8	1716	0.0057
Right Hemi*Parabelt	2.9	1716	0.004

Table 7 – Mixed effects model of peak decoding onset time in the CSD signal, using a 50 ms bin (significant terms only)

Model: Peak Time ~ Area + Layer + Hemisphere + Area*Layer + Layer*Hemisphere + Hemisphere*Area + (1|Channel)			
Term:	t-statistic	Degrees of freedom	P-value

Right hemi	-2.7	1716	0.0078
Parabelt	2.4	1716	0.019

Table 8 – Mixed effects model of peak decoding onset time in the MUA signal, using a 50 ms bin (significant terms only)

Model: Peak Time ~ Area + Layer + Hemisphere + Area*Layer + Layer*Hemisphere + Hemisphere*Area + (1|Channel)			
Term:	t-statistic	Degrees of freedom	P-value
Parabelt	4.7	1716	$3*10^{-6}$
Granular	2.1	1716	0.03
Infragranular	2.4	1716	0.01
Granular layer*Parabelt	-3.2	1716	0.002
Infragranular layer*Parabelt	-2.6	1716	0.007

Discussion

The present report builds on previous work on how sounds are encoded across the auditory cortical hierarchy. Previous NHP studies have described hierarchical progressions in stimulus selectivity that coincide with the hierarchical increase in response latency in primary, secondary, and tertiary (parabelt) auditory cortex (Camalier et al., 2012). For instance, NHP studies of AM encoding have discovered notable differences in encoding between primary and secondary auditory cortex. Namely, while primary AC encodes AM depth with an increasing rate code, secondary belt auditory cortex exhibits an opponent (increasing/decreasing) code for AM (Downer et al., 2017; Johnson et al., 2024). Consistent with this, fMRI studies have noted some degree of selectivity for longer temporal intervals in nonprimary auditory cortices, particularly in the parabelt (Dheerendra et al., 2021). An extended temporal integration window, and broad frequency tuning likely contributes to the parabelt's enhanced response, relative to core AC, to harmonic sounds with varying degrees of temporal coherence across frequency channels, termed

“auditory figure ground” stimuli (Schneider et al., 2018). Across studies of the cortical hierarchy, there appears to be a notable reduction in response to simple stimuli, such as tones, and an increase in selectivity for complex stimuli with characteristics, e.g. low modulation frequencies. This likely underpins enhanced selectivity for species-specific communication sounds, and associated visual stimuli, as one ascends the hierarchy (Mackey et al., 2024b).

Previous laminar studies. The results presented here provide the first study of AM encoding utilizing laminar resolution at multiple levels across the NHP auditory cortical hierarchy. Previous laminar studies in cat primary auditory cortex found that the extragranular layers preferentially encode slower AM relative to the thalamorecipient granular layer (Atencio and Schreiner, 2010), and that spectrotemporal feature selectivity is hierarchical i.e. Granular > Supragranular > Infragranular (Atencio et al., 2009). This is consistent with what we observed for MUA in core primary AC, though interestingly the parabelt appeared to differ in this respect (Figure 3 E-H). Nonetheless, these studies, in combination with the present report, describe some of the only data on the intracortical circuits underlying temporal sound envelope encoding in auditory cortex. The present report now builds on this knowledge by quantifying laminar-resolved AM encoding in the parabelt.

Task engagement. Though the NHPs in the present report were not engaged in a task, previous studies have suggested that stimulus complexity and task relevance exert substantial influence on stimulus responses, particularly in nonprimary cortical areas. In secondary and tertiary auditory cortices of ferrets (the dorsal posterior ectosylvian gyrus, and rostral ventral posterior auditory field, respectively), hierarchical differences, such as the progressive influence of task-related factors, have been well-studied (Atiani et al., 2014; Elgueda et al., 2019). A similar hierarchical progression is present between primary and secondary auditory cortex of NHPs discriminating AM noise, namely, single-unit discriminability via rate and temporal codes are enhanced by task-engagement (Niwa et al., 2015). Entrainment to rhythmic stimuli also appears to be behaviorally-gated in this same sense (O’Connell et al., 2015). Thus, if our NHPs had been engaged in a task, particularly one that indexed AM feature discrimination, we would expect enhanced encoding relative to what we report here.

Decoding studies. Previous decoding studies offer insight into the present results (Downer et al., 2021; Tovar et al., 2020). Tovar et al. (2020) leveraged multivariate pattern analysis across channels to reveal that orientation information is relatively well preserved across all layer compartments in V1, and eye-of-origin information decreases in the infragranular layers. Tovar et al. also noted that repetition suppression, which is greatest in the extragranular layers, contributed to their enhanced decoding of stimulus repetition. This may be analogized to our results, where the granular layer often exhibited the greatest mutual information about different modulation frequencies, likely due to its rapid response characteristics. Though this pattern was most notable in core AC, and the parabelt displayed a less clear granular > supragranular > infragranular gradient in AM encoding. Another study that utilized decoding analysis that is relevant to the present results is Downer et al. (2021). Downer et al. utilized decoding analyses on simulated and empirical neuronal populations from single-unit recordings, and revealed robust population-level synchronization to a wide range of modulation frequencies, 4-512 Hz (Downer et al., 2021). This is consistent with our MUA measures in core AC. Both studies suggest code-agnostic measures (e.g. linear discriminant analysis in the present report) on local population responses can reveal information not readily apparent in single-unit measures using phase-locking or an average rate code. Such code-agnostic measures also avoid commonly noted problems with vector strength and the Rayleigh test for synchrony, namely potential for false discoveries (Yin et al., 2011), and parametric statistical assumptions (Kajikawa and Hackett, 2005).

Interhemispheric differences. Our analysis demonstrates similar hierarchical and inter-hemispheric differences to human subjects in Boemio et al. (2005), and our results resemble previous NHP studies demonstrating sound-driven hemispheric asymmetries in the lateral sulcus, and the extent of the superior temporal gyrus, including the temporal pole (Poremba et al., 2003, 2004). Our results suggest the suprgranular layers are the main contributing layer to the left-hemisphere dominance in encoding rapid temporal fluctuations, which aligns with other that has defined supragranular sources underlying noninvasive measures in fMRI and MEG studies (e.g. (Kajikawa et al., 2024; Leszczyński et al., 2020)). Additionally, a previous NHP study has similarly demonstrated enhanced entrainment of low frequency (~1.6 Hz) oscillations in left supragranular A1, which corroborates our results (O'Connell et al., 2015).

REFERENCES

Atencio, C.A., and Schreiner, C.E. (2010). Laminar diversity of dynamic sound processing in cat primary auditory cortex. J. Neurophysiol. *103*, 192–205.

Atencio, C.A., Sharpee, T.O., and Schreiner, C.E. (2009). Hierarchical computation in the canonical auditory cortical circuit. Proc. Natl. Acad. Sci. U. S. A. *106*, 21894–21899.

Atiani, S., David, S. V., Elgueda, D., Locastro, M., Radtke-Schuller, S., Shamma, S.A., and Fritz, J.B. (2014). Emergent Selectivity for Task-Relevant Stimuli in Higher-Order Auditory Cortex. Neuron *82*, 486–499.

Banno, T., Shirley, H., Fishman, Y.I., and Cohen, Y.E. (2023). Changes in neural readout of response magnitude during auditory streaming do not correlate with behavioral choice in the auditory cortex. Cell Rep. *42*, 113493.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B *57*, 289–300.

Boemio, A., Fromm, S., Braun, A., and Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. Nat. Neurosci. *8*, 389–395.

Camalier, C.R., D'Angelo, W.R., Sterbing-D'Angelo, S.J., de la Mothe, L.A., and Hackett, T.A. (2012). Neural latencies across auditory cortex of macaque support a dorsal stream supramodal timing advantage in primates. Proc. Natl. Acad. Sci. *109*, 18168–18173.

Dheerendra, P., Baumann, S., Joly, O., Balezeau, F., Petkov, C.I., Thiele, A., and Griffiths, T.D. (2021). The Representation of Time Windows in Primate Auditory Cortex. Cereb. Cortex 1–13.

Ding, N., Patel, A.D., Chen, L., Butler, H., Luo, C., and Poeppel, D. (2017). Temporal modulations in speech and music. Neurosci. Biobehav. Rev. *81*, 181–187.

Downer, J.D., Niwa, M., and Sutter, M.L. (2017). Hierarchical differences in population coding within auditory cortex. J. Neurophysiol. *118*, 717–731.

Downer, J.D., Bigelow, J., Runfeldt, M.J., and Malone, B.J. (2021). Temporally precise population coding of dynamic sounds by auditory cortex. J. Neurophysiol. *126*, 148–169.

Elgueda, D., Duque, D., Radtke-schuller, S., Yin, P., David, S. V., Shamma, S.A., and Fritz, J.B. (2019). State-dependent encoding of sound and behavioral meaning in a tertiary region of the

ferret auditory cortex. Nat. Neurosci. *22*.

Fishman, Y.I., Reser, D.H., Arezzo, J.C., and Steinschneider, M. (2001). Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. *151*, 167–187.

Fishman, Y.I., Arezzo, J.C., and Steinschneider, M. (2004). Auditory stream segregation in monkey auditory cortex: effects of frequency separation, presentation rate, and tone duration. J. Acoust. Soc. Am. *116*, 1656–1670.

Freeman, J.A., and Nicholson, C.H. (1975). Experimental optimization of current source-density technique for anuran cerebellum. J. Neurophysiol. *38*, 369–382.

Fries, P., and Maris, E. (2022). What to do if N is two? J. Cogn. Neurosci. *34*, 1114–1118.

Hackett, T.A., Preuss, T.M., and Kaas, J.H. (2001). Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. J. Comp. Neurol. *441*, 197–222.

Hackett, T.A., de la Mothe, L.A., Camalier, C.R., Falchier, A., Lakatos, P., Kajikawa, Y., and Schroeder, C.E. (2014). Feedforward and feedback projections of caudal belt and parabelt areas of auditory cortex: Refining the hierarchical model. Front. Neurosci. *8*, 1–21.

Hamilton, L.S., Oganian, Y., Hall, J., and Chang, E.F. (2021). Parallel and distributed encoding of speech across human auditory cortex. Cell 1–14.

Henry, K.S., Neilans, E.G., Abrams, K.S., Idrobo, F., and Carney, L.H. (2016). Neural correlates of behavioral amplitude modulation sensitivity in the budgerigar midbrain. J. Neurophysiol. *115*, 1905–1916.

Henry, K.S., Abrams, K.S., Forst, J., Mender, M.J., Neilans, E.G., Idrobo, F., and Carney, L.H. (2017). Midbrain Synchrony to Envelope Structure Supports Behavioral Sensitivity to Single-Formant Vowel-Like Sounds in Noise. JARO - J. Assoc. Res. Otolaryngol. *18*, 165–181.

Hoglen, N.E.G., Larimer, P., Phillips, E.A.K., Malone, B.J., and Hasenstaub, A.R. (2018). Amplitude modulation coding in awake mice and squirrel monkeys. J. Neurophysiol. *119*, 1753–1766.

Johnson, J.S., Yin, P., O'Connor, K.N., and Sutter, M.L. (2012). Ability of primary auditory cortical neurons to detect amplitude modulation with rate and temporal codes: Neurometric analysis. J. Neurophysiol. *107*, 3325–3341.

Johnson, J.S., Niwa, M., O'Connor, K., Malone, B.J., and Sutter, M.L. (2024). Hierarchical emergence of opponent coding in auditory belt cortex. BioRxiv 1–37.

Joris, P.X., and Yin, T.C.T. (1992). Responses to amplitude modulated tones in the auditory nerve of the cat. J. Acoust. Soc. Am. *91*, 215–232.

Joris, P.X., Schreiner, C.E., and Rees, A. (2004). Neural Processing of Amplitude-Modulated Sounds. Physiol. Rev. *84*, 541–577.

Kajikawa, Y., and Hackett, T.A. (2005). Entropy analysis of neuronal spike train synchrony. J. Neurosci. Methods *149*, 90–93.

Kajikawa, Y., and Schroeder, C.E. (2011). How Local Is the Local Field Potential? Neuron *72*, 847–858.

Kajikawa, Y., Frey, S., Ross, D., Falchier, A., Hackett, T.A., and Schroeder, C.E. (2015). Auditory Properties in the Parabelt Regions of the Superior Temporal Gyrus in the Awake Macaque Monkey: An Initial Survey. J. Neurosci. *35*, 4140–4150.

Kajikawa, Y., Smiley, J.F., and Schroeder, C.E. (2017). Primary generators of visually evoked field potentials recorded in the macaque auditory cortex. J. Neurosci. *37*, 10139–10153.

Kajikawa, Y., Mackey, C.A., and O'Connell, M.N. (2024). Laminar pattern of sensory-evoked dynamic high-frequency oscillatory activity in the macaque auditory cortex. Cereb. Cortex *34*.

Lakatos, P., Shah, A.S., Knuth, K.H., Ulbert, I., Karmos, G., and Schroeder, C.E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. J. Neurophysiol. *94*, 1904–1911.

Lakatos, P., O'Connell, M.N., Barczak, A., Mills, A., Javitt, D.C., and Schroeder, C.E. (2009). The Leading Sense: Supramodal Control of Neurophysiological Context by Attention. Neuron *64*, 419–430.

Lakatos, P., Musacchia, G., O'Connel, M.N., Falchier, A.Y., Javitt, D.C., and Schroeder, C.E. (2013). The Spectrotemporal Filter Mechanism of Auditory Selective Attention Supplemental Info. Neuron *77*, 750–761.

Langner, G., and Schreiner, C.E. (1988). Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. J. Neurophysiol. *60*, 1799–1822.

Leszczyński, M., Barczak, A., Kajikawa, Y., Ulbert, I., Falchier, A.Y., Tal, I., Haegens, S., Melloni, L., Knight, R.T., and Schroeder, C.E. (2020). Dissociation of broadband high-frequency activity and neuronal firing in the neocortex. Sci. Adv. 1–13.

Lu, H., Mehta, A.H., and Oxenham, A.J. (2022). Methodological considerations when measuring and analyzing auditory steady-state responses with multi-channel EEG. Curr. Res. Neurobiol. *3*,

100061.
Mackey, C.A., Hauser, S., Schoenhaut, A.M., Temghare, N., and Ramachandran, R. (2024a). Hierarchical differences in the encoding of amplitude modulation in the subcortical auditory system of awake nonhuman primates. J. Neurophysiol. *132*, 1098–1114.
Mackey, C.A., O'Connell, M.N., Hackett, T.A., Schroeder, C.E., and Kajikawa, Y. (2024b). Laminar organization of visual responses in core and parabelt auditory cortex. Cereb. Cortex *34*.
Mitzdorf, U. (1985). Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. Physiol. Rev. *65*, 37–100.
Moody, D.B. (1994). Detection and discrimination of amplitude-modulated signals by macaque monkeys. J. Acoust. Soc. Am. *95*, 3499–3510.
Nelson, P.C., and Carney, L.H. (2007). Neural rate and timing cues for detection and discrimination of amplitude-modulated tones in the awake rabbit inferior colliculus. J. Neurophysiol. *97*, 522–539.
Niwa, M., O'Connor, K.N., Engall, E., Johnson, J.S., and Sutter, M.L. (2015). Hierarchical effects of task engagement on amplitude modulation encoding in auditory cortex. J. Neurophysiol. *113*, 307–327.
O'Connell, M.N., Barczak, A., Schroeder, C.E., and Lakatos, P. (2014). Layer Specific Sharpening of Frequency Tuning by Selective Attention in Primary Auditory Cortex. J. Neurosci. *34*, 16496–16508.
O'Connell, M.N., Barczak, A., Ross, D., McGinnis, T., Schroeder, C.E., and Lakatos, P. (2015). Multi-scale entrainment of coupled neuronal oscillations in primary auditory cortex. Front. Hum. Neurosci. *9*, 1–16.
Penikis, K.B., and Sanes, D.H. (2023). A redundant cortical code for speech envelope. J. Neurosci. *43*, 93–112.
Poremba, A., Saunder, R.C., Crane, A.M., Cook, M., Sokoloff, L., and Mishkin, M. (2003). Functional mapping of the primate auditory system. Science (80-.). *299*, 568–572.
Poremba, A., Malloy, M., Saunders, R.C., Carson, R.E., Herscovitch, P., and Mishkin, M. (2004). Species-specific calls evoke asymmetric activity in the monkey's temporal poles. Nature *427*, 448–451.
Rhode, W.S., Roth, G.L., and Recio-spinoso, A. (2010). Response properties of cochlear nucleus neurons in monkeys. Hear. Res. *259*, 1–15.

Rosen, S. (1992). Temporal information in speech. Phil. Trans. R. Soc. Lond. B *336*, 367–373.

Sayles, M., Füllgrabe, C., and Winter, I.M. (2013). Neurometric amplitude-modulation detection threshold in the guinea-pig ventral cochlear nucleus. J. Physiol. *591*, 3401–3419.

Schneider, F., Dheerendra, P., Balezeau, F., Ortiz-Rios, M., Kikuchi, Y., Petkov, C.I., Thiele, A., and Griffiths, T.D. (2018). Auditory figure-ground analysis in rostral belt and parabelt of the macaque monkey. Sci. Rep. *8*, 1–8.

Schroeder, C.E., and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. Trends Neurosci. *32*, 9–18.

Shi, K., Quass, G.L., Rogalla, M.M., Ford, A.N., Czarny, J.E., and Apostolides, P.F. (2024). Population coding of time-varying sounds in the nonlemniscal inferior colliculus. J. Neurophysiol. *131*, 842–864.

Steinschneider, M., Tenke, C.E., Schroeder, C.E., Javitt, D.C., Simpson, G. V, Arezzo, J.C., and Vaughan Jr, H.G. (1992). Cellular generators of the cortical auditory evoked potential initial component. Electroencephalogr. Clin. Neurophysiol. Potentials Sect. *84*, 196–200.

Steinschneider, M., Reser, D.H., Fishman, Y.I., Schroeder, C.E., and Arezzo, J.C. (1998). Click train encoding in primary auditory cortex of the awake monkey: evidence for two mechanisms subserving pitch perception. J. Acoust. Soc. Am. *104*, 2935–2955.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., and Pernier, J. (1996). Stimulus Specificity of Phase-Locked and Non-Phase-Locked 40 Hz Visual Responses in Human. J. Neurosci. *16*, 4240 LP – 4249.

Tovar, D.A., Westerberg, J.A., Cox, M.A., Dougherty, K., Carlson, T.A., Wallace, M.T., and Maier, A. (2020). Stimulus Feature-Specific Information Flow Along the Columnar Cortical Microcircuit Revealed by Multivariate Laminar Spiking Analysis. Front. Syst. Neurosci. *14*, 1–14.

Wang, X., Lu, T., Bendor, D., and Bartlett, E.L. (2008). Neural coding of temporal information in auditory thalamus and cortex. Neuroscience *154*, 294–303.

Yao, J.D., and Sanes, D.H. (2021). Temporal Encoding is Required for Categorization, But Not Discrimination. Cereb. Cortex *31*, 2886–2897.

Yin, P., Johnson, J.S., O'Connor, K.N., and Sutter, M.L. (2011). Coding of amplitude modulation in primary auditory cortex. J. Neurophysiol. *105*, 582–600.

Zion Golumbic, E., Ding, N., Bickel, S., Lakatos, P., Schevon, C.A., McKhann, G.M., Goodman,

R.R., Emerson, R., Mehta, A.D., Simon, J.Z., et al. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a “cocktail party.” Neuron *77*, 980–991.