Title: Evidence for time division multiplexing of multiple simultaneous items in a sensory coding bottleneck

Short title: Fluctuations as part of the neural code

Authors: V. C. Caruso[1,2,3,4]†, J. T. Mohl[1,2,3,4], C. Glynn[7,8], J. Lee[1,2,3,4]‡, S. M. Willett[1,2,3,4], A. Zaman[6], R. Estrada[5,9], S. Tokdar[1,6]†, J. M. Groh[1,2,3,4]†

Keywords: neural coding, time division multiplexing, statistical methods, perception

ABSTRACT

How the brain preserves information about multiple simultaneous items is poorly understood. We provide evidence that single neurons can represent multiple different stimuli by interleaving different signals across time (aka time-division multiplexing). We recorded single units in an auditory region, the inferior colliculus, while monkeys localized 1 or 2 simultaneous sounds. During dual-sound trials, we found that some neurons fluctuated between firing rates observed for each single sound, either on a whole-trial or on a sub-trial timescale. These fluctuations were correlated in pairs of neurons, could be predicted by the state of local field potentials prior to sound onset, and, in one monkey, predicted which sound would be reported first. Alternation between activity patterns corresponding to each of multiple items may be a general strategy to enhance the brain processing capacity, suggesting a potential connection between such disparate phenomena as variable neural firing, neural oscillations, and limits in attentional/memory capacity.

In the natural world many stimuli or events occur at the same time. Sensory neurons in the brain are broadly tuned and potentially responsive to more than one such stimulus, raising the question of how information about multiple simultaneous items can be preserved. We investigated whether the brain solves this problem at the neuronal level using *time division multiplexing*, i.e. via signals fluctuating between the activity patterns evoked by each stimulus alone (Figure 1A). Such time division multiplexing is commonly used in telecommunications. Similar strategies have been postulated to occur in some form in the brain [1-11], but empirical evidence and statistical methods of assessment are presently lacking.

Specifically, we tested how neurons in the monkey inferior colliculus (IC) respond when two sounds are presented simultaneously from different locations and both must be behaviorally reported. The IC is an early [12, 13]) and nearly obligatory station along the auditory pathway [14], and it is thought to be essential for accurate sound localization behavior [15-17]. Importantly, the IC is thought to encode sound location via a monotonic firing rate (meter) code in which individual neurons respond broadly across a range of spatial positions but the level of activity is correlated with the location of the sound [18-21]. We developed a novel statistical approach to assess whether signal fluctuations in the IC serve to interleave information about both sounds at a variety of time scales.

We found that a subpopulation of IC neurons exhibited fluctuating activity consistent with switching between individual sound responses at different time scales. Such activity fluctuations were masked in conventional aggregate analysis of spiking activity (trial-and-time pooled spike counts), which indicated that average dual-sound activity levels were intermediate between the average single sound responses. The state of the network prior to dual-stimulus onset (assessed by local field potentials) predicted the slower whole-trial fluctuations, and in turn, these fluctuations could predict which sound location the monkey reported first. These data support time-division multiplexing as a viable strategy for encoding simultaneously presented stimuli. Finally, we discuss how a multiplexed code can be read out, as well as the broad implications of activity fluctuations for interpreting variability and other aspects of neural encoding.

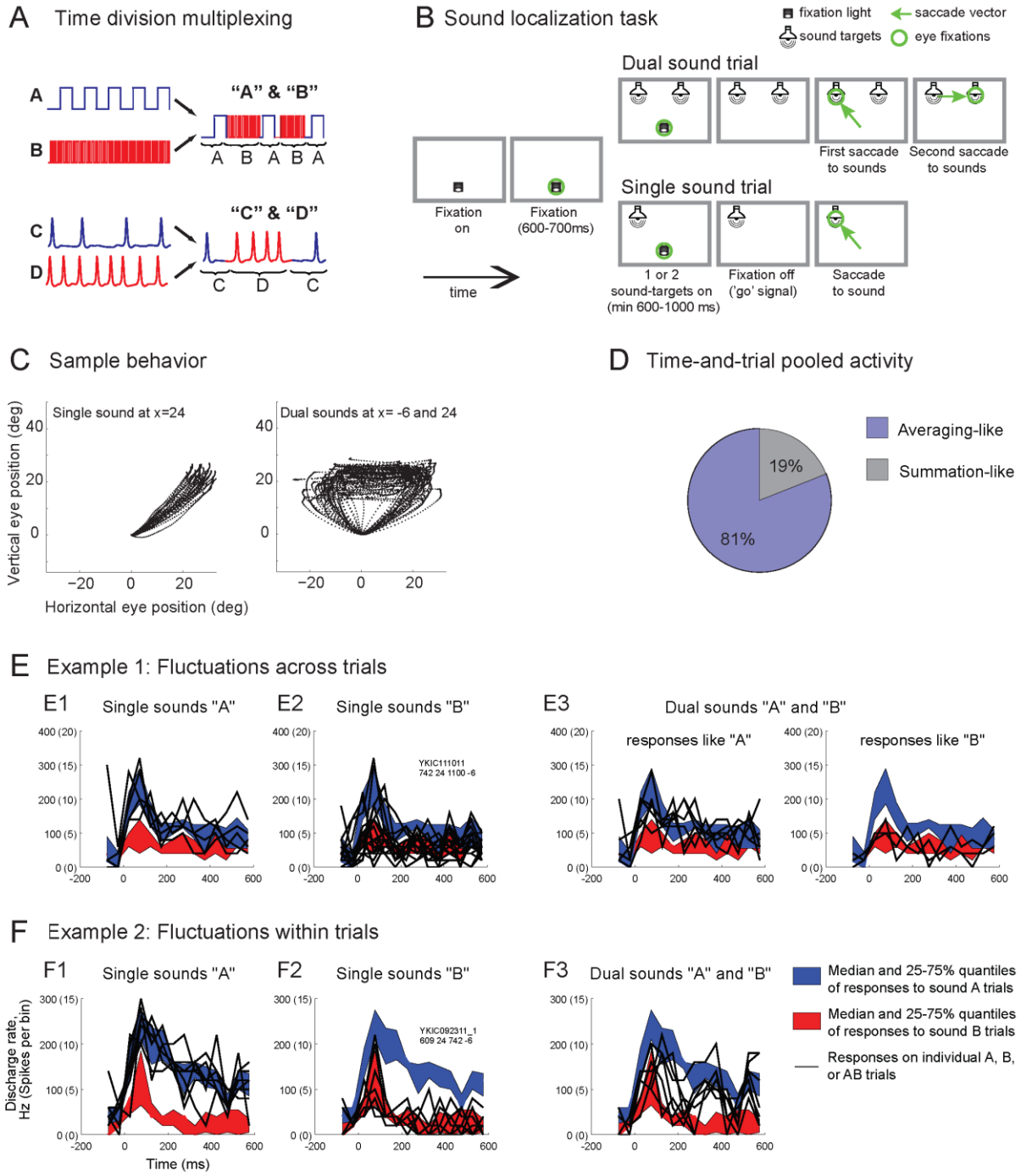


Figure 1. Schematics of time division multiplexing, task and visualization of individual trials' responses for two example cells. (A) *In telecommunications, multiple signals (A and B) can be conveyed along a single channel by interleaving samples of each, thus increasing the amount of information transmitted by a single physical resource. Here we investigated whether the brain might employ a similar strategy: do neurons encode multiple items (C and D) using spike trains*

that alternate between the firing rates corresponding to each item, at some unknown time scale? ***(B)*** *Sound localization task. Two monkeys were successfully trained to report one or two simultaneous (bandlimited noise) sounds by saccading at them. See figure S2 for accuracy.* ***(C)*** *Eye traces of the saccades towards one (left) or two targets (right) during a sample session.* ***(D)*** *Time-and-trial aggregated dual-sound responses resemble the averaging more than the summation of single sound responses. For 81% of the triplets tested, the absolute values of the Z-score of each dual-sound response relative to the average were smaller than those relative to the sum.* ***(E-F)*** *Visualization of individual trials of two IC neurons in which dual-sound responses alternate between firing rates corresponding to single sounds, across trials for the neuron in E, or within trials for the neuron in F. In each panel, the red and blue shaded areas indicate the median and central 50% of the data on the single sound trials. The black traces are the individual trials, for single sound and dual sound trials as indicated above the panel. For the neuron in E1-2-3, individual traces on dual sound trials were classified based on whether they matched the responses to single sounds A and B (A vs. B assignment score, see Methods) and are plotted in two separate panels accordingly. For the neuron in F1-2-3, the fluctuations occurred faster, within trials, and are plotted in the same panel.*

RESULTS

1. Monkeys can report the locations of two simultaneous sounds, but conventional time-and-trial pooled activity does not account for this behavior

We first tested whether monkeys can perceptually preserve information about multiple sounds presented simultaneously. Two monkeys performed a localization task (Figure 1B) in which they made eye movements to each of the sounds they heard: one saccade on single-sound trials and two saccades in sequence on dual-sound trials (Figure 1C and S1). The sounds were separated horizontally by 30 degrees and consisted of band-limited noise with different central frequencies. The sounds were thus physically distinguishable in principle, and humans are able to do so [22-24]. The monkeys learned the task successfully, and, like humans, typically performed better when the frequency separation between the two sounds was larger (Figure 1C, Figure S1, ~72 vs. ~77% correct for frequency differences of 3.4 vs. 6.8 semitones).

If the monkeys can report the locations of two sounds presented simultaneously, it follows that their brains, and the IC in particular, must preserve information about both sound items. To investigate the neural basis of this perceptual ability, we recorded 166 single units from the left and right IC of two monkeys (see Methods), and evaluated the activity patterns evoked on dual-sound trials (involving a given pair of sound locations and frequencies, 'AB') in comparison to the activity patterns evoked on the matching single sound trials ('A' and 'B'). We refer to such related sets of stimuli as '**triplets'**.

Conventional analysis of spike data typically involves two simplifications: spikes are counted within a fairly long window of time, such as a few hundred milliseconds, and activity is pooled across trials for statistical analysis. If IC neurons interleave signals related to each of the two sounds, then they might appear to show "averaging" responses on dual (or "AB") trials when activity is pooled across time and across trials. But they should not appear to show "summation" responses, i.e. in which the responses on dual-sound trials resemble the sum of the responses exhibited on single-sound trials involving the component sounds. Such summation has been observed in some neural populations in areas such as primary visual cortex [25, 26], the hippocampus [27], and the superior colliculus [28] when multiple stimuli are presented.

And indeed, IC dual sound responses do not generally appear to sum the responses to the component single sounds. Using an analysis similar to that of [28], dual-sound responses were converted to Z-scores relative to either the sum or the average of the corresponding single-sound responses (see Methods). For 81% of the tested triplets, the Z-score values relative to the average were smaller than those relative to the sum, indicating that the responses more closely resembled averaging (Figure 1D, see also Supplementary Figure 2 and 3)

However, such apparent averaging response patterns are inconsistent with the behavioral results: if the neurons truly responded at an average firing rate, then presumably the monkeys should respond to dual sounds as if there were only a single sound at the midpoint of the two sources (Supplementary Figure 2A). Since monkeys can indicate the locations of *both* sounds (Figure 1C), time-division multiplexing might provide a better explanation for so-called averaging response patterns.

2. Fluctuations in responses to dual sounds appear consistent with time-division multiplexing at various time scales

To determine whether neural activity fluctuates within and/or between trials, creating an overall averaging response but retaining information about each sound at distinct moments, we developed a series of statistical analyses that test for the presence of various forms of alternation in firing rates. Several unknown parameters must be taken into consideration when testing for activity fluctuations, specifically, the time scale, repeatability, and potential correlations across the neural population. We made minimal assumptions about the time scale at which neurons might alternate between encoding each stimulus, and considered that switching behaviors might vary from trial to trial and/or across time within a trial, as suggested by the visual inspection of single unit responses.

2.1 Visualization

Figure 1E-F shows the activity of two example neurons on dual-sound trials compared to their matched single-sound trials. The colored backgrounds illustrate the median and 25-75% quantiles of the activity on single-sound trials, in 50ms time bins. Superimposed on these backgrounds is the activity on individual trials. Individual single-sound (A alone, B alone) trials align well with their corresponding 25-75% quantiles, by definition (Figure 1E1-2 and 1F1-2). But on dual-sound (AB) trials, for any given trial or time bin, some individual traces correspond well to one of the component sound's 25-75% quantiles, and on other trials or time bins they correspond well to the 25-75% quantiles of the other component sound. For the neuron in Figure1E, there are whole trials in which the activity match that evoked by sound "A" alone and others in which it better corresponds to that evoked by sound "B" alone. For the neuron in Figure 1F, the firing pattern on dual-sound trials appears to switch back and forth between the levels observed for sounds A and B as the trial unfolds. In short, for these two examples, the activity on dual-sound AB trials does not appear to occur at a consistent value intermediate between those evoked on single-sound A and B trials, but can fluctuate between those levels at a range of time scales. The following analyses will quantitatively assess these fluctuations across trials and/or time.

2.2. Statistical analysis of whole trial spike counts provides evidence consistent with multiplexing

If neurons alternate firing rates at the time scale of trials, as appears to be the case for the neuron in Figure 1E, then the spike counts from dual-sound responses should resemble a mix of spike counts from each of the component single-sound responses. We statistically tested this

hypothesis against other reasonable competing possibilities using the subset of triplets whose spike counts on single sound A and B trials could be well modeled by Poisson distributions with statistically different mean rates λ^A and λ^B (N=363, see methods for details).

The competing scenarios to describe the corresponding dual sound trials were:

(a) **Mixture**: The spike counts observed on individual trials are best described as draws from a mixture of Poi(λ^A) and Poi(λ^B) (Figure 2A). This is consistent with multiplexing across trials.

(b) **Intermediate**: A single Poisson distribution best describes the spike counts, and this Poisson has a rate λ^{AB} between λ^A and λ^B (Figure 2B). This is consistent with either multiplexing at faster, sub-trial time scales or with true averaging/normalization.

(c) **Outside**: Again, a single Poisson distribution best describes the spike counts, but the rate λ^{AB} is outside the range of λ^A and λ^B (i.e it is greater than both or less than both; Figure 2C). Summation-type responses would be captured under this heading, as would inhibitory interactions.

(d) **Single**: A single Poisson describes the dual-sound trial spike counts, but the rate λ^{AB} is equal to one of the single-sound rates λ^A or λ^B (Figure 2D). A winner- (or loser-)-take-all pattern would fit this category.

In summary, these four models capture the spectrum of possibilities at the whole-trial time scale. For each triplet, we computed the posterior probabilities of the four competing models as a (Bayesian) measure of goodness of fit to the data [29] (see Methods). In other words, for each triplet the model with the highest posterior probability can be interpreted as the best fit, and the posterior probability itself is a measure of the confidence in the model. We consider three levels of confidence from high to low (posterior probability >0.95, >0.50, and no threshold on the confidence level, that is, posterior probability > the minimum, 0.33, figure 2F).

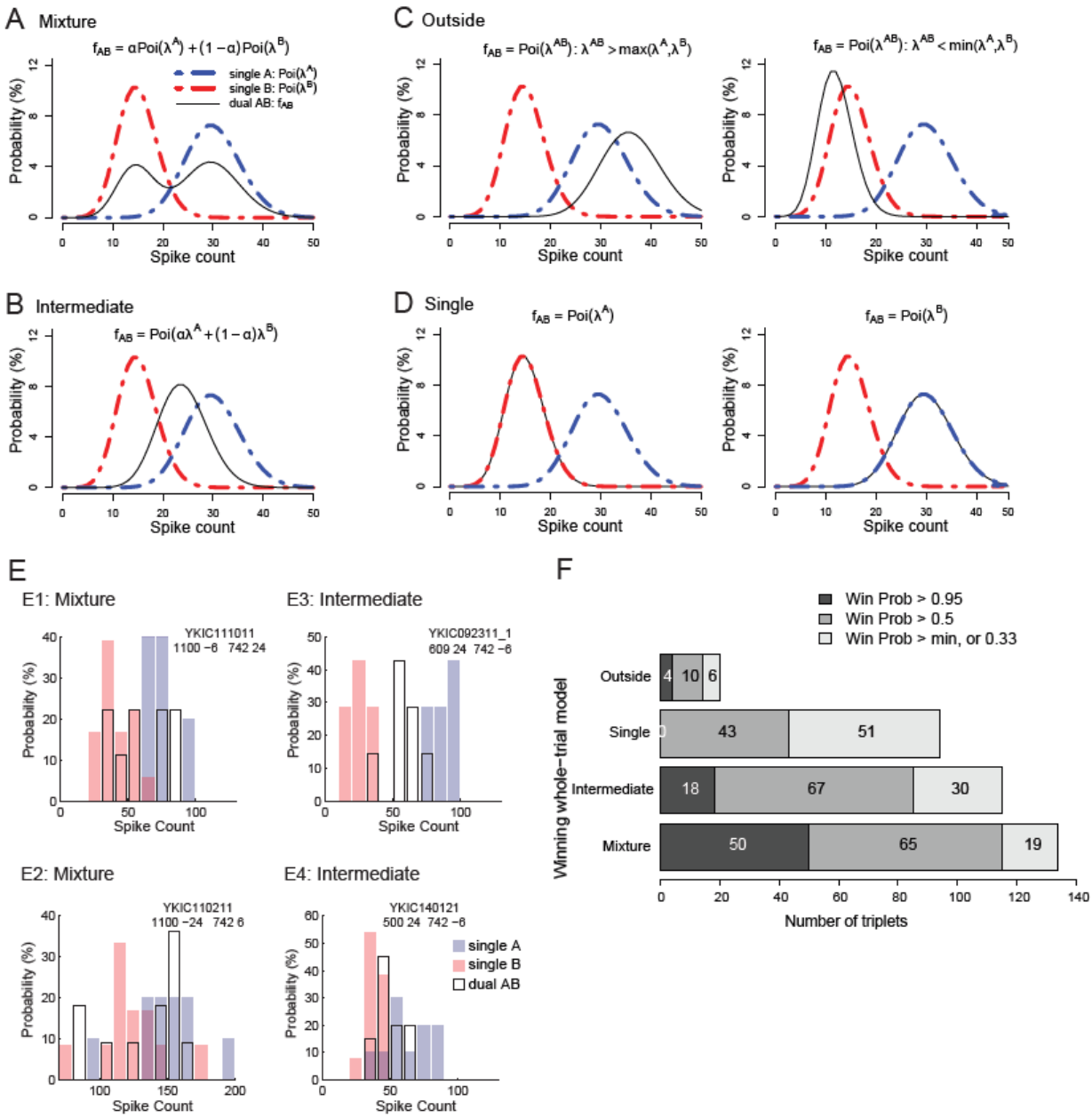


Figure 2. Whole-trial analysis. *Panels A to D show the four models that could describe the distribution of spike counts on individual dual-sound trials (0-600 or 0-1000 ms after sound onset, see Methods).* ***(A) mixture*** *of the Poisson distributions of spike counts for the component single-sound trials,* ***(B) intermediate*** *Poisson distribution, with rate between the rates of single sounds responses,* ***(C) outside****, Poisson distribution with rate larger or lower than the rates of single-sounds responses,* ***(D) single****, Poisson distribution with rate equal to one of the two single sound rates.* ***(E)*** *Four examples of spike count distributions for triplets classified as mixtures or*

intermediates. Red and blue shades indicate distributions of spike counts for single sounds; black outlines indicate distributions for dual sounds. The triplets in ***E1-2*** *were classified as "mixture" with winning probability >0.95 (E1 is the same triplet as in Figure 1E, E2 is the same triplet as in Figure 3B1). Triplets in* ***E3-4*** *were classified as "intermediate" with winning probability >0.95 (E3 is the same triplet as in Figure 1F and figure 3B2 and E4 is the same as in Figure 3B3).* ***(F)*** *Population results of the whole-trial analysis. Shading indicates the confidence level of the assignment of individual triplets to winning models.*

The mixture model had the strongest support of any of the models tested. For such triplets, the spike counts on dual-sound trials were better fit by a mixture of the single-sound Poisson distributions than by any single Poisson distribution (Figure 2 F, bar labeled "mixture"). These response patterns are potentially consistent with time division multiplexing at the level of individual trials; the neurons illustrated in Figures 1E (same as figure 2E1) and in figure 2E2 met these criteria. Of the 72 triplets in which one model had a winning probability >0.95, 50 or 69% were categorized as mixtures.

For the next largest category, the best fitting model involved a unique λ^{AB} between λ^{A} and λ^{B} (Figure 2F, bar labeled "intermediate"). These triplets are ambiguous: they could exhibit a true intermediate firing rate on the dual-sound trials, or they could simply show alternation at a time scale more rapid than individual trials (the neuron illustrated in Figure 1F (same as figure 2E3) and the one in figure 2E4 (same as figure 3B3) were classified as "intermediate"). Of the 72 triplets in which one model had a winning probability >0.95, 18 or ~25% were categorized this way.

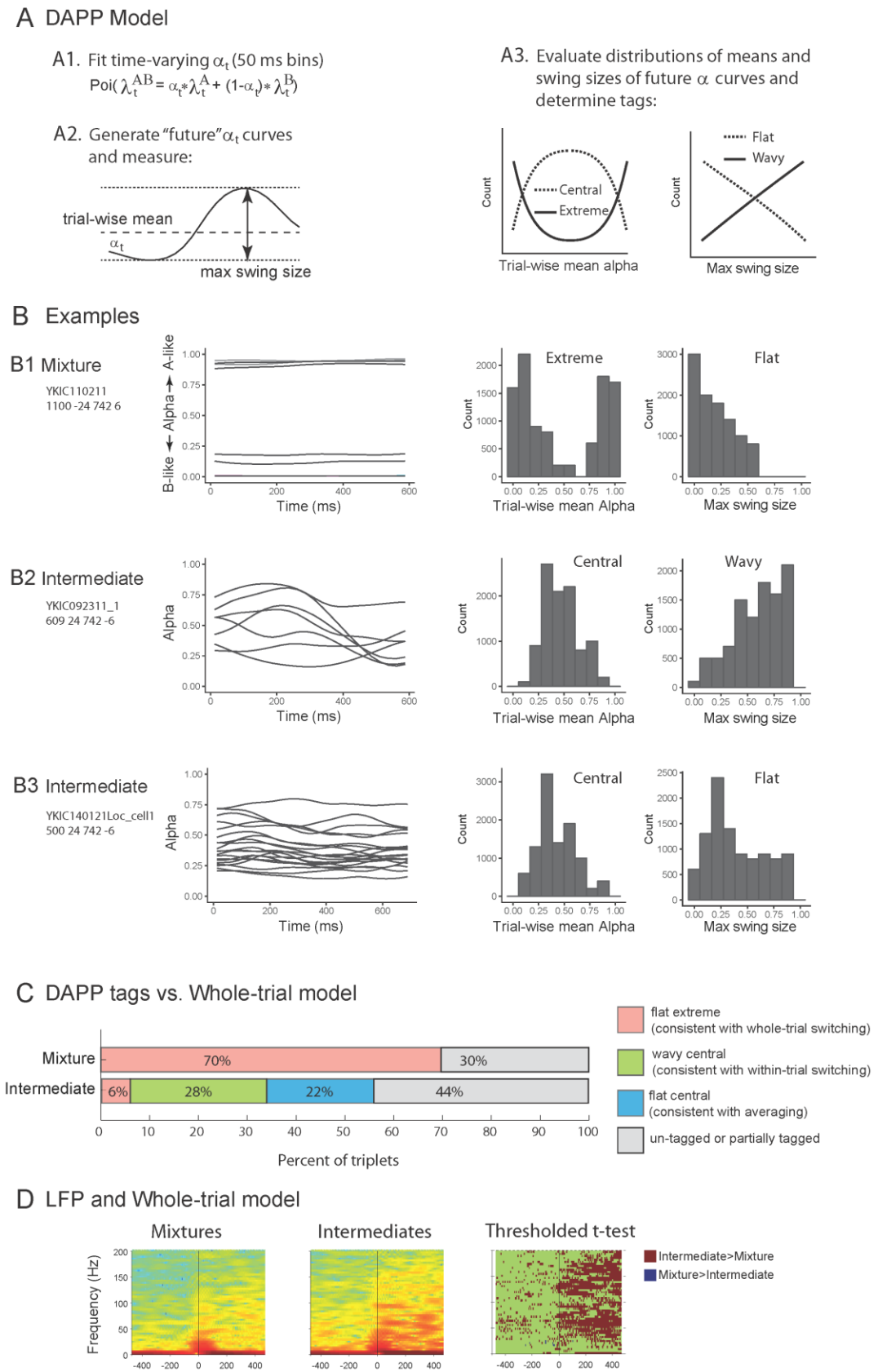
The remaining triplets were categorized as "single", or $\lambda^{AB} = \lambda^{A}$ or λ^{B} (a narrowly defined category that consequently did not produce any winning model probabilities >0.95) or "outside", λ^{AB} greater or less than both λ^{A} and λ^{B}. "Single" can be thought of as a winner-take-all response pattern. "Outside" may be consistent with a modest degree of summation in the neural population, particularly as λ^{AB} was generally greater than both λ^{A} and λ^{B} in this subgroup. The small prevalence of the single and outside groups compared to the mixture and intermediates is in agreement with the aggregated analysis of spiking data (Figures S2-3).

2.3 Statistical analysis of within-trial temporal patterns suggests presence of both fast multiplexing and stable averaging

We next evaluated whether firing patterns fluctuated or remained stable across time within a trial. In particular, might triplets categorized as "intermediate" in the whole trial analysis show evidence of fluctuating activity on a faster time scale?

We developed a novel statistical approach to study temporal patterns of the spike trains, the **Dynamic Admixture of Poisson Process (DAPP)** model. The computation is schematized in figure 3A (see also Methods). We focused on the same 363 triplets selected above. For each triplet, spike trains from individual single sound trials were assumed to be independent realizations of a nonhomogeneous Poisson process with unknown time-dependent firing rates ($\lambda^A(t)$ and $\lambda^B(t)$ for sounds A and B Dual sound trials were modelled as a weighted combination: $\lambda^{AB}(t) = \alpha(t)\lambda^A(t) + \big(1 - \alpha(t)\big)\lambda^B(t)$ (Figure 3A1).The weight function $\alpha(t)$ was unique to each dual sound trial and quantified the relative contribution of sound A on that trial at time *t*, while $1 - \alpha(t)$ quantified the complementary contribution of sound B. Thus, the dynamics of the $\alpha(t)$ function characterize the dynamics of each dual-sound trial (see examples in figure 3B).

For each selected triplet, we used Bayesian analysis to estimate the function valued model parameters (see Methods) and predict the $\alpha(t)$ curves that the corresponding cell was likely to produce on future dual sound AB trials (figure 3A2). Each predicted curve was summarized by two features: its time average over the response period of a given trial and its maximum swing size, that is, the difference between its highest peak and lowest trough on that trial. The triplet was then subjected to a two-way classification based on the distribution of these two features over the predicted curves (DAPP "tags", Figure 3A3). The triplet was categorized as "wavy" vs. "flat" depending on whether the distribution of the maximum swing size peaked at high or low values, and as "central" vs. "extreme" according to whether the distribution of the time average $\alpha(t)$ had a peak close to 0.5 or had one to two peaks at the extreme values of 0 and 1. In addition to this main classification scheme, triplets were subcategorized as exhibiting "symmetric" or "skewed" response patterns, indicating whether the $\alpha(t)$ curves reflected roughly equal contributions from the stimulus A and stimulus B response patterns or whether one or the other tended to dominate (figure S4).


A DAPP Model
A1. Fit time-varying α_t (50 ms bins)
Poi($\lambda_t^{AB} = \alpha_t * \lambda_t^A + (1-\alpha_t) * \lambda_t^B$)
A2. Generate "future" α_t curves and measure:
trial-wise mean
α_t
max swing size
A3. Evaluate distributions of means and swing sizes of future α curves and determine tags:
Count
Central
Extreme
Trial-wise mean alpha
Flat
Wavy
Max swing size
B Examples
B1 Mixture
YKIC110211
1100 -24 742 6
B-like ← Alpha → A-like
Time (ms)
Extreme
Flat
Trial-wise mean Alpha
Max swing size
B2 Intermediate
YKIC092311_1
609 24 742 -6
Alpha
Central
Wavy
B3 Intermediate
YKIC140121Loc_cell1
500 24 742 -6
C DAPP tags vs. Whole-trial model
Mixture
Intermediate
70%
30%
6%
28%
22%
44%
Percent of triplets
flat extreme (consistent with whole-trial switching)
wavy central (consistent with within-trial switching)
flat central (consistent with averaging)
un-tagged or partially tagged
D LFP and Whole-trial model
Mixtures
Intermediates
Thresholded t-test
Frequency (Hz)
Intermediate>Mixture
Mixture>Intermediate

238

Figure 3. Dynamic Admixture Point Process (DAPP) model: rationale, results, and relationship to local field potential (A) *The DAPP model fits smoothly time-varying weights (α and (1- α)) capturing the relative contribution of A- and B-like response distributions to each AB dual sound trial (A1). The dynamic tendencies of the α curves were then used to generate projected new α curves for hypothetical future draws from this distribution. The "waviness" and central tendencies were quantified by computing the max swing size and trial-wise mean for an individual trial draw from the distribution (panel A2). Low max swing sizes indicate "flat" curves and higher values indicate "wavy" ones (panel A3-right). Similarly, the distribution of trial-wise means could be bimodal ("extreme") or unimodal ("central") (panel A3-left).* ***(B)*** *Fit alphas for three example triplets (triplets in B1-2-3 are the same as in Figure 2E2-3-4 respectively) and the distribution of trial-wise means and max swing sizes for future draws from the alpha curve generator.* ***(C)*** *The pattern of DAPP results extended the whole-trial analysis results. Triplets categorized as "Mixtures" with a win probability > 0.95 tended to be tagged as "flat/extreme" (as example in B1). Triplets categorized as "Intermediates" fell in two different main groups, "wavy-central" (as example in B2) and "flat-central" (as example in B3). Information about the "skewed vs. symmetric" tag is not showed. See Supplementary Table 1 and Figures S4 and S5 for a complete listing of all the tag combinations and additional analyses.* ***(D)*** *Average local field potentials showed different patterns across the whole trial "mixture" vs. "intermediate" classification, with the "intermediate" group tending to show greater power at a range of oscillatory frequencies. The right panel shows a thresholded (p<0.05) depiction of whether these differences were statistically significant according to a two-tailed t-test for each time point and frequency combination.*

The DAPP tags confirmed and extended the results of the whole-trial analysis. Figure 3C shows the different distributions of DAPP tags for the triplets categorized as "intermediate" and "mixture" in the whole-trial analysis with winning probability >0.95 (see also figures S4-5).

"Mixture" triplets tended to be classified as "flat-extreme" (70%), that is, the dynamics of these dual sound responses were flat in time and matching either the responses to A (α ≈ 1) or B (α ≈ 0) on different trials. In addition, the distribution of the average α values tended to be either symmetric or unlabeled with regard to symmetry, a sanity check that excludes winner-take-all responses (these would be characterized by flat-extreme-skewed responses, Supplementary Table 1, Figure S4).

In contrast, “intermediate” triplets showed a combination of two types of labeling patterns relevant to our hypothesis. Some (22%) were classified as “flat-central” (and symmetric), indicating stable α between 0 and 1, that is, stable firing at roughly the average of the responses evoked by each sound separately (see example cell in figure 3B3). One triplet was classified as “flat-extreme” suggesting a stable α very close to either 0 or 1. Together these two firing patterns are consistent with some form of normalization occurring in this subpopulation. Other triplets (28%) were classified as “wavy-central” (and symmetric) indicating responses that fluctuated symmetrically around a central value (see example cell in figure 3B2).

In short, the DAPP analysis shows that the dynamics of the “mixture” responses were consistent with fluctuations at the level of whole trials like the cell in figure 3B1, whereas the dynamics of the “intermediate” responses classified as “wavy-central” are suggestive of a neural code that could “switch” relatively rapidly between responses consistent with two single stimuli.

2.4. Local field potential shows greater oscillatory energy associated with “intermediate” classification.

Consistent with this statistical evidence for activity fluctuations at the sub-trial timescale in the “intermediate” category, we also found that the local field potential (LFP) at such sites showed greater oscillatory activity. Figure 3D shows the average LFP power spectrum for dual trials of triplets categorized as “mixtures” vs. those categorized as “intermediates” and their statistical comparison (right panel, two-tailed t-test between the LFP power spectrum of dual trials classified as Intermediate and that of dual trials classified as mixtures, for each time point and frequency combination). The LFP for “intermediate” sites showed higher energy across a range of frequencies, including frequencies well above the 20 Hz (50 ms) frequency range that we were able to evaluate at the spike-count single unit level. This suggests there may be activity fluctuations occurring at frequencies higher than we were able to test with the current analysis method.

3. Activity fluctuations appear coordinated across the neural population and predict behavior

We next considered the question of whether and how activity fluctuations are coordinated across the neural population, in two ways: (1) by evaluating activity correlations across time within

trials between pairs of simultaneously recorded neurons, and (2) by evaluating whether the state of the local field potential prior to sound onset predicts between-trial fluctuations in activity [e.g. 30, 31]. Finally, we determined whether the observed fluctuations are of functional relevance for the execution of the task by testing the relation between the trial-by-trial fluctuation and which target the monkey looked at first on that trial.

3.1 Neural pairs show positive within trial correlations: To evaluate *within-trial* switching correlations, we evaluated the neuron-to-neuron correlation between how "A-like" vs. how "B-like" the responses were on a time bin by time bin basis on individual trials, in a total of 91 pairs of triplet conditions from 34 pairs of simultaneously recorded neurons (from among the 363 triplets used for the previous analyses). For each 50 ms bin of a dual-sound trial in a given triplet, we assigned a probability score between 0 and 1 that the spike count in the bin was drawn from the Poisson distribution with rate equaling the bin's sound A rate, and the complementary probability to the same being drawn from the Poisson distribution with rate equaling the bin's sound B rate (Figure 4A; see Methods: A vs. B assignment scores). We normalized these probabilities by converting them to Z-scores within a given time bin but across trials, to minimize the contribution of shared correlations due to stimulus responsiveness or changes in motivational state across time [32]. We then calculated the neuron-to-neuron correlation coefficients between the normalized assignment scores across time bins within each trial, i.e. one correlation coefficient value estimated per trial. This analysis is conceptually similar to conventional cross-correlation analysis of spike trains in neural pairs, but does not focus on precise timing of spikes or the relative latency between them [33, 34].

The observed correlations were generally positive, indicating that the activity was coordinated within the neural population. Figure 4 illustrates analysis of the dual-sound trials for a particular triplet in an example pair of neurons (A), and the distribution of the mean neuron-to-neuron correlations in the population for all the dual-sound conditions (B). The distribution of mean correlation coefficients was skewed positive (t-test, $p = 6.8 \times 10^{-6}$). Similar results were obtained when the raw spike counts were analyzed rather than the assignment scores (Supplementary Figure 6). This was the case even though we included triplets that were not categorized as showing "wavy" behavior in the DAPP analysis. It may be that coordinated activity fluctuations occur in more neurons than those that met our statistical criteria.

3.2 State of LFP prior to sound onset predicts between-trial fluctuations: To determine whether the state of the local field potential prior to sound onset predicts between-trial fluctuations in activity, we analyzed the LFP data recorded simultaneously with single unit spiking data. We combined data across triplets, creating two groups of trials based on whether the whole-trial spike count on a given dual-sound trial more closely resembled the responses evoked by sound A alone (where A is the contralateral sound) or sound B alone (see Methods: A vs. B assignment scores). Figure 4C shows the average LFP for the two groups of dual-sound trials. We quantified differences between these two groups with a t-test in the 600ms windows *before* and *after* sound onset (each trial contributed one mean LFP value in each time window). As expected, the LFP signals statistically differed after sound onset in these two trial groupings (red vs. blue traces, time period 0-600 ms, p-val = 1.0474×10^{-05}). But the LFP signals also differed *prior* to sound onset (p-val = 0.0064), suggesting that the state of activity in the local network surrounding an individual neuron at the time of sound onset is predictive of whether the neuron "encodes" the contra-lateral or the ipsi-lateral sound on that particular trial.

3.3 Neural fluctuations predict behavior: If fluctuations in neural activity are coordinated across the population, it follows that there should be a relationship between variability in neural activity and behavior. Accordingly, we investigated whether the activity on individual trials predicted whether the monkey would look first to sound "A" or sound "B" on that trial. We focused here on the whole-trial time scale as the most robust measure of activity fluctuations in our data. As noted in the Methods, we trained the monkeys on sequential sounds first and this training strategy tended to promote performing the task in a stereotyped sequence. Partway through neural data collection, we provided monkey Y with additional training on the non-sequential task, after which that monkey began displaying less stereotypical behavior and sometimes saccaded first to A and sometimes first to B for a given AB dual sound combination (see Figure 4D for example). We then analyzed recording sessions after this training (N=73 triplets) and we found that at both the whole trial and sub-trial time scales, the activity of individual neurons was predictive of what saccade sequence the monkey would choose on that particular trial. Specifically, the average dual sound AB assignment score for a given triplet was computed separately for trials in which the first saccade was toward A vs. toward B. The average scores statistically differed between the two

groups of dual-sound trials (t-test, pval = 5 x 10^-9, Figure 4E) and in the expected direction, with more A-like scores occurring on trials in which the monkey looked at A first. This relationship was also present when looking at finer, 50 ms bin time scales (Figure 4F).

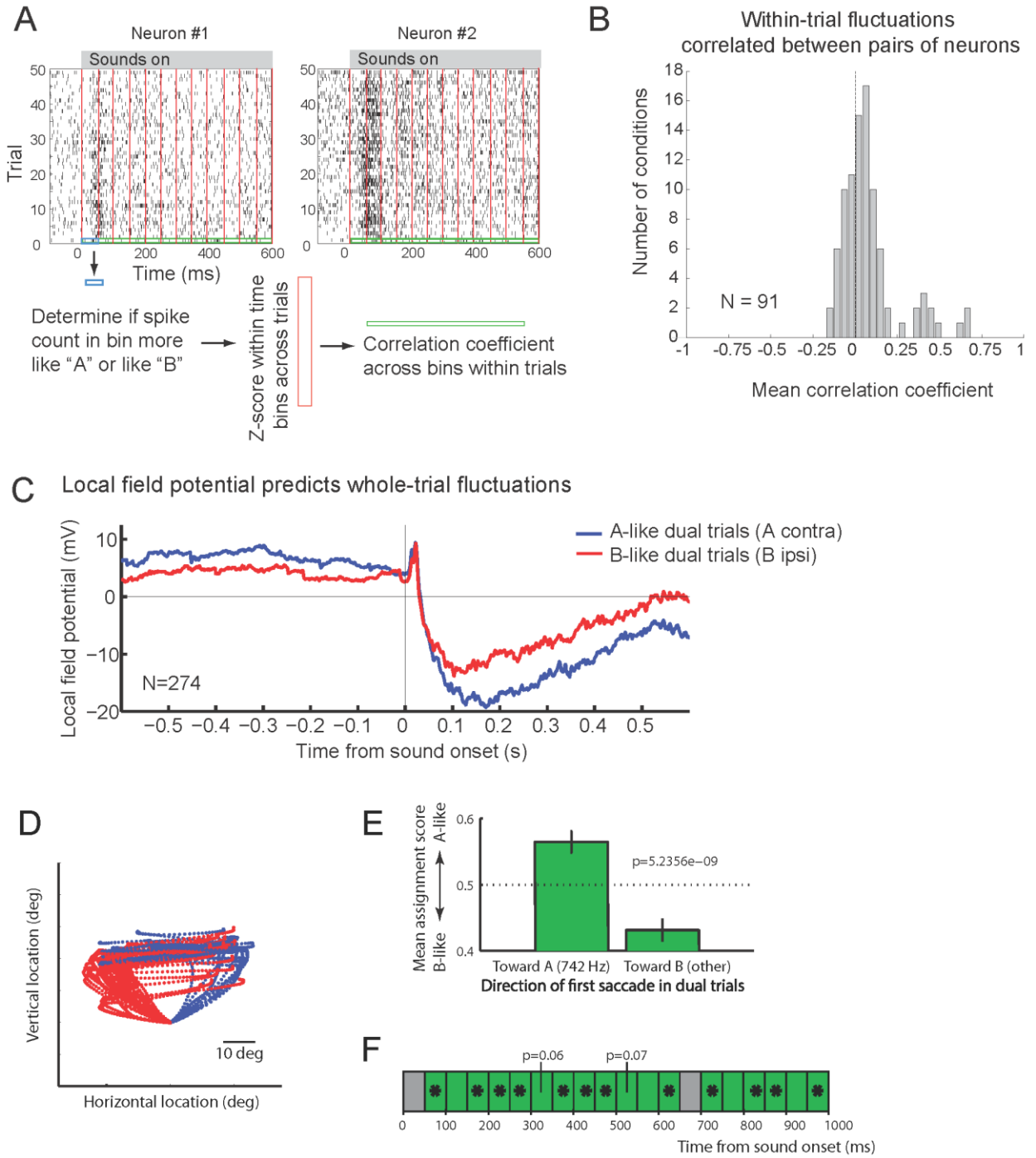


Figure 4. Within-trial fluctuations are correlated between pairs of neurons. Whole-trial fluctuations can be predicted by the state of the local field potential prior to sound onset and in turn predict which stimulus is reported first. ***(A-B)*** *Pairs of neurons recorded simultaneously tended to show positive correlations with each other.* ***(A)*** *Schematics of the analysis. Raster plots of two neurons recorded simultaneously; trials shown are for a particular set of dual-sound conditions. We evaluated the spike count in a given 50 ms time bin, trial, and member of the neuron pair for a given set of dual-sound conditions to determine if it was more similar to the spikes evoked during that bin on the corresponding sound "A" alone or "B" alone trials (blue box). We then converted these A vs. B assignment probabilities to a Z-score based on the mean and standard deviation of the assignment probabilities in that time bin on the other trials that involved the same stimulus conditions (red box). Finally, we computed the correlation coefficient between the set of Z score values for a given trial between the pair of simultaneously recorded neurons (green box).* ***(B)*** *Across the population of pairs of triplets recorded simultaneously, the distribution of mean correlation coefficients tended to be positive (t-test comparing the mean correlation coefficients to zero;* $p = 6.8 \times 10^{-6}$*).* ***(C)*** *To assess the relation between the LFP* prior *to sound onset and the spiking response* after *sound onset, we assigned each dual-sound trial to one of two groups, A-like or B-like, based on whether the spike count matched the response to single sound A or B (see Methods) and then compared the average LFP of the two groups. The average LFP for the 1902 A-like trials (A=contralateral sound) is in blue, the average LFP for the 1618 B-like trials is in red. The traces are significantly different in the 600ms before sound onset and in the 600ms after sound onset (t-test, p-val < 0.01).* ***(D-F)*** *The target of the first saccade on dual sound trials is predicted by the spike count during sound presentation.* ***(D)*** *Eye trajectories during dual-sound trials to the same pair of single sounds (one triplet). The traces are color-coded based on which of the two sounds the monkey looked at first in the response sequence. For clarity, all traces are aligned on a common starting position despite some variation in fixation accuracy.* ***(E)*** *The average assignment score of trials in which the monkey looked at sound A first is more A-like than that of trials in which the monkey looked at sound B first. Bars indicate SEM.* ***(F)*** *The relationship between assignment score and first saccade target was also evident at the scale of 50 ms bins (green = positive correlation; * indicate p<0.05 for t-test of assignment score on A-first vs. B-first trials).*

DISCUSSION

Our results show that the activity patterns of IC neurons fluctuate, and that these fluctuations may be consistent with encoding of multiple items in the same processing channels (i.e. the set of neural spike trains occurring in the IC). The time scale of these fluctuations ranges from the level of individual trials down to at least 50 ms bins within a trial. The fluctuations are positively correlated across pairs of neurons (at least, those recorded within the IC on a given side of the brain), are reflective of the state of local field potentials at the time of sound onset, and are predictive of the behavioral response to follow. These fluctuations may be a necessary element of the neural code, permitting the representation of more than one stimulus across time within a neural ensemble, and they may serve as a critical element of the coding of natural scenes replete with many simultaneous stimuli.

A strength of our statistical approach is that we do not simply assess unimodality/bimodality (e.g. [10]) but anchor our consideration of the dual stimulus response distributions to the observed distributions for single stimulus trials. However, there are several limitations to the present statistical approach. First, the analyses could only be conducted on a subset of the data, requiring a good fit of a Poisson distribution to the single-sound trials and adequate separation of the responses on those trials. For the moment, it is unknown whether any of the excluded data exhibit meaningful response fluctuations. In principle, the modeling approach can be extended to other types of response distributions which should reduce the amount of data that is excluded. Second, the full range of time scales at which fluctuations occur is still undetermined. We focused on the whole-trial timescale and on 50ms bins -- biologically plausible for both the task and the stimuli. However, any faster fluctuations would likely have been (erroneously) categorized as flat-central in the DAPP model. Third, our statistical approach based on the DAPP model involves a categorization step that summarizes the dominant features of a triplet. If a neuron sometimes behaves as a “flat-extreme” type and sometimes as a “wavy-central” type for a given triplet of conditions, it would likely be categorized as ambiguous. In other words, even though the DAPP model can pick up composite response patterns, the results we present ignore the existence of any such patterns.

The observed fluctuations have broad implications because they provide a novel account linking a number of other well-known aspects of brain function under a common explanation.

First, it is widely recognized that neural firing patterns are highly variable. This variability is often thought to reflect some fundamental inability of neurons to code information accurately. Here, we suggest that some of this variability may actually reflect interleaved periods of potentially quite accurate coding of different items. What else individual neurons may commonly be coding for in experiments involving presentation of only one stimulus at a time is not known, but possibilities include stimuli not deliberately presented by the experimenter, memories of previous stimuli, or mental imagery as suggested by the theory of embodied cognition [35]. In the present study, we were able to demonstrate signal in these fluctuations by virtue of statistical tests comparing each of the trial types in A-B-AB triplets, but it may be the case that fluctuations were occurring in the single stimulus trials as well. We could not test this because our analysis required having as benchmarks the response distributions corresponding to the potentially encoded items.

Second, as a concept, time division multiplexing provides insight into why limitations in certain types of cognition exist. Working memory capacity is limited; attention filters stimuli to allow in depth processing of a selected set of items. These limitations may stem from using the same population of neurons for each attended or remembered item. If this is the case, then the puzzle becomes why these limits are often greater than one. Time-division multiplexing suggests that cycling between different items across time allows evading what might otherwise be a one-item limit [2]. Here, we investigated only two time scales, 50 ms and whole trials. Future work will be needed to more fully explore the time scales on which this occurs and to tie the resulting information on duty cycle to perceptual capacity.

Third, brain oscillations are ubiquitous, have been linked specifically to attentional and memory processes [31, e.g. 36, see also 37], and have been suggested as indicating multiplexing [2-8]. Oscillations indicate that neural activity fluctuates, although they capture only the portion of such fluctuation that is coordinated across the underlying neural population and is regular in time. It remains to be determined to what degree oscillations in field potentials reflect the activity of neural circuits that control such temporal coordination in other neural populations vs. the activity of the neural circuits subject to the effects of such coordination. In a highly interconnected system such as the brain, both are likely to occur.

In the case of our particular experimental paradigm, several additional questions arise. How do signals related to different items come to be multiplexed? Are they later de-multiplexed? If so, how?

To some degree, sounds are multiplexed in the world. That is, the sound waves from multiple sources sum in the world and are never purely distinct from one another. The air pressure waves arriving at each ear reflect the combined contribution of all sound sources. However, if the IC's neural fluctuations were driven by the sound signals arriving at the ears, then individual neurons should always respond the same way on every trial, and they do not. Instead, it seems likely that the externally-multiplexed sound waves interact with neural circuit states at the time that the incoming signal arrives to govern how individual neurons respond on a moment by moment basis.

Where and how signals may be de-multiplexed critically depend on the nature of the representation to which a de-multiplexed output could be written. In barn owls, which have maps of auditory space, the coding bottleneck intrinsic to meter/rate coding does not occur, and two sounds produce two separate active populations [38-41]. Such distinct peaks suggest that the multiplexed-in-the-air signals have been de-multiplexed and segregated into two hills of activity.

In primates and several other mammals, neural representations of space employ meters (rate codes) rather than maps throughout the pathway from sound input to eye movement output, as far as is currently known [18-21, 42-46]. This is the case even at the level of the superior colliculus [47], an oculomotor structure which has a well-deserved reputation for mapping when activity is evoked by non-auditory stimuli [48, 49].

Given that different types of codes exist in different species, and given that coding format is not known in all the circumstances in which multiplexing might apply (e.g. attention, working memory), we developed two different models to illustrate a range of different de-multiplexing possibilities (Figure 5) based on the nature of the recipient representation. In the first (Figure 5A), a multiplexed signal in a meter is converted into two hills of activity in a map, using a basic architecture involving graded thresholds and inhibitory interneurons suggested previously [50]. Adding an integration mechanism such as local positive feedback loops would then serve to latch activity "on" at the appropriate locations in the map, producing a more sustained firing pattern. No clock signal is necessary for this model.

In the second model (Figure 5B), there are multiple output channels, each capable of encoding one item. An oscillating circuit that knows about the timing of the input gates signals to each output channel at the appropriate moments. As in the first model, a local positive feedback mechanism acts to sustain the activity during the gaps in the input. This model thus retains the

efficient coding format of a meter but requires a controlling signal with knowledge of when to latch input flow through to each output channel. In our data, fluctuations were coordinates across pairs of simultaneously recorded neurons, in agreement with this model. It is possible that within-trial fluctuating units lie at the input stage of such a circuit, and that between-trial fluctuating units actually lie at the output stage. A given unit might be allocated to either the "A" or the "B" pools based on state of the network (as detected by the LFP measurements) on different trials.

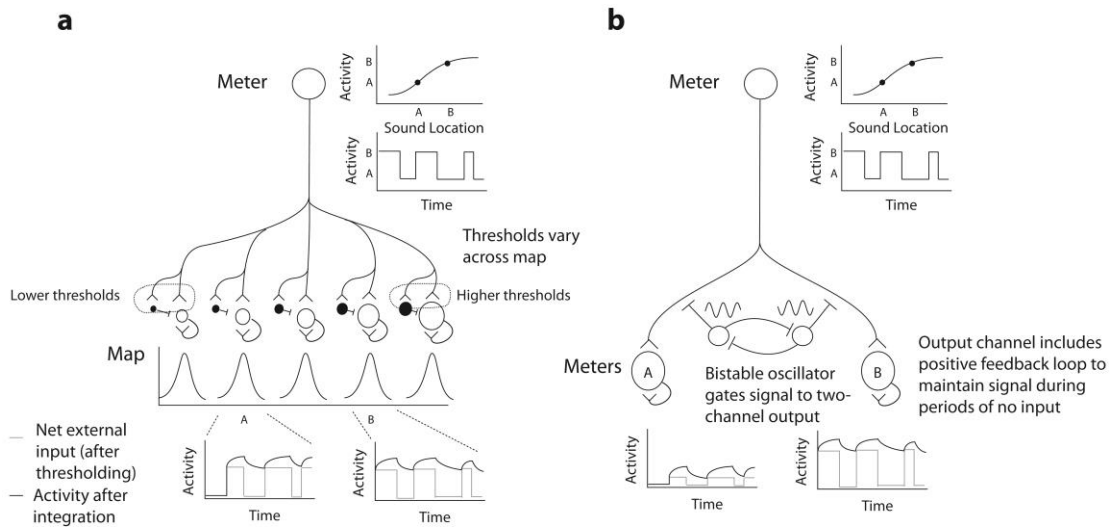


Figure 5. Two possible mechanisms for de-multiplexing a fluctuating signal. *A clock signal that knows about coding transitions is not necessarily needed if signals are read out into a map* ***(a)****, but is required if signals are retained in a meter or rate-coded format* ***(b)****.*

An important unresolved question posed by our study is whether multiplexing may be a general mechanism that is commonly at play to enhance the total processing power of the brain. The statistical tools developed here can be applied to any "triplet" data. Additional studies with both single stimulus conditions, to define the distributions of signals, and dual stimulus conditions, to evaluate fluctuations between membership in those distributions, will be important for delineating the extent of this phenomenon. Digging under the hood of the time-and-trial pooled activity to look at activity patterns on a moment by moment basis will be essential to advancing our understanding of how the brain operates dynamically to maximize its processing power.

REFERENCES

1. Meister, M., *Multineuronal codes in retinal signaling.* Proc Natl Acad Sci U S A, 1996. **93**(2): p. 609-14.
2. Lisman, J.E. and M.A. Idiart, *Storage of 7 +/- 2 short-term memories in oscillatory subcycles.* Science, 1995. **267**(5203): p. 1512-5.
3. Hoppensteadt, F.C. and E.M. Izhikevich, *Thalamo-cortical interactions modeled by weakly connected oscillators: could the brain use FM radio principles?* Biosystems, 1998. **48**(1-3): p. 85-94.
4. Akam, T. and D.M. Kullmann, *Oscillatory multiplexing of population codes for selective communication in the mammalian brain.* Nat Rev Neurosci, 2014. **15**(2): p. 111-22.
5. Tallon-Baudry, C., *The roles of gamma-band oscillatory synchrony in human visual cognition.* Front Biosci (Landmark Ed), 2009. **14**: p. 321-32.
6. Lisman, J.E. and O. Jensen, *The theta-gamma neural code.* Neuron, 2013. **77**(6): p. 1002-16.
7. Siegel, M., M.R. Warden, and E.K. Miller, *Phase-dependent neuronal coding of objects in short-term memory.* Proc Natl Acad Sci U S A, 2009. **106**(50): p. 21341-6.
8. McLelland, D. and R. VanRullen, *Theta-Gamma Coding Meets Communication-through-Coherence: Neuronal Oscillatory Multiplexing Theories Reconciled.* PLoS Comput Biol, 2016. **12**(10): p. e1005162.
9. Jezek, K., et al., *Theta-paced flickering between place-cell maps in the hippocampus.* Nature, 2011. **478**(7368): p. 246-9.
10. Li, K., et al., *Neurons in Primate Visual Cortex Alternate between Responses to Multiple Stimuli in Their Receptive Field.* Front Comput Neurosci, 2016. **10**: p. 141.
11. Panzeri, S., et al., *Sensory neural codes using multiplexed temporal scales.* Trends Neurosci, 2010. **33**(3): p. 111-20.
12. Adams, J.C., *Ascending projections to the inferior colliculus.* J Comp Neurol, 1979. **183**(3): p. 519-38.
13. Moore, R.Y. and J.M. Goldberg, *Ascending projections of the inferior colliculus in the cat.* J Comp Neurol, 1963. **121**: p. 109-35.
14. Aitkin, L.M. and S.C. Phillips, *Is the inferior colliculus an obligatory relay in the cat auditory system?* Neuroscience Letters, 1984. **44**(3): p. 259-64.
15. Jenkins, W.M. and R.B. Masterton, *Sound localization: effects of unilateral lesions in central auditory system.* J Neurophysiol, 1982. **47**(6): p. 987-1016.
16. Kelly, J.B. and G.L. Kavanagh, *Sound localization after unilateral lesions of inferior colliculus in the ferret (Mustela putorius).* Journal of Neurophysiology, 1994. **71**(3): p. 1078-1087.
17. Litovsky, R.Y., B.J. Fligor, and M.J. Tramo, *Functional role of the human inferior colliculus in binaural hearing.* Hear Res, 2002. **165**(1-2): p. 177-88.
18. Groh, J.M., K.A. Kelly, and A.M. Underhill, *A monotonic code for sound azimuth in primate inferior colliculus.* Journal of Cognitive Neuroscience, 2003. **15**: p. 1217-1231.
19. McAlpine, D. and B. Grothe, *Sound localization and delay lines--do mammals fit the model?* Trends Neurosci, 2003. **26**(7): p. 347-50.
20. Grothe, B., M. Pecka, and D. McAlpine, *Mechanisms of sound localization in mammals.* Physiol Rev, 2010. **90**(3): p. 983-1012.
21. Groh, J.M., *Moving with maps and meters*, in *Making space: How the brain knows where things are*. 2014, Harvard University Press. p. 143-159.
22. Perrott, D.R., *Discrimination of the spatial distribution of concurrently active sound sources: some experiments with stereophonic arrays.* J Acoust Soc Am, 1984. **76**(6): p. 1704-12.
23. Perrott, D.R., *Concurrent minimum audible angle: a re-examination of the concept of auditory spatial acuity.* J Acoust Soc Am, 1984. **75**(4): p. 1201-6.
24. Best, V., A. van Schaik, and S. Carlile, *Separation of concurrent broadband sound sources by human listeners.* J Acoust Soc Am, 2004. **115**(1): p. 324-36.
25. Carandini, M., D.J. Heeger, and J.A. Movshon, *Linearity and normalization in simple cells of the macaque primary visual cortex.* J Neurosci, 1997. **17**(21): p. 8621-44.
26. Movshon, J.A., I.D. Thompson, and D.J. Tolhurst, *Spatial summation in the receptive fields of simple cells in the cat's striate cortex.* J Physiol, 1978. **283**: p. 53-77.
27. Cash, S. and R. Yuste, *Linear summation of excitatory inputs by CA1 pyramidal neurons.* Neuron, 1999. **22**(2): p. 383-94.
28. Alvarado, J.C., et al., *Multisensory versus unisensory integration: contrasting modes in the superior colliculus.* J Neurophysiol, 2007. **97**(5): p. 3193-205.

29. Berger, J.O. and L.R. Pericchi, *The Intrinsic Bayes Factor for Model Selection and Prediction.* Journal of the American Statistical Association, 1996. **91**(433): p. 109-122.
30. Haslinger, R., et al., *Analysis of LFP phase predicts sensory response of barrel cortex.* J Neurophysiol, 2006. **96**(3): p. 1658-63.
31. Vanrullen, R., et al., *Ongoing EEG Phase as a Trial-by-Trial Predictor of Perceptual and Attentional Variability.* Front Psychol, 2011. **2**: p. 60.
32. Metzger, R.R., et al., *Effects of reward and behavioral context on neural activity in the primate inferior colliculus.* Journal of Neuroscience, 2006. **26**(28): p. 7468-76.
33. Gerstein, G.L., D.H. Perkel, and K.N. Subramian, *Identification of functionally related neural assemblies.* Brain Res., 1978. **140**: p. 43-62.
34. Alonso, J.M., W.M. Usrey, and R.C. Reid, *Precisely correlated firing in cells of the lateral geniculate nucleus.* Nature, 1996. **383**(6603): p. 815-9.
35. Barsalou, L.W., *Grounded cognition.* Annu Rev Psychol, 2008. **59**: p. 617-45.
36. Raghavachari, S., et al., *Gating of human theta oscillations by a working memory task.* J Neurosci, 2001. **21**(9): p. 3175-83.
37. Engel, T.A., et al., *Selective modulation of cortical state during spatial attention.* Science, 2016. **354**(6316): p. 1140-1144.
38. Baxter, C.S., B.S. Nelson, and T.T. Takahashi, *The role of envelope shape in the localization of multiple sound sources and echoes in the barn owl.* J Neurophysiol, 2013. **109**(4): p. 924-31.
39. Takahashi, T.T., et al., *Object localization in cluttered acoustical environments.* Biol Cybern, 2008. **98**(6): p. 579-86.
40. Keller, C.H. and T.T. Takahashi, *Localization and identification of concurrent sounds in the owl's auditory space map.* J Neurosci, 2005. **25**(45): p. 10446-61.
41. Takahashi, T.T. and C.H. Keller, *Representation of multiple sound sources in the owl's auditory space map.* Journal of Neuroscience., 1994. **14**(8): p. 4780-4793.
42. Salminen, N.H., et al., *A population rate code of auditory space in the human cortex.* PLoS One, 2009. **4**(10): p. e7600.
43. Salminen, N.H., et al., *The neural code for interaural time difference in human auditory cortex.* J Acoust Soc Am, 2010. **127**(2): p. EL60-5.
44. Werner-Reiss, U. and J.M. Groh, *A rate code for sound azimuth in monkey auditory cortex: implications for human neuroimaging studies.* Journal of Neuroscience, 2008. **28**(14): p. 3747-58.
45. Woods, T.M., et al., *Effects of stimulus azimuth and intensity on the single-neuron activity in the auditory cortex of the alert macaque monkey.* J Neurophysiol, 2006. **96**(6): p. 3323-37.
46. Ashida, G. and C.E. Carr, *Sound localization: Jeffress and beyond.* Curr Opin Neurobiol, 2011. **21**(5): p. 745-51.
47. Lee, J. and J.M. Groh, *Different stimuli, different spatial codes: a visual map and an auditory rate code for oculomotor space in the primate superior colliculus.* PLoS One, 2014. **9**(1): p. e85017.
48. Robinson, D.A., *Eye movements evoked by collicular stimulation in the alert monkey.* Vision Res, 1972. **12**: p. 1795-1807.
49. Mohler, C.W. and R.H. Wurtz, *Organization of monkey superior colliculus: intermediate layer cells discharging before eye movements.* Journal of Neurophysiology, 1976. **39**(4): p. 722-44.
50. Groh, J.M. and D.L. Sparks, *Two models for transforming auditory signals from head-centered to eye-centered coordinates.* Biological Cybernetics, 1992. **67**(4): p. 291-302.
51. Caruso, V.C., et al., *Is multiplexing a general strategy for encoding multiple items in the brain? Evidence from a visual cortical face area and a subcortical auditory area.* Soc Neurosci. Abstr., 2015.
52. Bulkin, D.A. and J.M. Groh, *Systematic mapping of the monkey inferior colliculus reveals enhanced low frequency sound representation.* Journal of Neurophysiology, 2011. **105**(4): p. 1785-97.
53. Barton, E.J. and D.L. Sparks, *Saccades to remembered targets exhibit enhanced orbital position effects in monkeys.* Vision Res, 2001. **41**(18): p. 2393-406.
54. Metzger, R.R., et al., *Auditory saccades from different eye positions in the monkey: implications for coordinate transformations.* Journal of Neurophysiology, 2004. **92**(4): p. 2622-7.
55. Berger, J., *The case for objective Bayesian analysis.* Bayesian Analysis, 2006. **1.3**: p. 385-402.
56. Berger, J.O. and L.R. Pericchi, *The intrinsic Bayes factor for model selection and prediction.* Journal of the American Statistical Association, 1996. **91.433**: p. 109-122.
57. Escobar, M.D. and M. West, *Bayesian Density-Estimation and Inference Using Mixtures.* Journal of the American Statistical Association, 1995. **90**(430): p. 577-588.

SUPPLEMENTARY MATERIALS:

MATERIALS AND METHODS

General procedures

All procedures conformed to the guidelines of the National Institutes of Health (NIH Pub. No. 86-23, Revised 1985) and were approved by the Institutional Animal Care and Use Committee of the university. Two adult rhesus monkeys (Macaca mulatta) participated (monkey P, and monkey Y, both female). Under general anesthesia and in sterile surgery we first implanted a head post holder to restrain the head and a scleral search coil to track eye movements (Robinson 1963; Judge et al. 1980). After recovery with suitable analgesics and veterinary care, we trained the monkeys in the experimental task. In a second surgery, we implanted a recording cylinder (2 cm diameter) over the right (monkey Y) or left (monkey Y, P) IC respectively. We determined the location of the cylinder with stereotactic coordinates and verified it with MRI scans [e.g. 52].

Behavioral task and training

<u>Sound localization task:</u>

The monkeys performed a single- or dual-sound localization task (Figure 2A) by making saccades toward one or two simultaneously-presented auditory targets with one or two saccades as appropriate. All sound targets were located in front of the monkey at eye level; the horizontal location, frequency and intensity were varied pseudorandomly as described below (Recording Procedures). Each trial began with 600-700ms of fixation of a visual stimulus (light emitting-diode, LED, located straight ahead and 10-14° below the speakers). During fixation we presented one sound (single-sound trials) or two simultaneous sounds (dual-sound trials). After a fixation time of either 600-800 (Data Set I, some of Data Set II) or 1000-1100 (remainder of Data Set II), the fixation light was extinguished and the monkey was required to make a single saccade on single-sound trials or a sequence of two saccades (in either order) on dual-sound trials. Trials were considered correct if each saccade was directed within 10-17.5 degrees horizontally and 20-40 degrees vertically of a target [due to vertical inaccuracies in localizing non-visual targets in primates, 53] and if the gaze was maintained on the final target for 100 – 200 ms. On correct trials monkeys were rewarded with juice drops.

Training

Training was accomplished in three stages. We initially trained the monkeys to report the location of single visual targets by saccading to them. We then introduced single auditory targets. As these were novel and unexpected in the silent experimental booth monkeys readily saccaded to them [54]. To help the monkeys calibrate their auditory saccades, a visual feedback was added on trials where the auditory saccade was not initiated correctly within 700 ms. The feedback was presented only at the most peripheral target locations (+/- 24 degree) and only for a few initial days of training. Finally, we trained monkey to localize dual-sound targets. Initially we presented the two sounds sequentially in a specific order, then we gradually reduced the temporal gap between them until the sounds were simultaneous.

In the final version of the task, monkeys were allowed to look at the targets in either order, as noted above. However, due to the initial training with sequential sounds, they retained stereotyped patterns of saccades in which they tended to look first to whichever sound location had been presented first during the sequential and partial overlap stages of training. Monkey P was trained with more central target locations (e.g. -6 or 6 degree targets) initially occurring first and more peripheral targets (e.g. -24 or 24 degree targets) occurring second, and monkey Y was trained with sounds initially occurring in the opposite sequence. Midway through neural data collection, we provided additional training to monkey Y to encourage free choice of which sound to look at first. This allowed us to investigate the relationship between each behavioral response and the neural representation at that moment.

Recording procedure and strategy

General procedure

Recordings were made with one or two tungsten electrodes (FHC, impedance between 1 and 3 MΩ at 1 kHz). Each electrode was lodged in a stainless-steel guide tube (manually advanced through the dura) and controlled independently with an oil hydraulic pulse micropositioner (Narishige International USA, Inc. and NAN INSTRUMENTS LTD, Israel). First, we localized the IC (and isolated single neurons) while the monkey listened passively to sounds of different frequencies. We then collected single unit spiking activity and local field potential while the monkey performed the single- and dual-sound localization tasks. We used a Multichannel

Acquisition Processor (MAP system, Plexon Inc., Dallas, TX) and Sort Client software. The single unit spiking activity was filtered between 150 Hz and 8 kHz and sampled at 20 kHz, while the LFP signal was filtered between 0.7 and 300 Hz and sampled at either 20kHz or 1kHz (see Local Field Potential). Data were collected as long as the neurons were well isolated and the monkey performed the tasks

Neural signals were recorded primarily from two functionally-defined subregions of the IC, the low frequency area and the tonotopic area [52]. Neurons in the low frequency tuned area generally respond best to low frequencies and there is little heterogeneity in tuning, whereas neurons recorded in the tonotopic area had best frequencies that could be either low or high depending on the position of the recording electrode.

Data Sets and Auditory Stimuli: Locations, Frequencies, and Levels

The spiking activity of 166 single neurons was recorded in two datasets involving the same task but differing in which sound levels and frequencies were included. A total of 68 of these neurons were recorded as pairs from separate electrodes positioned in the IC on the same side of the brain at a minimum spatial separation of 2 mm. Local field potentials (LFP) were also recorded from 87 of these recording sites.

In both datasets, the sounds consisted of bandpass noise with a bandwidth of +/- 200 Hz. On dual-sound trials, the sounds were delivered from pairs of locations (24 degrees and -6 degrees), and (-24 and +6 degrees) i.e. 30 degrees apart. The two sounds differed in frequency, with one of the two sounds having a 742 Hz center frequency and the other differing by at least 0.285 octaves or multiples of this distance. Single-sound trials involved the same set of locations and frequencies as on dual-sound trials, but with only a single sound presented at a time. All sounds were "**frozen**" within an individual session; that is, all trials with a given set of auditory parameters involved the same time series signal delivered to the relevant speaker.

In data set I (N=98 neurons), the sounds presented on dual-sound trials were 742 Hz and a sound from the set (500, 609, 903, 1100 Hz); these frequencies were ±0.285 octave or ± 0.57 octaves above or below 742 Hz, or ±3.4 and 6.8 semitones. Combining two sounds will produce a combination that is louder than either component. Sound levels were therefore calibrated to provide two sets of conditions: dual sounds for which the component sounds involve the *same signals* to the audio speakers as on single-sound trials, producing a *louder* dual sound, and dual

sounds for which the level of the component sounds was *reduced* so that the overall *loudness* was the *same* on dual as on single trials. The levels used for the components were 51 and 55 dB, producing sound levels of minimum 55 or maximum 60 dB on dual-sound trials. The same-signal comparison involved using the 55 dB component levels, singly and on dual-sound trials. The same-loudness comparison involved using the 55 dB levels on single-sound trials and the 51 dB levels for the components of dual-sound trials. Calibrations were performed using a microphone (Bruel and Kjaer 2237 sound level meter) placed at the position normally occupied by the animal's head.

Because results did not differ substantively when comparisons were made between same-signal and same-loudness conditions (Figure S2 vs. Figure S3), we pooled across sound levels for subsequent analyses, and we dispensed with the multiple sound levels for data set II (monkey Y only, N=68 neurons), using either 50 or 55 dB levels for all components. We also incorporated additional sound frequencies, [1340 1632 1988 Hz], to improve the odds that responses to each of the component sounds differed significantly. Again, one of the two sounds on dual-sound trials was 742 Hz; the other sound frequency was either from the original list of [500 609 903 1100] or from the new frequencies. Most of the neurons in this data set were tested with [500 742 1632].

<u>Cell Inclusion/Exclusion criteria and trial counts</u>

The N=166 neurons (N=98 from Data Set I and N=68 from Data Set II) included for analysis were drawn from a larger set of 325 neurons. Neurons were excluded from analysis if the neuron proved unresponsive to sound (Student's t-test, spike counts during the 600 ms after sound onset compared to the same period immediately prior to sound onset, one-tailed, $p>0.05$), or if there were too few correct trials (mininum of five correct trials for each of the components [A, B, and AB trials] that formed a given "triplet" of conditions or if there were technical problems during data collection (e.g. problems with random interleaving of conditions or with computer crashes). The average number of correct trials for a given set of stimulus conditions in the included dataset (N=166) was 10.5 trials. The total number of included triplets was 1484.

Data Analysis

All analyses concerned correctly performed trials.

Analysis of activity pooled across time and/or trials: Summation and Averaging

To evaluate IC activity using conventional analysis methods that pool across time and/or across trials, we counted action potentials during two standard time periods. The baseline period (*Base*) was the 600ms period before target onset, and the sensory-related target period (*Resp*) was the 600ms period after target onset (i.e. ending before, or at the time of, the offset of the fixation light, Figure 1B).

Summation/Averaging Indices: We quantified the activity on dual-sound trials in comparison to the sum and the average of the activity on single-sound trials, expressed in units of standard deviation (Z-scores), similar to a method used by [28]. Specifically, we calculated,

$$PredictedSum_{A,B,} = mean(Resp_A) + mean(Resp_B) - mean(Base_{A,B}) \quad (1)$$

and

$$PredictedAvg_{A,B,} = (mean(Resp_A) + mean(Resp_B))/2 \quad (2)$$

where $Resp_A$ and $Resp_B$ were the number of spikes of a given neuron for a given set of single-sound conditions *A* and *B* (location, frequency, and intensity) that matched the component sounds of the dual-sound trials being evaluated. As the "response" may actually include a contribution from spontaneous baseline activity, we subtracted the mean of the baseline activity for the single sounds ($Base_{A,B}$). Without this subtraction, the predicted sum would be artificially high because two "copies" of baseline activity are included under the guise of the response activity.

The Z scores for the dual-sound trials were computed by subtracting these predicted values from the mean of the dual-sound trials (mean($Resp_{AB}$)) and dividing by the mean of the standard deviations of the responses on single-sound trials:

$$Zsum_{AB} = \frac{mean(Resp_{AB}) - PredictedSum_{A,B}}{mean(std(Resp_A), std(Resp_B))} \quad (3)$$

and

$$ZAvg_{AB} = \frac{mean(Resp_{AB}) - PredictedAvg_{A,B}}{mean(std(Resp_A), std(Resp_B))} \quad (4)$$

If the dual response was within +/- 1.96 of the predicted sum or predicted average, we could say the actual dual response was within the 95% confidence intervals for addition or averaging of two single responses, respectively.

Analyses of fluctuations in neural firing across and within-trials, and inclusion criteria

Our statistical tests for fluctuations in neural firing were conducted on triplets, or related sets of single and dual-sound trials (A, B, AB trials). To evaluate whether neural activity fluctuates *across* trials in a fashion consistent with switching between firing patterns representing the component sounds, we evaluated the Poisson characteristics of the spike trains on matching dual and single-sound trials (triplets: AB, A and B). Spike train data from each trial was summarized by the total spike count between 0-600ms or 0-1000 ms from sound onset (i.e. whatever the minimum duration of the overlap between fixation and sound presentation was for that recorded neuron, see section Events of Task). We modeled the distribution of spike counts in response to single sounds A and B as Poisson distributions with unknown rates λ^A, denoted $Poi(\lambda^A)$, and λ^B, denoted $Poi(\lambda^B)$. Four hypotheses were considered for the distribution of sound AB spike counts:

1. a mixture distribution $\alpha \cdot Poi(\lambda^A) + (1 - \alpha) \cdot Poi(\lambda^B)$ with an unknown mixing weight α ("mixture")
2. a single $Poi(\lambda^{AB})$ with some λ^{AB} in between λ^A and λ^B ("intermediate")
3. a single $Poi(\lambda^{AB})$ where λ^{AB} is either larger or smaller than both λ^A and λ^B ("outside")
4. a single $Poi(\lambda^{AB})$ where λ^{AB} exactly equals one of λ^A and λ^B ("single")

Relative plausibility of these competing hypotheses was assessed by computing their posterior probabilities with equal prior weights (1/4) assigned to the models, and with default Jeffreys' prior [55] on model specific Poisson rate parameters, and a uniform prior on the mixing weight parameter α. Posterior model probabilities were calculated by computation of relevant intrinsic Bayes factors [56].

Triplets were excluded if either of the following applied: 1) the Poisson assumption on A and B trial counts was not supported by data; or 2) λ^A and λ^B were not well separated. To test the Poisson assumption on single-sound trials A and B of a given triplet, we used an approximate chi-square goodness of fit test with Monte Carlo p-value calculation. For each sound type, we estimated the Poisson rate by averaging counts across trials. Equal probability bins were constructed from the quantiles of this estimated Poisson distribution, with number of bins determined by expected count of 5 trials in each bin or at least 3 bins -- whichever resulted in more bins. A lack-of-fit statistic was calculated by summing across all bins the ratio of the square of the difference between observed and expected bin counts to the expected bin count. Ten thousand Monte Carlo samples of Poisson counts, with sample size given by the observed number of trials, were generated from the estimated

Poisson distribution and the lack-of-fit statistic was calculated from each one of these samples. P-value was calculated as the proportion of these Monte Carlo samples with lack-of-fit statistic larger than the statistic value from the observed data. Poisson assumption was considered invalid if the resulting Monte Carlo p-value < 0.1.

For triplets with valid Poisson assumption on sound A and B spike counts, we tested for substantial separation between λ^A and λ^B, by calculating the intrinsic Bayes factor of the model $\lambda^A \neq \lambda^B$ against $\lambda^A = \lambda^B$ with the non-informative Jeffreys' prior on the λ parameters: λ^A, λ^B or their common value. The triplet was considered well separated in its single sounds if the logarithm of the intrinsic Bayes factor equaled 3 or more, which is the same as saying the posterior probability of $\lambda^B \neq \lambda^A$ exceeded 95% when a-priori the two models were given 50-50 chance.

It should be noted that the sensitivity/specificity of detection were not equal across the four competing hypotheses. Because the "single" hypothesis is a limiting case of each of the other three hypotheses, the method's sensitivity to this response pattern is lower than for the other three possible outcomes. This was verified on simulated data for which the truth could be known; simulated "mixtures", "intermediates", and "outsides" were commonly correctly categorized with >95% confidence, whereas "singles" were correctly categorized but at a lower level of confidence.

Dynamic Admixture Point Process Model

To evaluate whether neural activity fluctuates *within* trials, we developed a novel analysis method we call a Dynamic Admixture Point Process model (DAPP) which characterized the dynamics of spike trains on dual sound trials as an admixture of those occurring on single sound trials. The analysis was carried out by binning time into moderately small time intervals. Given a predetermined bin-width $w = T/C$ for some integer C, we divided the response period into contiguous time intervals $I_1 = [0;w)$; $I_2 = [w; 2w)\ldots$ $I_C = [(C-1)w,Cw)$ and reduced each trial to a C-dimensional vector of bin counts $(X^e_{j1},\ldots,X^e_{jC})$ for $e \in \{A;B;AB\}$ and $j = 1,\ldots, n_e$. Mathematically, $X^e_{jC} = N^e_j(Ic)$. The results reported in Section 2 were based on $w = 50$ (with time measured in ms and T = 600 or 1000), but we also repeated the analyses with $w = 25$ and noticed little difference.

Our model for the bin counts was the following. Below we denote by t^*_c the mid-point $(c - 1/2)w$ of sub-interval I_c.

1. $X_{jc}^{e} \sim Poi\big(w \cdot \lambda^{e}(t_c^*)\big)$, $e \in \{A, B\}$, $c \in \{1, \dots, C\}$, $j \in \{1, \dots, n_e\}$. We assume both $\lambda^A(t)$ and $\lambda^B(t)$are smooth functions over $t \in [0, T]$.

2. $X_{jc}^{AB} \sim Poi(w \cdot \lambda_j(t_c^*))$, where $\lambda_j(t) = \lambda_j(t) = \alpha_j(t)\lambda^A(t) + \{1 - \alpha_j(t)\}\lambda^B(t)$ with $\alpha_j: [O, T] \to (0,1)$ being unknown smooth functions.

We modeled $\alpha_j(t) = S(\eta_j(t))$, where $S(t) = 1/(1 + e^{-t})$, and, each $\eta_j(t)$ was taken to be a (smooth) Gaussian process with $E\{\eta_j(t)\} \equiv \phi_j$, $Var\{\eta_j(t)\} \equiv \psi_j$, and, $Cor\{\eta_j(t), \eta_j(t')\} = \exp\{-0.5(t - t')^2/\ell_j^2\}$. The three parameters (ϕ_j, ψ_j, ℓ_j) respectively encoded the *long-term average value*, the *total swing magnitude* and the *waviness* of the $\alpha_j(t)$ curve. While the temporal imprint carried by each α_j was allowed to be distinct, we enforced the dual trials to share dynamic patterns by assuming $(\phi_j, \psi_j, \ell_j), j = 1, \dots, n_{AB}$, were drawn from a common, unknown probability distribution P, which we called a dynamic pattern generator and viewed as a characteristic of the triplet to be estimated from the data.

To facilitate estimation of P, we assumed it decomposed as $P = P_{\phi\psi} \times P_\ell$, where $P_{\phi\psi}$ was an unknown distribution on $(-\infty, \infty) \times (0, \infty)$ generating (ϕ_j, ψ_j), and, P_ℓ was an unknown distribution on $(0, \infty)$ generating ℓ_j. To simplify computation, we restricted ℓ_j to take only a finitely many positive values, representative of the waviness range we are interested in (in our analyses, we took these representative values to be {75, 125, 200, 300, 500}, all in ms). This restricted P_ℓ to be a finite dimensional probability vector.

We performed an approximate Bayesian estimation of model parameters. Note that only $\lambda^A(t)$ and $\lambda^B(t)$ were informed by the single sound trial data. All other model parameters were informed only by the dual sound trial data conditionally on the knowledge of $\lambda^A(t)$ and $\lambda^B(t)$. To take advantage of this partial factorization of information sources, we first smoothed each set of single sound trial data to construct a conditional gamma prior for the corresponding $\lambda^e(t_c^*), e \in \{A, B\}, c = 1, \dots, C$, where the gamma distribution's mean and standard deviation were matched with the estimate and standard error of $\lambda^e(t_c^*)$. A formal Bayesian estimation was then carried out on all model parameters jointly by (a) using only the dual sound trial data, (b) utilizing the conditional gamma priors on $\lambda^A(t)$ and $\lambda^B(t)$, and, (c) assuming a Dirichlet process prior [57] on $P_{\phi\psi}$ and an ordinary Dirichlet prior on P_ℓ. This final step involved a Markov chain Monte Carlo computation whose details will be reported in a separate paper.

Next, the estimate of the generator P was utilized to repeatedly simulate $\alpha(t)$ functions for hypothetical, new dual trials for the triplet. For each simulated $\alpha(t)$ curve, we computed its maximum swing size $|\alpha| = max_t\, \alpha(t) - min_t \alpha(t)$, and, time aggregated average value $\bar{\alpha} = \int_0^T \alpha(t)dt\,/T$. The waviness index of the triplet was computed as the odds of seeing an $\alpha(t)$ function exhibiting a swing of at least 50% between its peak and trough:

$$r_w = \frac{P(|\alpha|>0.5)}{P(|\alpha|<0.5)}$$

where P denotes the sampling proportion of the simulated α draws. The triplet's extremeness index was computed as the odds of seeing an $\alpha(t)$ function with its long-term average $\bar{\alpha}$ being closer to the mid-way mark of 50% than the extremes:

$$r_c = \frac{P(\bar{\alpha} \in (0.25,0.75))}{P(\bar{\alpha} \notin (0.25,0.75))}.$$

The two indices were then thresholded to generate a 2-way classification of all triplets. On waviness, a triplet was categorized as "wavy", "flat" or "ambiguous" according to whether $r_w > 1.3$, $r_w < 0.77$, or, $0.77 \leq r_w \leq 1.3$, respectively On extremeness, the categories were "central", "extreme", or, "ambiguous" according to whether $r_c > 3.24$, $r_c < 1.68$, or, $1.68 \leq r_c \leq 3.24$, respectively.

In addition to the flat/wavy and extreme/central classification, a third parameter was evaluated for each triplet: the degree of skewness in the distribution of $\overline{\alpha_*}$:

$$r_S = max\left(\left\{\frac{P(\overline{\alpha_*}<0.5)}{P(\overline{\alpha_*}>0.5)}, \frac{P(\overline{\alpha_*}>0.5)}{P(\overline{\alpha_*}<0.5)}\right\}\right.$$

which ranges in (1,∞). Each triplet's flat/wavy/central/extreme tag could then be subcategorized as either "skewed" or "symmetric" depending on whether $r_S > 4$ or $r_S < 2$ (with no label in the middle). This subcategorization step was useful for distinguishing the dynamic admixtures associated with the whole-trial categorizations of "single" and "outside" from "intermediate" and "mixture", with "single" and "outside" tending to be classified as "skewed". Supplementary Table 1 and Figures S4 and 5 give the full results of the main 2-way classification together with the symmetry/skewness subclassification, cross tabulated with the classification done under the whole trial spike count analysis.

<u>A vs. B assignment scores: individual neurons, pairs of neurons, local field potential, and behavioral prediction</u>

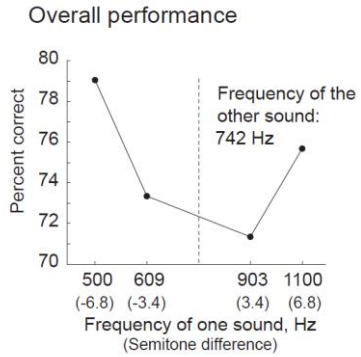
A vs. B assignment scores were computed for several analyses (the example shown in Figure 1E-F; pairs of recorded neurons; the relationship between spiking activity and local field potential; and the relationship between saccade sequences and spiking activity). For each triplet, every dual-sound trial received an "A-like" score and a "B-like" score, either for the entire response window (600-1000 ms after sound onset) or for 50 ms time bins. The scores were computed as the posterior probability that the spike count in each dual-sound trial was drawn from the Poisson distribution of single-sound spike counts,

For the pairs analysis, the A vs. B assignment scores were computed within each 50 ms time bin independently for each pair of neurons recorded simultaneously. The scores were normalized across trials by subtracting the mean score and dividing by the standard deviation of scores for that bin (a Z-score in units of standard deviation). Only conditions for which both recorded neurons exhibited reasonably different responses to the "A" vs. the "B" sound and for which there were at least 5 correct trials for A, B, and AB trials were included (t-test, $p < 0.05$). A total of 206 conditions were included in this analysis.

<u>Local field potential analysis</u>

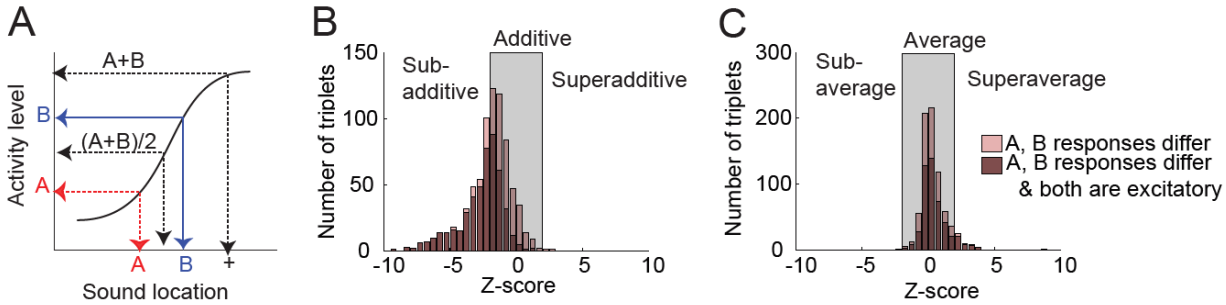
We analyzed the local field potential from 87 sites in both monkeys (30 sites from monkey P's left IC, 31 sites from monkey Y's right IC and 26 sites from monkey Y's left IC). The LFP acquisition was either recorded in discrete temporal epochs encompassing behavioral trials (roughly 1.2 to 2 seconds long) and at a sampling rate of 20 kHz (Dataset I, part of Dataset II), or as a continuous LFP signal during each session, at a sampling rate of 20 kHz or 1kHz (rest of Dataset II). We standardized the LFP signals by trimming the continuous LFP into single trial intervals and down-sampling all signals to 1 kHz. The MAP system filters LFP signals between 0.7 and 300 Hz; no additional filtering was applied. For each site we subtracted the overall mean LFP value calculated over the entire session, to remove any DC shifts, and we excluded trials that exceeded 500mV. For each triplet, we assigned individual dual-sound trials to two groups based on the total spike count in a 600 ms response window (see Methods: A vs. B assignment scores). The average LFP was then compared across the two groups in two 600 ms windows before and after sound onset (baseline and response periods). The results reported here refer to these mean-normalized LFP signals. We obtained similar results when the amplitude of each trial's LFP was scaled as a proportion of the maximum response within the session.

SUPPLEMENTARY FIGURES AND TABLES



Supplementary figure 1: monkey can localize two sounds.

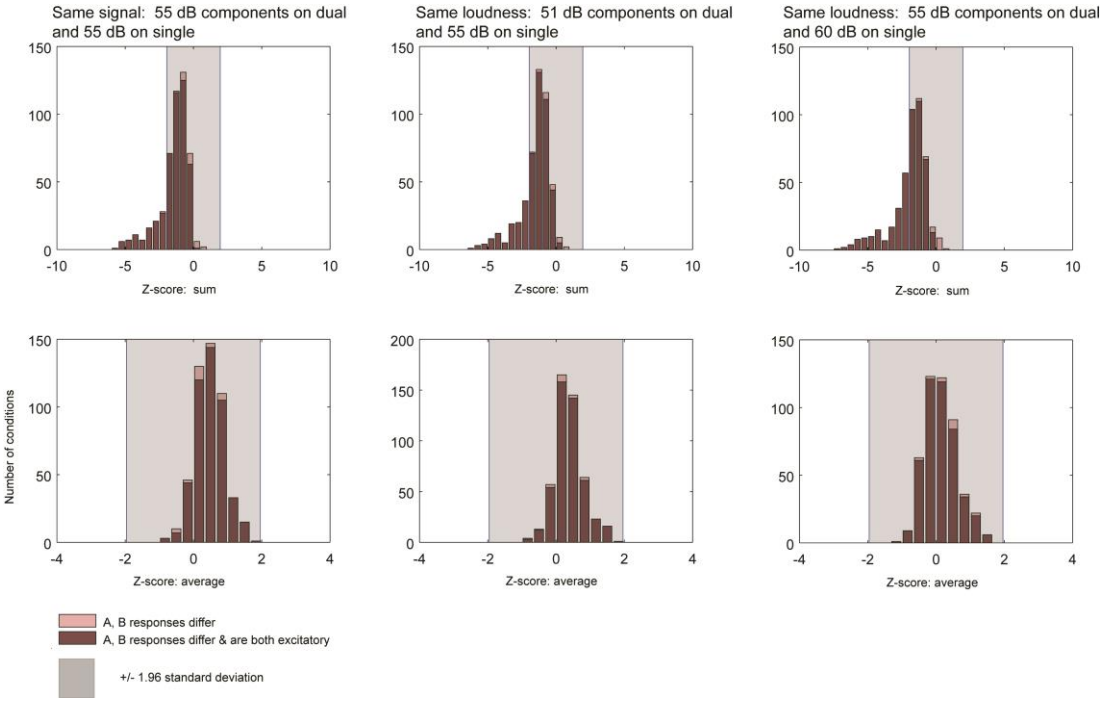
Average accuracy for the dual-sound targets was overall over 70% correct. Performance was better for larger frequency separations.



Supplementary figure 2: Results of conventional analyses pooling across time and trials are consistent with averaging but not summation.

(A) *Schematic activity patterns of IC neurons in response to single sounds A and B, and predicted response for dual sounds if the neuron sums inputs corresponding to individual sounds (A+B) or averages them ((A+B)/2).* ***(B)*** *Observed Z-scores of activity on dual-sound trials differ from the sum.* ***(C)*** *Observed Z-scores of activity on dual-sound trials correspond well to the average. The shaded areas indicate Z score values of +/- 1.96 (95% confidence intervals) on the sum (B) and on the average (C). This analysis was conducted on "triplets" of single- and dual-sound trials with a given set of locations and frequencies, pooling across intensities. Triplets were included if the single-sound responses differed (light bars, two-tailed t-test, $p<0.05$, $n=761$); results were*

similar when single-sound responses were different and both excitatory (dark bars, one-tailed t-test, p<0.05, n=486).



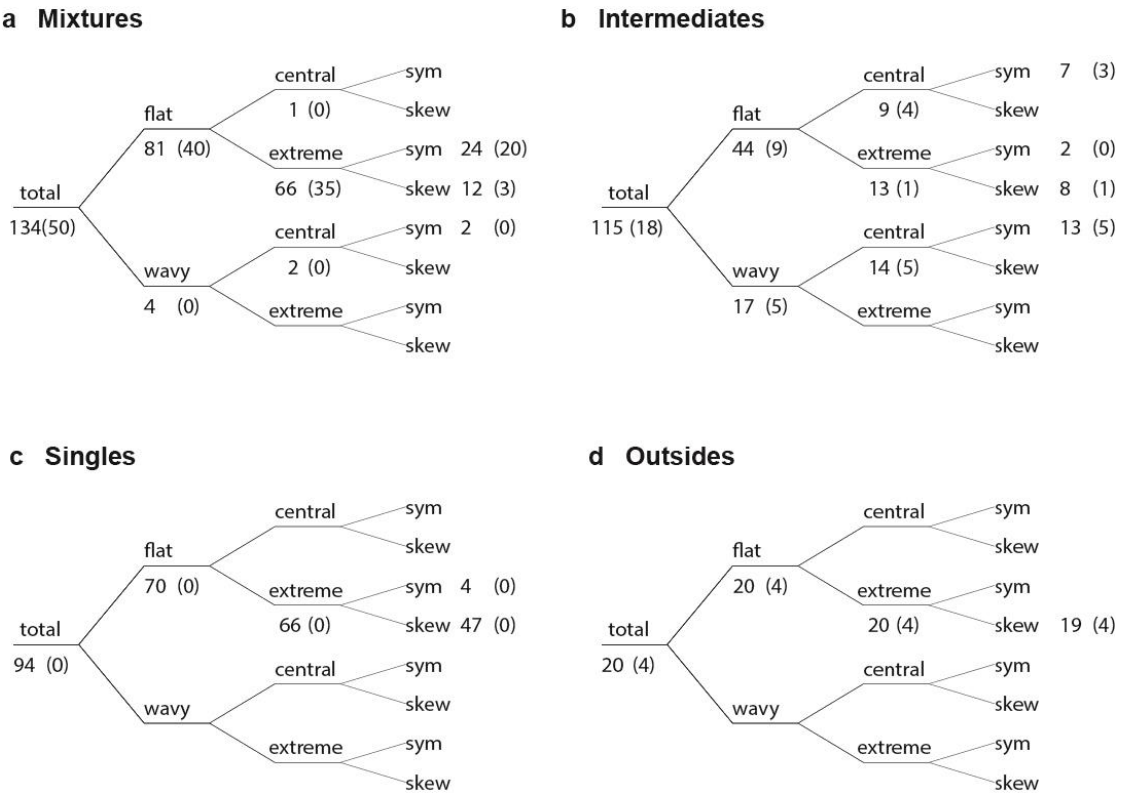
Supplementary Figure 3. Dual-sound responses do not depend on sound intensities

Same analysis as supplementary figure 2, but comparing dual-sound trials to single-sound trials that used the same signal sent to the speakers (left column), or for which the signal was amplified on single-sound trials to match the dual sounds in loudness (middle and right columns). The results are essentially identical to each other.

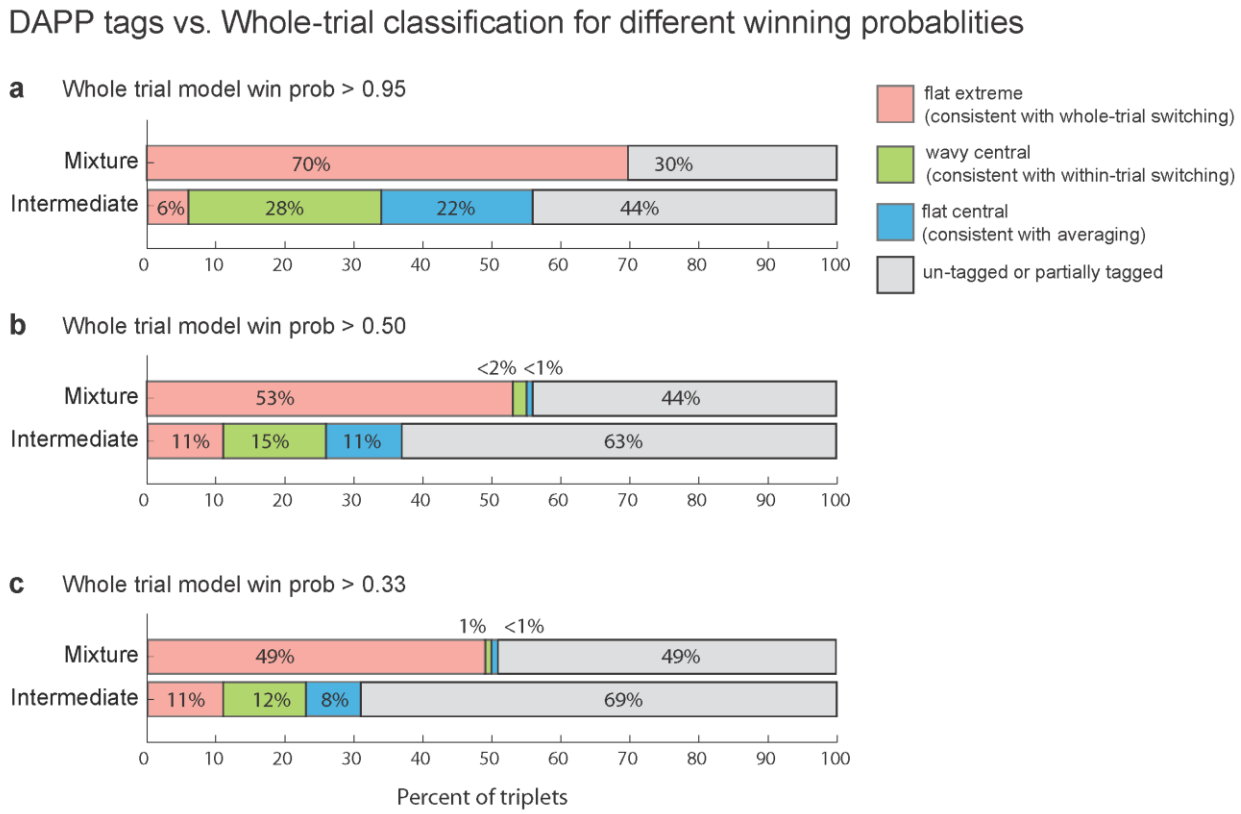
	Intermediate	Mixture	Outside	Single	Total
Wavy-central+symmetric	13 (5)	2			15 (5)
Wavy-central	1				1
Wavy+symmetric	4	2			6
Flat-central+symmetric	7 (3)				7 (3)
Flat-central	2 (1)	1			3 (1)
Flat-extreme+skewed	8 (1)	12 (3)	19 (4)	46	85 (8)
Flat-extreme+symmetric	2	22 (19)		4	28 (19)
Flat-extreme	3	31 (13)	1	16	51 (13)

Flat+skewed	4 (2)				4 (2)
Flat+symmetric	6	11 (4)		4	21 (4)
Flat	13 (2)	1		2	16 (2)
Extreme+symmetric	1	4 (4)		1	6 (4)
Extreme		3 (2)		1	4 (2)
Central+symmetric	9 (2)	1		2	12 (2)
Symmetric	36 (1)	42 (5)		14	92 (6)
Skewed	1 (1)				1 (1)
Ambiguous	5	2		3	10
Total	115 (18)	134 (50)	20 (4)	93	362 (72)

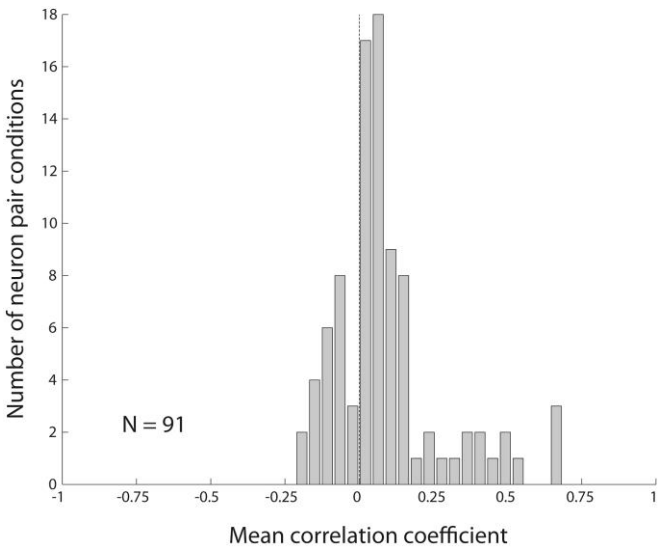
Supplementary Table 1. *Complete listing of "tags" applied in the DAPP analysis. Categories not shown (e.g. wavy-central+skewed) did not receive any members. The numbers in parentheses are for triplets that produced a winning model in the whole trial analysis with a probablility > 95%. The total is 362, since for one "single" triplet, the DAPP analysis failed to compute the necessary metrics.*



Supplementary Figure 4. *Hierarchical depiction of the DAPP tags, considering first flat vs. wavy, then central vs. extreme, then symmetric vs. skewed, as a function of whole trial model classification. As in Supplementary Table 1, numbers in parentheses are for triplets that produced a winning model in the whole trial analysis with a probability > 95% and the total included here is 362. For clarity, labels "0(0)" are not shown.*



Supplementary Figure 5. *Relationship between the three relevant DAPP tag combinations (flat-extreme, wavy-central, and flat-central) and whole-trial Poisson classification, for different levels of winning probability for the whole-trial analysis. Panel A is identical to Figure 3C in the main text.*



Supplementary Figure 6. *Same analysis as in Figure 4B, but using spike counts in each bin instead of A vs. B assignment scores.*