

## Representing the dynamics of natural marmoset vocal behaviors in frontal cortex

### Highlights

- Neurons in marmoset monkey PFC and PMC recorded during natural conversations
- GLM and PSTH were applied to quantify neural activity in continuous behavior
- Model-based approach robustly outperformed more traditional analyses
- Neurons in behavior-related functional clusters were distributed throughout PFC/PMC

### Authors

Jingwen Li, Mikio C. Aoi, Cory T. Miller

### Correspondence

jil286@ucsd.edu (J.L.),  
cormiller@ucsd.edu (C.T.M.)

### In brief

Li and colleagues applied model-based and traditional analyses to characterize single-neuron responses in the frontal cortex while marmosets engaged in their natural conversational exchanges. Results showed that the population supported nearly all facets of this ethological behavior through an anatomically distributed—but functionally modular—pattern of neural activity.



## Report

# Representing the dynamics of natural marmoset vocal behaviors in frontal cortex

Jingwen Li,<sup>1,\*</sup> Mikio C. Aoi,<sup>2,3,4,5,6</sup> and Cory T. Miller<sup>1,4,5,6,7,\*</sup>

<sup>1</sup>Cortical Systems & Behavior Lab, University of California, San Diego, La Jolla, CA 92093, USA

<sup>2</sup>Department of Neurobiology, University of California, San Diego, La Jolla, CA 92093, USA

<sup>3</sup>Halıcıoğlu Data Science Institute, University of California, San Diego, La Jolla, CA 92093, USA

<sup>4</sup>Neurosciences Graduate Program, University of California, San Diego, La Jolla, CA 92093, USA

<sup>5</sup>These authors contributed equally

<sup>6</sup>Senior author

<sup>7</sup>Lead contact

\*Correspondence: [jil286@ucsd.edu](mailto:jil286@ucsd.edu) (J.L.), [corymiller@ucsd.edu](mailto:corymiller@ucsd.edu) (C.T.M.)

<https://doi.org/10.1016/j.neuron.2024.08.020>

## SUMMARY

Here, we tested the respective contributions of primate premotor and prefrontal cortex to support vocal behavior. We applied a model-based generalized linear model (GLM) analysis that better accounts for the inherent variance in natural, continuous behaviors to characterize the activity of neurons throughout the frontal cortex as freely moving marmosets engaged in conversational exchanges. While analyses revealed functional clusters of neural activity related to the different processes involved in the vocal behavior, these clusters did not map to subfields of prefrontal or premotor cortex, as has been observed in more conventional task-based paradigms. Our results suggest a distributed functional organization for the myriad neural mechanisms underlying natural social interactions and have implications for our concepts of the role that frontal cortex plays in governing ethological behaviors in primates.

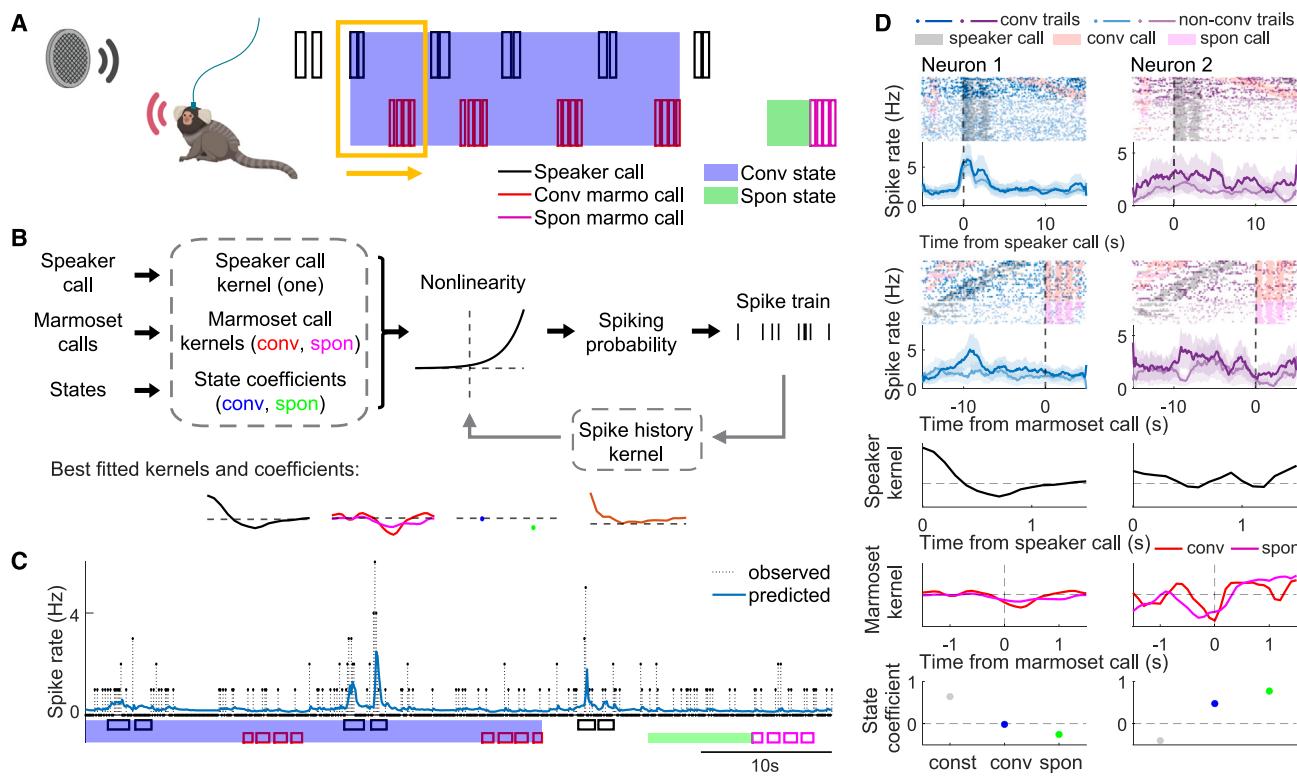
## INTRODUCTION

Natural behaviors typically vary along many dimensions—including their structure, timing, frequency, and occurrence with respect to other behaviors. This seemingly indomitable character of ethology has long been a key bottleneck to neuroscience<sup>1</sup> because of the difficulty it poses to explicating the relationship between the sources of this variance and patterns of neural activity. But as interest in natural behaviors grows, so too does the impetus to overcome these challenges and better understand their neural basis.<sup>2–4</sup> This issue is particularly pertinent to primate frontal cortex—i.e., prefrontal cortex (PFC) and premotor cortex (PMC)—because of the integral role these substrates play in governing a range of complex behavioral and cognitive processes.<sup>5–14</sup> Much of our understanding of primate frontal cortex has been from studies applying conventional, head-fixed task-based paradigms, revealing functions that would reasonably be assumed to be governed by the same subfields when primates engage in the analogous natural behaviors. Evidence to this end, however, is not clear. For example, a within-neuron comparison of marmoset PFC in both conventional and natural behavior contexts found that responses in the former context were not predictive of the latter, despite the stimuli being identical.<sup>15</sup> One potential explanation may be that changes in the state of frontal cortex during natural behaviors<sup>15,16</sup> or the inherent variability of ethological behaviors<sup>1,4</sup>

may mask the subfield specializations evident in more conventional paradigms. Alternatively, features of natural behaviors—such as the occurrence of sequential events<sup>1</sup>—may drive patterns of neural activity distinct to those contexts.<sup>17,18</sup> To reconcile these issues, a more refined analysis approach may be needed to investigate primate frontal cortex activity during natural, ethological behaviors.

To more effectively characterize the role of primate frontal cortex during natural behavior, here we applied a generalized linear model (GLM)-based approach to quantify neural activity during naturally occurring species-typical antiphonal calling exchanges in freely moving marmosets.<sup>19,20</sup> Such model-based analyses offer several advantages to this end because they can reveal relationships between neural responses and overlapping external covariates (sensory, motor, state, etc.) along with spike history and account for variability in the organization of behaviors, thereby capturing facets of brain/behavior interactions that descriptive statistics, such as mean firing rate, cannot.<sup>21–24</sup> Given the breadth of variability incumbent to most natural behaviors—including antiphonal calling—model-based approaches have the potential to be a powerful tool for revealing properties of the brain that may not be apparent otherwise. Neurophysiological studies of marmoset vocal behavior have shown that PFC and PMC neurons are responsive during different facets of the antiphonal exchanges,<sup>25</sup> including a latent social state.<sup>15,16</sup> However, the number of neurons involved was





**Figure 1. PSTHs and GLM-based analyses were applied on activity of neurons in the frontal cortex as marmosets engaged in conversational exchanges**

(A) Interactive playback design. Marmoset calls were categorized as conversational (red) or spontaneous (pink) calls depending on whether the marmoset responded to a speaker call (black) or initiated a call, respectively. The period during an interactive communication was labeled as conversational state (blue), and the period before spontaneous calls was labeled as spontaneous state (green). The yellow box indicated the sliding window used in GLM for the continuous recording.

(B) GLM-based analysis framework. Behavioral events, internal states, and spike history were included as regressors in the GLM model to get best-fitted kernels and coefficients (in gray dashed box).

(C) Observed and GLM predicted spike rates throughout a continuous recording.

(D) PSTHs and GLM kernels of two example neurons. Neuron 1 (area “8av”) exhibited a significant increase in response to hearing a vocalization when analyzed by the PSTH aligned by conspecific call onset as well as a significant GLM kernel. Neuron 2 (area “8av”) exhibited decreased activity during call production in both PSTHs and GLM marmoset call kernels, as well as significant GLM state coefficients quantifying the changes in the activity related to the animal’s conversational and spontaneous states. From top to bottom: spike raster and PSTHs aligned by hearing calls and producing calls, respectively; hearing call kernel; marmoset call kernels (red and pink for conversational and spontaneous call kernels, respectively); and state coefficients. In spike raster, darker dots correspond to conversational trials and lighter dots correspond to non-conversational (aligned by hearing calls) or spontaneous (aligned by producing calls) trials; gray, red, and pink shaded areas are time periods of speaker calls, conversational calls, and spontaneous calls, respectively. In PSTHs, darker lines are PSTHs for conversational trials, and lighter lines are non-conversational or spontaneous trials; shades represent standard error.

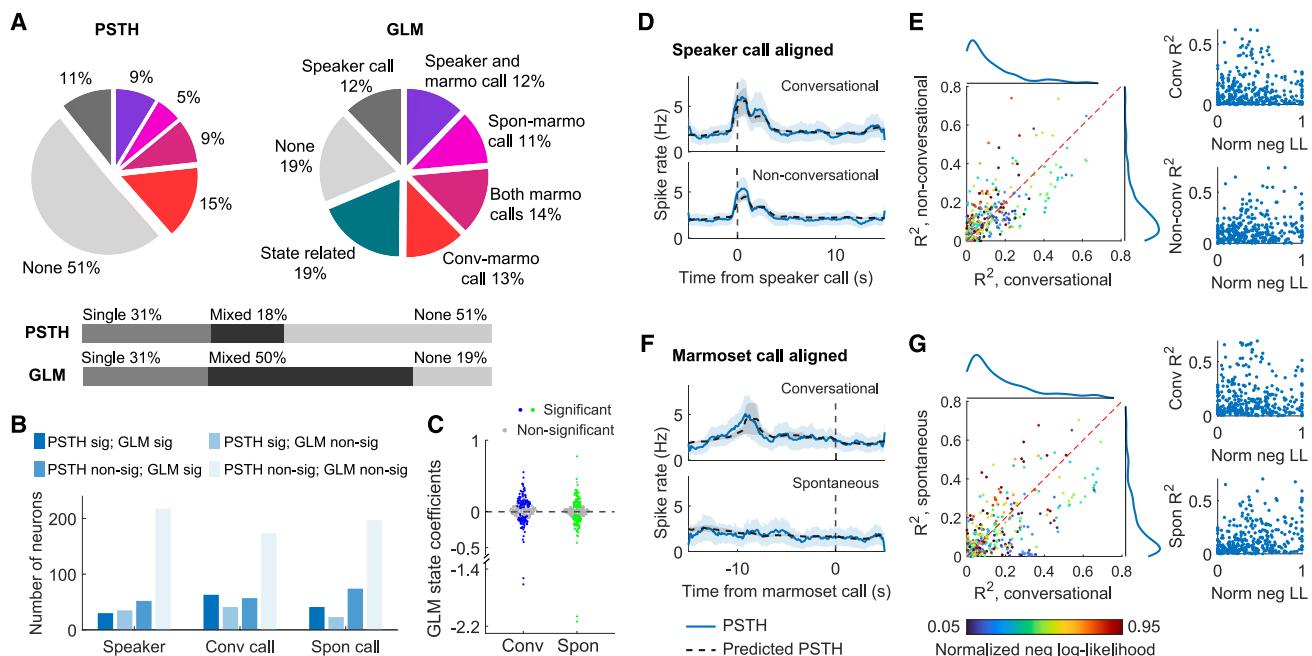
relatively low (~35%), leaving open the possibility that the traditional analyses used in these studies lacked the quantitative power to fully characterize the role of frontal cortex in these natural social interactions.

## RESULTS

### GLMs can characterize continuous events in natural vocal behavior

We recorded the activity of single neurons in PMC and PFC while freely moving marmosets engaged in a natural, species-typical vocal behavior—antiphonal calling (Figure 1A; see STAR Methods). Neural activity was quantified using two different analysis approaches: peri-stimulus time histograms (PSTHs) and

GLM-based analysis. Although PSTH analyses have been used previously to quantify neural activity during this behavior,<sup>16,25</sup> whether the GLM could be applied given various dimensions of variability incumbent to these vocal exchanges was untested. To adapt the GLM-based analysis to our time-varying, naturalistic experimental paradigm, a set of task variables—behavioral events, internal behavioral states, and spike history—were included as regressors within a sliding window (Figures 1B and S1; STAR Methods), allowing for a continuous prediction of spike rate over the course of a session (Figure 1C). Behavioral analysis was performed in parallel to identify vocal-perception events (i.e., hearing calls) and vocal-motor events (i.e., producing calls) as well as the social context (i.e., antiphonal conversation, spontaneous calls, etc.)—and the animal’s underlying state



**Figure 2. GLM-based analysis identifies more neurons with significant functions and successfully recapitulates PSTHs**

- (A) Top: percentage of neurons in categories that significantly respond to communication events or internal behavior states identified by PSTHs and GLM, respectively. Bottom: percentage of neurons with single or mixed selectivity identified in PSTHs and GLM, respectively.
- (B) Number of neurons in four categories for hearing calls, emitting conversational calls, and spontaneous calls: significant in PSTH and significant in GLM; significant in PSTH and non-significant in GLM; non-significant in PSTH and significant in GLM; non-significant in PSTH and non-significant in GLM.
- (C) GLM coefficients for conversational (blue) and spontaneous (green) states.
- (D) Observed PSTH (blue solid line) and predicted PSTH from GLM (black dash line) aligned with hearing calls for an example neuron, grouped by conversational trials (top) and non-conversational trials (bottom). Shaded area represents standard error. The predicted PSTHs from GLM well recapitulated the observed PSTHs.
- (E) R<sup>2</sup> between the observed PSTH and the predicted PSTH from GLM of each neuron for non-conversational trials as a function of conversational trials ( $\text{corr} = 0.70, p < 10^{-3}$ ). Blue lines show distribution of R<sup>2</sup> for the corresponding axes. Color indicates the GLM negative log-likelihood of the neuron normalized to [0,1] among all the neurons. The GLM negative log-likelihood, indicating the overall fitting of GLM throughout the recording, is not correlated to the conversational R<sup>2</sup> and weakly correlated to the non-conversational R<sup>2</sup> ( $\text{corr} = 0.17, p = 0.002$ ).
- (F) Same as in (D) but aligned by producing calls for the same example neuron, grouped by conversational trials (top) and spontaneous trials (bottom).
- (G) Same as in (E) but aligned with producing calls. The R<sup>2</sup> in conversational and spontaneous contexts are correlated ( $\text{corr} = 0.63, p < 10^{-3}$ ). The GLM negative log-likelihood was not correlated to the conversational R<sup>2</sup> but correlated to the non-conversational R<sup>2</sup> ( $\text{corr} = 0.31, p < 10^{-3}$ ).

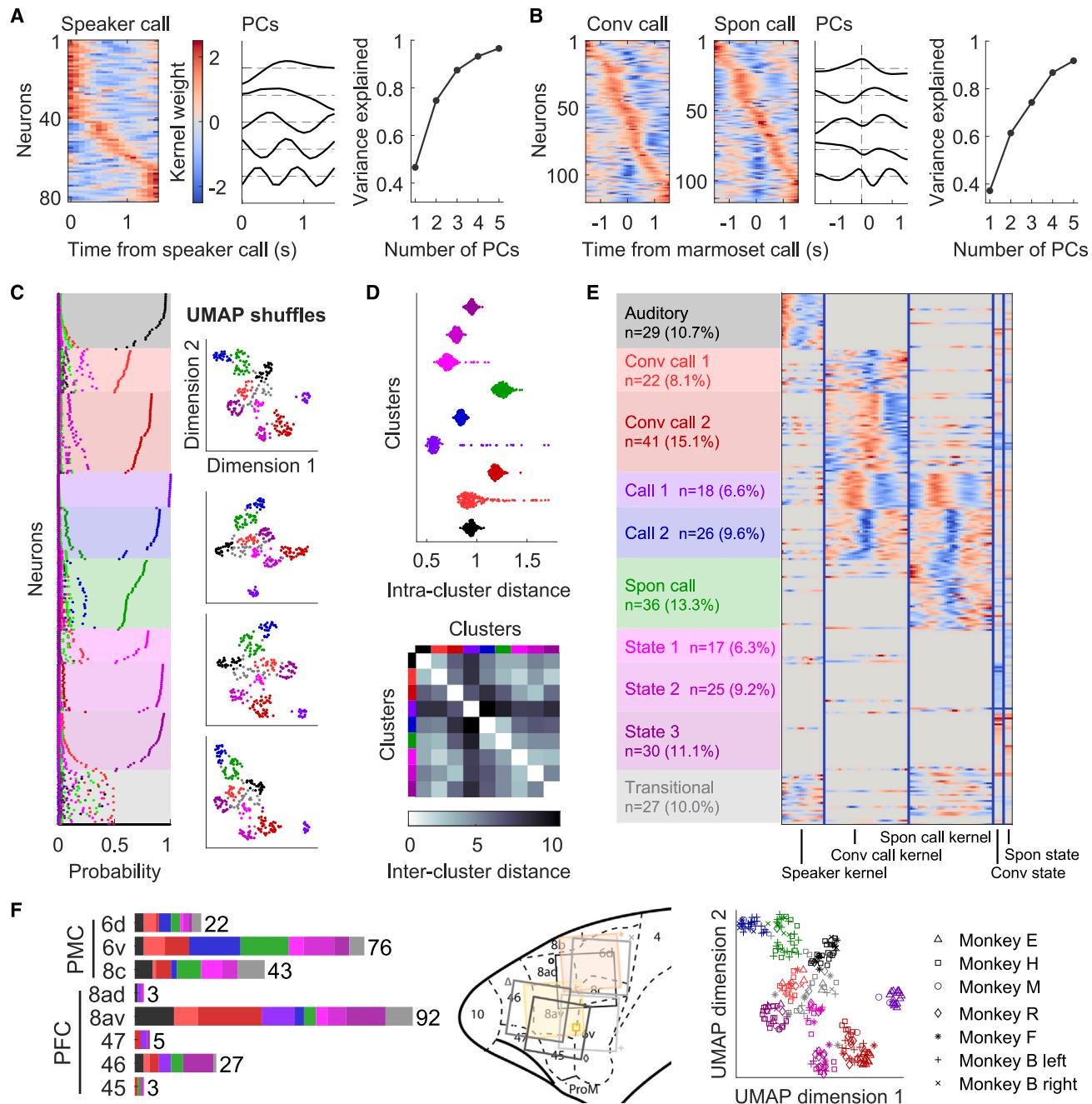
(i.e., conversational or spontaneous calling).<sup>15,16</sup> GLM-based analyses accurately model spike rates for neurons in trial-based tasks,<sup>21,22</sup> so it was important to confirm that our adaptation of this approach to a continuous natural behavior would yield a similar outcome. We reasoned that the sliding window would account for this variance and result in an improved characterization. Indeed, neuronal firing rates responding to hearing calls or producing calls characterized by GLM kernels were consistent with patterns observed in PSTHs (Figure 1D), thereby confirming that our modified GLM analysis could reliably capture multiple facets of a continuous primate social interaction (i.e., conversations).

#### GLM encapsulates natural brain/behavior interactions

Building on these observations, we applied both the PSTH- and GLM-based analysis to each neuron in the population to directly compare how these different quantitative approaches encapsulate interactions between this natural behavior and frontal cortex activity. As highlighted in Figure 2A, GLM analysis (81% with per-

mutation test at 0.05 significance level) identified more neurons with significant modulation by sensory, behavioral, or state events as well as mixed selectivity compared with the PSTH analyses (49% with Wilcoxon signed rank test at 0.05 significance level). Furthermore, neurons exhibiting significant changes in activity for hearing and/or producing calls in PSTHs were more likely to be identified significant in the GLM analysis than vice versa (Figure 2B). Moreover, the GLM revealed neurons in the population having significant state-related responses with large positive or negative GLM state coefficient values (Figure 2C), a feature of the natural behavior that was entirely undetected by the PSTH analysis. Overall, the GLM analysis was more sensitive to behavior-related neural modulations that were not evident in the overall change of firing rate before and after stimulus than in PSTHs (Figure S2).

To test whether neural activity captured in our model-based analysis was consistent with PSTHs, we used the predicted spike rate from the GLMs to reconstruct PSTHs (Figures 2D–2G). As shown in Figure 2D, the firing rate of a representative



**Figure 3. The population in frontal cortex forms a continuous spectrum with clusters of functions from dimensionality reduction and clustering analysis on GLM kernels and coefficients**

(A) PCA of significant speaker kernels. From left to right: significant speaker kernels; corresponding first 5 PCs; variance explained by the first 5 PCs.

(B) PCA of significant marmoset call kernels. From left to right: significant conversational and spontaneous marmoset call kernels; corresponding first 5 PCs; variance explained by the first 5 PCs.

(C) Clusters identified from the Gaussian mixture model (GMM) on the projection of each neuron's GLM kernels and coefficients on two UMAP dimensions. Left: probability of each neuron being classified in a cluster by GMM, grouped and sorted with the probability > 0.5 being in a dominant cluster. Right: four examples of UMAP projections. Each point is a neuron with color-indicating clusters.

(D) Top: intra-cluster distance for 200 runs of UMAP projections. The low variance of the intra-cluster distance indicates the robustness of the clustering results. Bottom: mean inter-cluster distance of 200 runs of UMAP projections.

(legend continued on next page)

neuron predicted by the GLM when a marmoset heard a call in both the conversational and non-conversational contexts was faithfully recapitulated by the GLM. Though the  $R^2$  of the neuron was correlated between the conversational and non-conversational contexts, its overall encoding performance throughout the recording represented by the negative log-likelihood of the GLM did not correlate with the conversational  $R^2$  and was weakly correlated with non-conversational  $R^2$  (Figure 2E). A similar result was evident during vocalization production for both reconstructing firing rates of individual neurons (Figure 2F) and robust encoding evident for the population (Figure 2G). Like the result above (e.g., Figure 2B), our model-based approach captured a greater extent of the brain/behavior interaction in these natural vocal exchanges. These findings demonstrate that the GLM-based analysis is a more powerful tool than traditional PSTH analyses to identify the myriad neural processes that underlie dynamic, continuous natural behaviors.

### Frontal cortex encodes functional clusters

Given the success of the GLM-based analysis to characterize natural behavior-related events, we next sought to determine the encoding strategy of primate frontal cortex for a suite of functions that may covary during the vocal behavior. Specifically, analyses were designed to test whether neurons in the frontal cortex are allocated to distinct functional roles or whether neurons form a continuous spectrum of various functions. We collectively analyzed all the GLM kernels and coefficients with dimensionality reduction and clustering analysis as a single dataset because of the high percentage of neurons with mixed selectivity in the population (Figure 3). The population structure of significant kernels corresponding to the animal hearing and emitting calls was examined using dimensionality reduction to extract key features of these significant kernels (Figures 3A and 3B). Principal-component analysis (PCA) applied to the significant GLM speaker call kernels revealed that the first 3 PCs explained more than 80% of the variance (Figure 3A), while the first 4 PCs explained more than 80% of the variance for the call production kernels (Figure 3B). When determining the population structure of functional roles, including higher PC dimensions would obscure the shared features among kernels, while fewer PCs omitted important features to capture clustering structure. Thus, we collectively constructed a multi-dimensional space with the projection of the speaker kernel on the first 3 PCs, the projection of the conversational or spontaneous call kernels on the first 4 PCs, and the conversational and spontaneous state coefficients for the following analysis (see STAR Methods). Next, we applied dimensionality reduction on the multi-dimensional space with uniform manifold approximation and projection (UMAP) and clustering analysis with a Gaussian mixture model (GMM) to quantify functional separability in the population (see STAR Methods). Across 200 UMAP runs, the GMM consistently identified nine distinct functional clusters, as well as a minority of neurons (10%) with ambiguous cluster identity (Figures 3C and

3D), demonstrating that neurons supporting this vocal behavior organize into separable modules in frontal cortex.

Interestingly, we observed more functional clusters than behavioral variables (Figure 3E). While only a single auditory cluster was evident for neurons responding to hearing calls, vocal production separated into multiple clusters, such as two clusters of neurons that respond specifically to producing calls only in conversations, two clusters of neurons that respond to producing any calls, and a cluster of neurons that respond specifically to producing spontaneous calls. Likewise, three clusters of neurons mainly associated with the internal states were also evident. Surprisingly, the functional cluster identity of each neuron did not correspond to the anatomical position of the electrodes in frontal cortex subfields, suggesting that functional identities were broadly spatially distributed (Figures 3F and S3). Neurons in each functional cluster were on average observed in 5 anatomically distinct areas of frontal cortex (Figure 3F left). Neurons in the “hearing” cluster, for example, were found in 6 different fields of PMC and PFC despite the fact that a direct projection from auditory cortex to vIPFC<sup>26</sup> is well known, suggesting that the functional clusters supporting this natural communication behavior may be influenced by processes distinct to social interactions.<sup>27,28</sup>

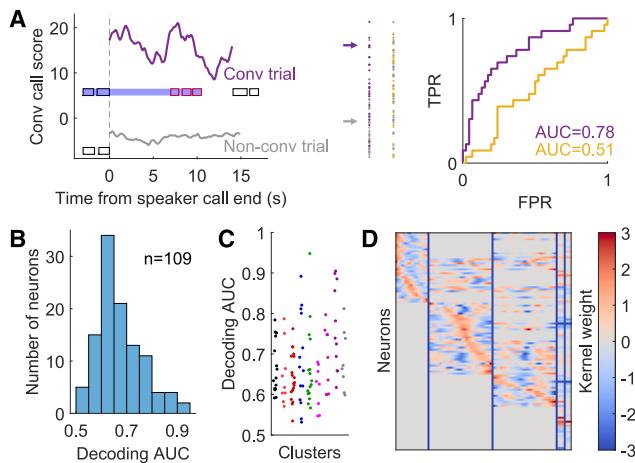
Overall, these analyses indicate that primate frontal cortex comprises functional clusters representing the full spectrum of natural behavior-related processes that occur during communication. Notably, separation of call production into multiple functional clusters suggests that a more dynamic suite of neural mechanisms underlies this process than a singular vocal-motor act,<sup>29</sup> potentially due to its close relationship with other processes in the behavioral sequence.<sup>18,30</sup>

### Natural behavior can be predicted from GLM kernels

The distinct functional clusters evident in marmoset frontal cortex are suggestive of the broad role that these neocortical substrates play in supporting vocal behavior. To explore the significance of neurons in these functional clusters, we next tested their respective contributions to predict whether a marmoset will produce a vocalization in response to hearing a conspecific call. Because kernels in each cluster varied in their behavior, state, and context-related activity, we reasoned that only some neurons carry predictive information about subjects' vocal behavior. For this analysis, we calculated a conversational call score (CCS) each time a conspecific call was heard. The CCS corresponded to the log-likelihood ratio of a GLM that assumed a marmoset emitted a conversational call—defined as a call produced in response to hearing a conspecific call—versus one assuming no response (Figure 4A; see STAR Methods). If the marmoset produced a conversational call after hearing a call, the GLM assumes that the emitted call should have a better fit to the data and yield a high CCS. Conversely, the GLMs should have a worse fit when assuming a conversational call exists and yield a low CCS if no response occurred. We selected the highest

(E) The GLM kernels and coefficients of neurons in each cluster organized by color. Left column lists the specific function demonstrated in each cluster identified from GMM.

(F) Left: number of neurons recorded in each cortex area where color indicates functional clusters. The number on the right indicates the total number of neurons in each area. Middle: topography of the arrays. Right: an example UMAP projections with symbol indicating arrays and color-indicating clusters.



**Figure 4. Using GLMs to decode a marmoset's response is successful for neurons from various clusters and functions**

(A) Decoding process illustrated by two example neurons. Left: conversational call score (CCS) measured by GLM negative log-likelihood as a function of time from speaker call end for a conversational trial (purple line) and a non-conversational trial (gray line) of a neuron. Middle: CCS for all trials of two neurons (purple and yellow). Each dot represents a score of a trial. Dots in purple or yellow are conversational trials and dots in gray are non-conversational trials. The two arrows represent the two trials in the left. Right: ROC curve and AUC of the two neurons in the middle.

(B) Distribution of decoding AUC for neurons with significant GLM kernels or coefficients and significant decoding performance. 109 neurons yielded significant decoding performance.

(C) Decoding AUC of neurons grouped in different clusters. Color of the clusters is the same as in Figure 3.

(D) GLM kernels and coefficients of the neurons with significant decoding performance.

score after the marmoset hears a call to represent how likely the marmoset was to engage in the conversation and calculated the receiver operating characteristic (ROC) curve and area under curve (AUC) for each neuron (Figure 4A). This analysis revealed a range of decoding performances among neurons, where 32.5% of neurons yielded significant decoding performance with some neurons' AUC as high as 0.8 and above (Figure 4B), indicating robust predictability of the behavioral event. The decoding performance of neurons in each functional cluster is not significantly different from each other, and neurons with high decoding performance were distributed across clusters (Figure 4C). Likewise, the population with the highest decoding performance consisted of neurons with diverse functions (Figure 4D). These results further support our observation that functions supporting the different facets of the natural behavior are anatomically distributed throughout frontal cortex.

## DISCUSSION

Here, we examined the respective contributions of primate PFC and PMC to processes underlying natural marmoset vocal behavior. To account for the inherent variability in continuous social interactions, such as antiphonal calling, a GLM-based analysis<sup>21,22</sup> was adapted to evaluate the activity of neurons throughout frontal cortex. This model-based analysis robustly

outperformed more traditional PSTH-based analyses to characterize the species-typical behavior, as it identified more neurons with significant behavior-related functions—i.e., hearing and producing calls in different social contexts—as well as state-related neural activity. This discrepancy likely occurred because the GLM considers the entire spike history related to an event but not necessarily aligned to the event—as in the PSTH analyses—which may be critical for capturing the dynamics of continuous ethological behaviors. Analyses also revealed distinct functional clusters in frontal cortex, encapsulating the different behavioral, state, and context-related properties. Notably, the functional clusters were distributed in an anatomically heterogeneous organization that contrasts with analogous data from more traditional task-based experiments, including those testing vocalization recognition and production.<sup>31–35</sup> These results suggest that primate frontal cortex activity supports nearly all facets of a natural, continuous vocal behavior through an anatomically distributed—but functionally modular—pattern of population activity.

Findings here were possible because of two key analysis innovations that may be integral to generalizing this approach to other natural behaviors. First, a sliding window was applied to the GLM analysis to study the encoding properties of cells on a moment-by-moment, but continuous basis. Second, rather than only interpreting the learned kernels for each cell independently, we also took a population-level view of the cell encoding properties. By performing dimensionality reduction and clustering analysis on the GLM kernels and coefficients, we revealed separable modules in the population playing distinct functional roles that spanned the breadth of the natural behavior. Notably, these clusters were not localized to frontal cortex subfields but rather formed a continuum of functions that were independent of anatomical position. This population encoding strategy may benefit the system in maximizing information propagation or functional robustness.<sup>36,37</sup> A balance of segregation and integration must be maintained in frontal cortex to flexibly adapt cognitive strategies, a particularly powerful computational mechanism for mediating social interactions that vary due to—sometimes unpredictable—changes in the social and ecological landscapes. In human vocal turn-taking, for example, distinct processes like listening and planning occur in parallel<sup>30,38</sup> by neural mechanisms optimized for these vocal exchanges.<sup>18</sup> As a learned vocal behavior that involves numerous cognitive processes, ranging from social categorization and decision-making to social monitoring,<sup>19,39–44</sup> marmoset antiphonal calling likely relies on mechanisms in frontal cortex<sup>35</sup> that similarly occur in parallel despite the serial nature of the behavioral interaction. While the pattern of activity observed here is different from previous observations of mixed selectivity in marmoset PFC across different contexts,<sup>15</sup> neural activity examined here was only during natural communication. It is possible that the functional role of these neurons may be consistent within a particular context but change in response to the demands of the immediate environment or behaviors (e.g., head direction, posture, hunger, etc.) not quantified here that covary with an aspect of the vocal behavior.<sup>45–47</sup>

Perspectives on the neural basis of primate behavior are evolving in parallel with the increased use of ethological paradigms<sup>17,24,48–50</sup> and as effects in conventional tasks are not necessarily recapitulated in the presumptive analogous natural

behaviors.<sup>15,51</sup> Our results here showed that—while functional clusters of continuous behavior-related processes are evident in the PFC and PMC of marmosets engaged in natural vocal interactions—the neurons within these clusters are widely distributed throughout frontal cortex, a stark contrast to the most directly analogous studies with conventional paradigms.<sup>31,32,34</sup> We conjecture that the distributed functional organization during this natural behavior was evident because the behavioral and cognitive processes that have traditionally been studied selectively through more targeted experimentation become more covaried in continuous, natural behaviors<sup>1</sup> and, as such, rely on a more distributed computational ensemble coding strategy in frontal cortex.<sup>52</sup> This result does have precedence in human PFC, as Broca's area was shown to be involved in more dynamic computations when studied in more complex conditions than had previously been recognized.<sup>53</sup> Because only a single animal was recorded in each test session, we were unable to appreciate the influence of synchronized changes in brain state that emerge in coordinated social interactions between two or more individuals<sup>27</sup>—including during human vocal exchanges<sup>54,55</sup>—on the functional clusters observed here. Moreover, because the number of neurons recorded in each test session was limited, the use of more sophisticated population models could not be applied to better understand the ensemble coding and network dynamics, including how different states are reflected in a dynamical system. Though admittedly limited in number, the few studies to directly compare neural responses in the primate brain between conventional and naturalistic paradigms have reported considerable differences<sup>15,51</sup> that may portend the need to evolve our conceptions of brain computations and behavior. Rather than be an impediment to understanding frontal cortex function—or the brain more generally—the behavioral variability and covariance inherent to ethological behaviors may be a uniquely powerful engine of discovery, as it is only in these contexts that at least some computational mechanisms needed to govern the more distinct primate cognitive faculties may emerge and be open to scientific inquiry.<sup>56–58</sup>

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Cory Miller ([cormiller@ucsd.edu](mailto:cormiller@ucsd.edu)).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

- Behavioral and neural data have been deposited at Dryad and are publicly available as of the date of publication. DOIs are listed in the [key resources table](#).
- All original code has been deposited at Dryad and is publicly available as of the date of publication. DOIs are listed in the [key resources table](#).
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

## ACKNOWLEDGMENTS

We thank Drs. Vlad Jovanovic, Sam Nummela, and Wren Thomas for assistance with data collection and Dr. Alexander Huk for discussion. This work was supported by NIH R01 DC012087 to C.T.M.

## AUTHOR CONTRIBUTIONS

Conceptualization: J.L. and C.T.M.; methodology: J.L., M.C.A., and C.T.M.; investigation: J.L.; data curation: J.L. and M.C.A.; writing: J.L., M.C.A., and C.T.M.; visualization: J.L. and M.C.A.; supervision: M.C.A. and C.T.M.; funding acquisition: C.T.M.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS](#)
  - Subjects
- [METHOD DETAILS](#)
  - Neurophysiology
  - Experiment design
- [QUANTIFICATION AND STATISTICAL ANALYSIS](#)
  - PSTH analysis
  - GLM based analysis
  - Clustering analysis
  - Decoding

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.neuron.2024.08.020>.

Received: March 19, 2024

Revised: July 26, 2024

Accepted: August 28, 2024

Published: September 23, 2024

## REFERENCES

- Miller, C.T., Gire, D., Hoke, K., Huk, A.C., Kelley, D., Leopold, D.A., Smear, M.C., Theunissen, F., Yartsev, M., and Niell, C.M. (2022). Natural behavior is the language of the brain. *Curr. Biol.* 32, R482–R493. <https://doi.org/10.1016/j.cub.2022.03.031>.
- Datta, S.R., Anderson, D.J., Branson, K., Perona, P., and Leifer, A. (2019). Computational Neuroethology: A Call to Action. *Neuron* 104, 11–24. <https://doi.org/10.1016/j.neuron.2019.09.038>.
- Dennis, E.J., El Hadj, A., Michael, A., Clemens, A., Tervo, D.R.G., Voigts, J., and Datta, S.R. (2021). Systems Neuroscience of Natural Behaviors in Rodents. *J. Neurosci.* 41, 911–919. <https://doi.org/10.1523/JNEUROSCI.1877-20.2020>.
- Pereira, T.D., Shaevitz, J.W., and Murthy, M. (2020). Quantifying behavior to understand the brain. *Nat. Neurosci.* 23, 1537–1549. <https://doi.org/10.1038/s41593-020-00734-z>.
- Miller, E.K., Lundqvist, M., and Bastos, A.M. (2018). Working memory 2.0. *Neuron* 100, 463–475. <https://doi.org/10.1016/j.neuron.2018.09.023>.
- Lara, A.H., and Wallis, J.D. (2015). The Role of Prefrontal Cortex in Working Memory: A Mini Review. *Front. Syst. Neurosci.* 9, 173. <https://doi.org/10.3389/fnsys.2015.00173>.
- Miller, E.K., Nieder, A., Freedman, D.J., and Wallis, J.D. (2003). Neural correlates of categories and concepts. *Curr. Opin. Neurobiol.* 13, 198–203. [https://doi.org/10.1016/s0959-4388\(03\)00037-0](https://doi.org/10.1016/s0959-4388(03)00037-0).
- Mansouri, F.A., Freedman, D.J., and Buckley, M.J. (2020). Emergence of abstract rules in the primate brain. *Nat. Rev. Neurosci.* 21, 595–610. <https://doi.org/10.1038/s41583-020-0364-5>.

9. Hernández, A., Zainos, A., and Romo, R. (2002). Temporal evolution of a decision-making process in medial premotor cortex. *Neuron* 33, 959–972. [https://doi.org/10.1016/s0896-6273\(02\)00613-x](https://doi.org/10.1016/s0896-6273(02)00613-x).
10. Wallis, J.D., and Miller, E.K. (2003). From rule to response: neuronal processes in the premotor and prefrontal cortex. *J. Neurophysiol.* 90, 1790–1806. <https://doi.org/10.1152/jn.00086.2003>.
11. Buschman, T.J. (2021). Balancing flexibility and interference in working memory. *Annu. Rev. Vis. Sci.* 7, 367–388. <https://doi.org/10.1146/annurev-vision-100419-104831>.
12. Fuster, J.M. (2008). *The Prefrontal Cortex* (Academic Press).
13. Petrides, M. (2005). Lateral prefrontal cortex: architectonic and functional organization. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 781–795. <https://doi.org/10.1098/rstb.2005.1631>.
14. Wise, S.P. (1985). The primate premotor cortex: Past, present, and preparatory. *Annu. Rev. Neurosci.* 8, 1–19. <https://doi.org/10.1146/annurev.ne.08.030185.000245>.
15. Jovanovic, V., Fishbein, A.R., de la Mothe, L., Lee, K.-F., and Miller, C.T. (2022). Behavioral context affects social signal representations within single primate frontal cortex neurons. *Neuron* 110, 1318–1326.e4. <https://doi.org/10.1016/j.neuron.2022.01.020>.
16. Nummela, S.U., Jovanovic, V., de la Mothe, L.A., and Miller, C.T. (2017). Social context-dependent activity in marmoset frontal cortex populations during natural conversations. *J. Neurosci.* 37, 7036–7047. <https://doi.org/10.1523/JNEUROSCI.0702-17.2017>.
17. Testard, C., Tremblay, S., Parodi, F., DiTullio, R.W., Acevedo-Ithier, A., Gardiner, K.L., Kording, K., and Platt, M.L. (2024). Neural signatures of natural behaviour in socializing macaques. *Nature* 628, 381–390. <https://doi.org/10.1038/s41586-024-07178-6>.
18. Castellucci, G.A., Kovach, C.K., Howard, M.A., 3rd, Greenlee, J.D.W., and Long, M.A. (2022). A speech planning network for interactive language use. *Nature* 602, 117–122. <https://doi.org/10.1038/s41586-021-04270-z>.
19. Grijseels, D.M., Fairbank, D.A., and Miller, C.T. (2024). A model of marmoset monkey vocal turn-taking. *Proc. Biol. Sci.* 291, 20240150. <https://doi.org/10.1098/rspb.2024.0150>.
20. Miller, C.T., and Wang, X. (2006). Sensory-motor interactions modulate a primate vocal behavior: antiphonal calling in common marmosets. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 192, 27–38. <https://doi.org/10.1007/s00359-005-0043-z>.
21. Park, I.M., Meister, M.L.R., Huk, A.C., and Pillow, J.W. (2014). Encoding and decoding in parietal cortex during sensorimotor decision-making. *Nat. Neurosci.* 17, 1395–1403. <https://doi.org/10.1038/nn.3800>.
22. Latimer, K.W., Yates, J.L., Meister, M.L.R., Huk, A.C., and Pillow, J.W. (2015). Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. *Science* 349, 184–187. <https://doi.org/10.1126/science.aaa4056>.
23. Aoi, M.C., Mante, V., and Pillow, J.W. (2020). Prefrontal cortex exhibits multidimensional dynamic encoding during decision-making. *Nat. Neurosci.* 23, 1410–1420. <https://doi.org/10.1038/s41593-020-0696-5>.
24. Voloh, B., Maisson, D.J.-N., Cervera, R.L., Conover, I., Zambre, M., Hayden, B., and Zimmermann, J. (2023). Hierarchical action encoding in prefrontal cortex of freely moving macaques. *Cell Rep.* 42, 113091. <https://doi.org/10.1016/j.celrep.2023.113091>.
25. Miller, C.T., Thomas, A.W., Nummela, S.U., and de la Mothe, L.A. (2015). Responses of primate frontal cortex neurons during natural vocal communication. *J. Neurophysiol.* 114, 1158–1171. <https://doi.org/10.1152/jn.01003.2014>.
26. Romanski, L.M., Bates, J.F., and Goldman-Rakic, P.S. (1999). Auditory belt and parabelt projections in the prefrontal cortex in the rhesus macaque. *J. Comp. Neurol.* 403, 141–157. [https://doi.org/10.1002/\(sici\)1096-9861\(19990111\)403:2<141::aid-cne1>3.0.co;2-v](https://doi.org/10.1002/(sici)1096-9861(19990111)403:2<141::aid-cne1>3.0.co;2-v).
27. Hasson, U., Ghazanfar, A.A., Galantucci, B., Garrod, S., and Keysers, C. (2012). Brain-to-brain coupling: A mechanism for creating and sharing a social world. *Trends Cogn. Sci.* 16, 114–121. <https://doi.org/10.1016/j.tics.2011.12.007>.
28. Freiwald, W.A. (2020). Social interaction networks in the primate brain. *Curr. Opin. Neurobiol.* 65, 49–58. <https://doi.org/10.1016/j.conb.2020.08.012>.
29. Khanna, A.R., Muñoz, W., Kim, Y.J., Kfir, Y., Paulk, A.C., Jamali, M., Cai, J., Mustroph, M.L., Caprara, I., Hardstone, R., et al. (2024). Single-neuronal elements of speech production in humans. *Nature* 626, 603–610. <https://doi.org/10.1038/s41586-023-06982-w>.
30. Bögels, S., Magyari, L., and Levinson, S.C. (2015). Neural signatures of response planning occur midway through an incoming question in conversation. *Sci. Rep.* 5, 12881. <https://doi.org/10.1038/srep12881>.
31. Romanski, L.M., Averbeck, B.B., and Diltz, M. (2005). Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. *J. Neurophysiol.* 93, 734–747. <https://doi.org/10.1152/jn.00675.2004>.
32. Hage, S.R., and Nieder, A. (2013). Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. *Nat. Commun.* 4, 2409. <https://doi.org/10.1038/ncomms3409>.
33. Jafari, A., Dureux, A., Zanini, A., Menon, R.S., Gilbert, K.M., and Everling, S. (2023). A vocalization-processing network in marmosets. *Cell Rep.* 42, 112526. <https://doi.org/10.1016/j.celrep.2023.112526>.
34. Plakke, B., Ng, C.W., and Poremba, A. (2013). Neural correlates of auditory recognition memory in primate lateral prefrontal cortex. *Neuroscience* 244, 62–76. <https://doi.org/10.1016/j.neuroscience.2013.04.002>.
35. Grijseels, D.M., Prendergast, B.J., Gorman, J.C., and Miller, C.T. (2023). The neurobiology of vocal communication in marmosets. *Ann. N. Y. Acad. Sci.* 1528, 13–28. <https://doi.org/10.1111/nyas.15057>.
36. Parker, P.R.L., Abe, E.T.T., Leonard, E.S.P., Martins, D.M., and Niell, C.M. (2022). Joint coding of visual input and eye/head position in V1 of freely moving mice. *Neuron* 110, 3897–3906.e5. <https://doi.org/10.1016/j.neuron.2022.08.029>.
37. Sit, K.K., and Goard, M.J. (2023). Coregistration of heading to visual cues in retrosplenial cortex. *Nat. Commun.* 14, 1992. <https://doi.org/10.1038/s41467-023-37704-5>.
38. Bögels, S. (2020). Neural correlates of turn-taking in the wild: Response planning starts early in free interviews. *Cognition* 203, 104347. <https://doi.org/10.1016/j.cognition.2020.104347>.
39. Burkart, J.M., Adriaense, J.E.C., Brügger, R.K., Miss, F.M., Wierucka, K., and van Schaik, C.P. (2022). A convergent interaction engine: vocal communication among marmoset monkeys. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 377, 20210098. <https://doi.org/10.1098/rstb.2021.0098>.
40. Chow, C.P., Mitchell, J.F., and Miller, C.T. (2015). Vocal turn-taking in a non-human primate is learned during ontogeny. *Proc. Biol. Sci.* 282, 20150069. <https://doi.org/10.1098/rspb.2015.0069>.
41. Toarmino, C.R., Wong, L.A., and Miller, C.T. (2017). Audience affects decision-making in a marmoset communication network. *Biol. Lett.* 13, 20160934. <https://doi.org/10.1098/rsbl.2016.0934>.
42. Miller, C.T., Beck, K., Meade, B., and Wang, X. (2009). Antiphonal call timing in marmosets is behaviorally significant: interactive playback experiments. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 195, 783–789. <https://doi.org/10.1007/s00359-009-0456-1>.
43. Pomberger, T., Risueno-Segovia, C., Gultekin, Y.B., Dohmen, D., and Hage, S.R. (2019). Cognitive control of complex motor behavior in marmoset monkeys. *Nat. Commun.* 10, 3796. <https://doi.org/10.1038/s41467-019-11714-8>.
44. Takahashi, D.Y., Liao, D.A., and Ghazanfar, A.A. (2017). Vocal learning via social reinforcement by infant marmoset monkeys. *Curr. Biol.* 27, 1844–1852.e6. <https://doi.org/10.1016/j.cub.2017.05.004>.
45. Musall, S., Kaufman, M.T., Juavinett, A.L., Gluf, S., and Churchland, A.K. (2019). Single-trial neural dynamics are dominated by richly varied movements. *Nat. Neurosci.* 22, 1677–1686. <https://doi.org/10.1038/s41593-019-0502-4>.

46. Steinmetz, N.A., Zatka-Haas, P., Carandini, M., and Harris, K.D. (2019). Distributed coding of choice, action and engagement across the mouse brain. *Nature* 576, 266–273. <https://doi.org/10.1038/s41586-019-1787-x>.
47. Allen, W.E., Chen, M.Z., Pichamoorthy, N., Tien, R.H., Pachitariu, M., Luo, L., and Deisseroth, K. (2019). Thirst regulates motivated behavior through modulation of brainwide neural population dynamics. *Science* 364, 253. <https://doi.org/10.1126/science.aav3932>.
48. Maisson, D.J.-N., Cervera, R.L., Voloh, B., Conover, I., Zambre, M., Zimmermann, J., and Hayden, B.Y. (2023). Widespread coding of navigational variables in prefrontal cortex. *Curr. Biol.* 33, 3478–3488.e3. <https://doi.org/10.1016/j.cub.2023.07.024>.
49. Shaw, L., Wang, K.H., and Mitchell, J. (2023). Fast prediction in marmoset reach-to-grasp movements for dynamic prey. *Curr. Biol.* 33, 2557–2565.e4. <https://doi.org/10.1016/j.cub.2023.05.032>.
50. Mao, D., Avila, E., Caziot, B., Laurens, J., Dickman, J.D., and Angelaki, D.E. (2021). Spatial modulation of hippocampal activity in freely moving macaques. *Neuron* 109, 3521–3534.e6.
51. McMahon, D.B.T., Russ, B.E., Elnaiem, H.D., Kurnikova, A.I., and Leopold, D.A. (2015). Single-Unit Activity during Natural Vision: Diversity, Consistency and Spatial Sensitivity among AF Face Patch Neurons. *J. Neurosci.* 35, 5537–5548. <https://doi.org/10.1523/JNEUROSCI.3825-14.2015>.
52. Sonneborn, A., Bartlett, L., Olson, R.J., Milton, R., and Abbas, A.I. (2024). Divergent subregional information processing in mouse prefrontal cortex during working memory. Preprint at bioRxiv. <https://doi.org/10.1101/2024.04.25.591167>.
53. Flinker, A., Korzeniewska, A., Shestyuk, A.Y., Franaszczuk, P.J., Dronkers, N.F., Knight, R.T., and Crone, N.E. (2015). Redefining the role of Broca's area in speech. *Proc. Natl. Acad. Sci. U S A* 112, 2871–2875. <https://doi.org/10.1073/pnas.1414491112>.
54. Lin, J.L., Imada, T., Meltzoff, A.N., Hiraishi, H., Ikeda, T., Takahashi, T., Hasegawa, C., Yoshimura, Y., Kikuchi, M., Hirata, M., et al. (2023). Dual-MEG interbrain synchronization during turn-taking verbal interactions between mothers and children. *Cereb. Cortex* 33, 4116–4134. <https://doi.org/10.1093/cercor/bhac330>.
55. Ahn, S., Cho, H., Kwon, M., Kim, K., Kwon, H., Kim, B.S., Chang, W.S., Chang, J.W., and Jun, S.C. (2018). Interbrain phase synchronization during turn-taking verbal interaction—a hyperscanning study using simultaneous EEG/MEG. *Hum. Brain Mapp.* 39, 171–188. <https://doi.org/10.1002/hbm.23834>.
56. Miller, E.K., and Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. <https://doi.org/10.1146/annurev.neuro.24.1.167>.
57. Martinez-Trujillo, J. (2022). Visual attention in the prefrontal cortex. *Annu. Rev. Vis. Sci.* 8, 407–425. <https://doi.org/10.1146/annurev-vision-100720-031711>.
58. Miller, E.K. (2000). The prefrontal cortex and cognitive control. *Nat. Rev. Neurosci.* 1, 59–65. <https://doi.org/10.1038/35036228>.
59. Miller, C.T. (2017). Why Marmosets? *Dev. Neurobiol.* 77, 237–243. <https://doi.org/10.1002/dneu.22483>.
60. Schiel, N., and Souto, A. (2017). The common marmoset: An overview of its natural history, ecology and behavior. *Dev. Neurobiol.* 77, 244–262. <https://doi.org/10.1002/dneu.22458>.
61. Miller, C.T., and Wren Thomas, A. (2012). Individual recognition during bouts of antiphonal calling in common marmosets. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 198, 337–346. <https://doi.org/10.1007/s00359-012-0712-7>.
62. Miller, C.T., Iguina, C.G., and Hauser, M.D. (2005). Processing vocal signals for recognition during antiphonal calling in tamarins. *Anim. Behav.* 69, 1387–1398. <https://doi.org/10.1016/j.anbehav.2004.08.021>.
63. Chari, T., and Pachter, L. (2023). The specious art of single-cell genomics. *PLOS Comput. Biol.* 19, e1011288. <https://doi.org/10.1371/journal.pcbi.1011288>.
64. Lause, J., Kobak, D., and Berens, P. (2024). The art of seeing the elephant in the room: 2D embeddings of single-cell data do make sense. Preprint at bioRxiv. <https://doi.org/10.1101/2024.03.26.586728>.

## STAR★METHODS

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Behavioral and neural data	This paper	Dryad: <a href="https://doi.org/10.5061/dryad.c2fqz61hj">https://doi.org/10.5061/dryad.c2fqz61hj</a>
Experimental models: Organisms/strains		
Marmoset	University of California San Diego	N/A
Software and algorithms		
Matlab 2021a	MathWorks	<a href="https://www.mathworks.com/products/new_products/release2021a.html">https://www.mathworks.com/products/new_products/release2021a.html</a>
Code	This paper	Dryad: <a href="https://doi.org/10.5061/dryad.c2fqz61hj">https://doi.org/10.5061/dryad.c2fqz61hj</a>

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

## Subjects

Experiments in this work were done with six adult common marmosets (*Callithrix jacchus*).<sup>59,60</sup> Monkey E, F, H and M were female and monkey B and R were male. All subjects were group housed and were at least 1.5 years old at time of implant. Monkey B had bilateral arrays implanted and all the other monkeys had a single array implanted. All experiments were performed in the Cortical Systems and Behavior Laboratory at University of California, San Diego (UCSD) and approved by the UCSD Institutional Animal Care and Use Committee.

## METHOD DETAILS

## Neurophysiology

The surgical procedure of the chronic microelectrode array implantation is described in our previous work (subjects H, E and M<sup>15</sup>, and subjects F, B and R<sup>25</sup>). The microelectrode array consists of 16 channel tungsten electrodes each housed in an independent guide tube in a 4\*4 mm grid (Neuralynx, Bozeman, MT). Each array was positioned on the surface of the brain in the chronic implant, and perpendicularly lowered to the laminar surface of neocortex with a calibrated Warp Drive pusher attached to the end of the guide tubes. The electrode was advanced 10-20  $\mu\text{m}$  twice a week over the course of the experiment sessions. The headstage preamplifier was connected to the arrays and attached to a tether to allow subjects to freely move in the experiment box. The tether was wrapped by a metal coil to prevent interference from the movement and vocalization of the subjects. Spike sorting was done off-line manually. First, a 300-9000 Hz filter and thresholding was applied across the entire recording session. Next, PCA was applied on the waveform with a 1 ms window and DBSCAN was used to cluster the waveforms automatically in the space of first 3 PCs. Last, units are selected with the criteria of signal-to-noise ratio (SNR)  $\geq 13$  dB and less than 1% inter-spike intervals violate the refractory period  $< 1$  ms. Overall, 335 single units met the criteria and were studied in this work. A detailed number of neurons and their recording site is presented in Table S1 and Figure S3. One of the limitations in the current work lays in the number of neurons simultaneously recorded in each session, where less than 3 neurons are simultaneously recorded in some sessions. It further leads to a lack of analyses and computational models to investigate dynamical structure in the population. Future work should focus on recording a larger number of neurons in the population simultaneously with high density multi-electrode arrays as well as exploiting models including population dynamics, such as GLM with coupling filters and state space models.

## Experiment design

The recordings took place in a 4x3 m test room. Subjects were placed in a 32x18x46 cm test cage that allows them to freely behave and engage in natural vocal communication behaviors, as in previous behavioral and neurophysiological experiments.<sup>15,61,62</sup> We employed an interactive playback paradigm to reproduce the natural vocal behavior of a marmoset and engage a live subject in antiphonal calling exchanges. The design of this paradigm is based on the natural statistical structure of marmoset turn-taking during antiphonal exchanges<sup>19,20,42</sup> and has been used frequently in our lab to investigate the perceptual, cognitive and neural processes underlying this communication behavior.<sup>15,16,41,42,61</sup> Briefly, when subjects produce a phee call, a responsive phee call was broadcasted in response to the subject's phee call within 2-4 s after the subjects' phee call offset. If subjects produce no phee calls for 45-60 s, a phee call was broadcasted to initiate a vocal interaction. Phee calls from subjects within 10 s after a broadcasted phee call are classified as a conversational response; subjects' phee calls outside the time window are classified as spontaneous phee calls. Duration of each session is typically 60-160 minutes (see Table S1). Detailed description of the behavioral paradigm in

this work can also be found in the previous work for Jovanovic et al.,<sup>15</sup> Nummela et al.,<sup>16</sup> and Miller et al.<sup>25</sup> In the current experiment model, we quantified only vocal related behaviors, such as call production, while behaviors not directly related to natural communication were omitted, such as head-direction and posture. Because sessions were not recorded by video, these non-vocal behaviors could not be quantified. A more naturalistic and integrative experiment recording other facets of behavior along with vocal behavior is encouraged in future work to further examine the neural basis of those facets in natural communication.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### PSTH analysis

PSTHs were plotted using 0.1 s time bins with a boxcar average filter on 15 data points. The PSTHs without smoothing filter were shown in [Figure S1A](#). The significance of neural responses in PSTH was tested on the firing rate during 2 s pre-stimulus time versus firing rate during stimulus with the Wilcoxon signed rank test at 0.05 significance level.

### GLM based analysis

The GLM based analysis is built upon the model:

$$y(t) \sim \text{Poisson} \left( \exp \left( \vec{k} * \vec{x}(t) + \vec{h} * \vec{y}_{\text{hist}}(t) \right) \right),$$

where  $y(t)$  is spike count at time  $t$ ,  $y_{\text{hist}}$  is the history of spiking up to time  $t$ ,  $x$  represents behavioral events and internal behavioral states,  $h$  and  $k$  represent the kernels or coefficients for the events or states, respectively. Behavioral events include hearing a call, producing a call in response to hearing a conspecific call, and producing a spontaneous call; internal behavioral states include conversational state, defined as the period that the marmoset is engaged in vocal interactions, and spontaneous state, defined as 8 s before the marmoset emits a spontaneous call. Justifications of the model design and parameters are explored with alternative models and behavioral analysis as shown and discussed in [Figure S1](#). It is noteworthy that in the current work, we defined the internal states as a fixed period of time according to the statistics of vocal behavior even though the stochastic nature of the monkey's vocal behavior shows no clear thresholds ([Figure S1D](#)). Future modelling work can address this limitation by combining GLM and Hidden Markov Models to exploit varying internal state time windows instead of a fixed time window, in order to better capture the stochastic nature of the marmoset vocal behavior.

The spike trains and behavioral data were discretized into 0.01 s bins. Thus, spike trains were converted to spike count time series, and behavior events and states were converted to binary time series as input to the model. A sliding window was used to collect samples at different time  $t$ . We used the canonical exponential nonlinearity for the activation function  $f$  and obtained kernels and coefficients by maximizing the log posterior likelihood with a smoothing prior:

$$\mathcal{L}(\omega) = \log P(s|\omega) - \lambda \sum (\omega_i - \omega_{i-1})^2,$$

where  $s$  is the spike rate in data,  $w$  represents the kernel coefficients, and  $\lambda$  is the penalty parameter. Notably, the regularization term penalizes differences in neighboring kernel coefficients and encourages smoothness of the kernel. This formulation is equivalent to a choosing a prior over the differences in neighboring coefficients as independent Gaussian random variables with zero mean and variance  $1/\lambda$ . For each neuron,  $\lambda$  was determined by fivefold cross-validation ([Figure S4](#)). Significance of the kernels and coefficients were determined by permutation test at 0.05 significance level. The null kernels and coefficients in the permutation test for each neuron were obtained by shuffling the alignment of spike trains and behavioral data 1000 times. This null model breaks the temporal association between spikes and behavior without altering the temporal autocorrelations of the time series. For the kernels, the permutation test compares the peak of each kernel with its corresponding null distribution; for the coefficients, the permutation test is performed on the coefficient and null distribution of coefficients.

### Clustering analysis

The purpose of this analysis was to use an unsupervised method for profiling the functional characteristics of the neural population. We therefore used a clustering analysis following nonlinear dimensionality reduction to determine if a consistent set of functional clusters would be evident. Clustering analysis of the population was applied on the neurons with at least 1 significant kernel or state coefficient. For each neuron, we took the projections on the first 3 PCs of the speaker call kernel, the first 4 PCs of the two marmoset call kernels, the conversational and the spontaneous coefficients to form a 13-dimensional space. The number of PCs used is determined by 80% variance explained by the PCs. We found that higher numbers of included PCs obscured the shared features among kernels and diminished clustering performance, while fewer PCs omitted important features to capture clustering structure. UMAP was applied to the PC loadings for nonlinear dimensionality reduction with two latent dimensions. Since our goal was to identify a stable set of functional clusters, and not to identify global relationships between clusters, we were not concerned with recently demonstrated failures of UMAP to preserve pairwise distances<sup>63</sup> (for a rejoinder see Lause et al.<sup>64</sup>).

A Gaussian Mixture Model (GMM) with 10 components was applied to the two UMAP dimensions. For each neuron and run of UMAP, fitting the GMM gives a posterior probability of cluster memberships. To account for variability in cluster identity and global cluster geometry across runs of UMAP, we developed an algorithm to match cluster identity across runs. Our rationale was that if

cluster membership across runs was consistent then we could reasonably claim that the neural population was in fact organized into functional clusters while if cluster membership was inconsistent then the existence of clusters in any single UMAP realization could be regarded as illusory. We first globally aligned UMAP coordinates across runs by minimizing the Procrustes distance between UMAP runs. We then aligned corresponding cluster identities between runs by minimizing the Wasserstein-2 distance between clusters and then averaged posterior cluster membership probabilities across all 200 UMAPs. This procedure is equivalent to approximate marginalization across the ensemble of UMAP runs. We hard-assigned a neuron to a cluster if there existed one dominant cluster identity (average posterior probability  $> 0.5$ ). Otherwise, the neuron was considered as “transitional” which lies in between clusters. To test the robustness of the clustering analysis, intra-cluster and inter-cluster distances were defined as the average distance of the coordinates within a cluster and the distance between the centers of two clusters, respectively.

**Decoding**

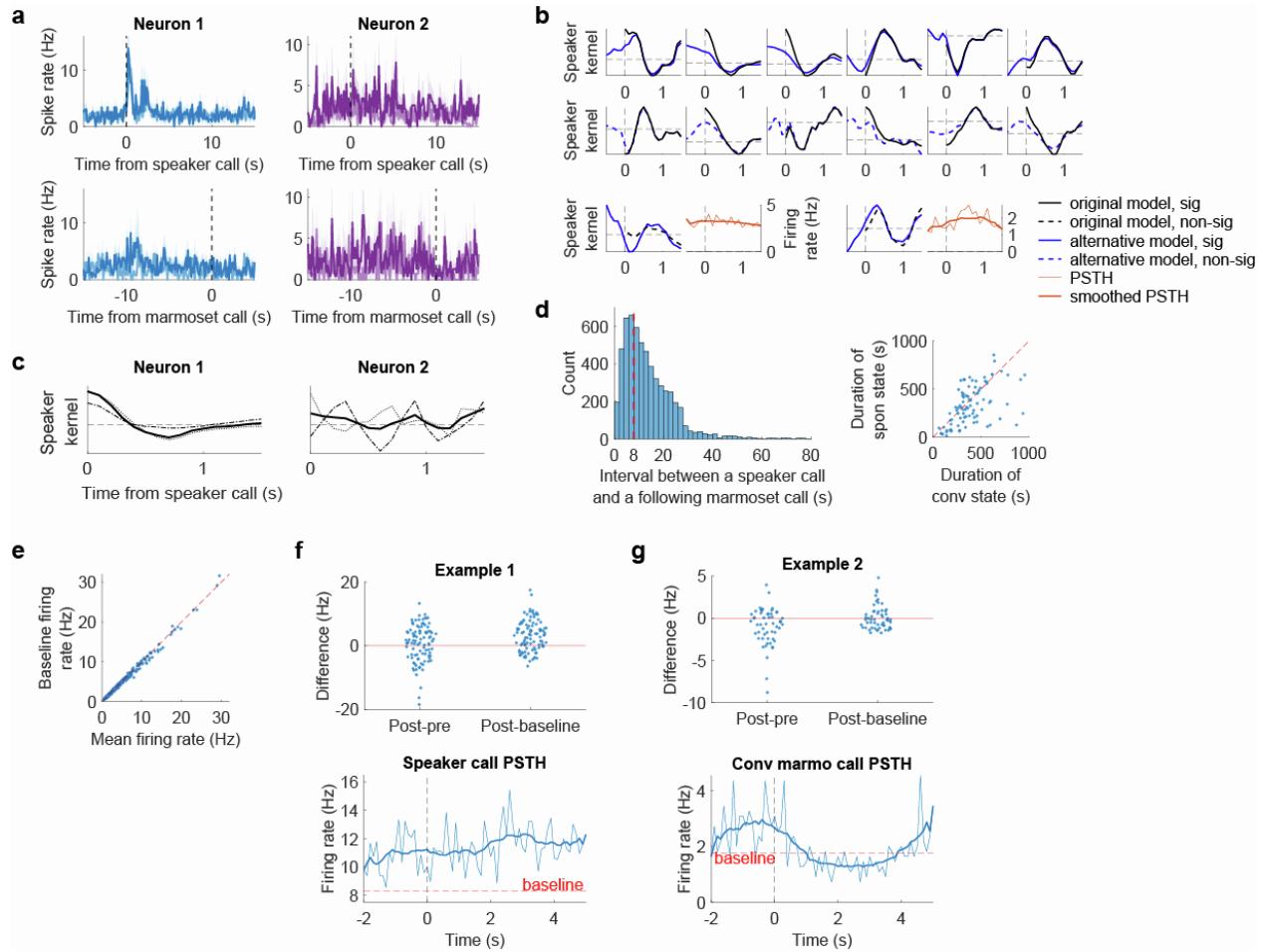
We defined a conversational call score (CCS) to represent how likely a call is produced by the marmoset after hearing a call. At each instance a conspecific call was heard, we fit 2 models; one where we assumed that the marmoset produced a call with 4 s duration at time  $t$  after the offset of the conspecific call was heard, and one where no call was produced. Each version of the model was fit with a GLM which yielded 2 series of log-likelihood values for  $t$  from 0-10 s or until another conspecific call was broadcasted to the marmoset with 0.1 s time bin. The CCS was defined as the maximum log-likelihood-ratio of the two models at the time after the conspecific call offset. Ground truth of the marmoset vocal behavior was used to calculate ROC and AUC of each neuron’s decoding performance. Significance of the decoding performance were determined by permutation test at 0.05 significance level. The null decoding performance in the permutation test was obtained by shuffling the alignment of spike trains and behavioral data 1000 times.

**Neuron, Volume 112**

**Supplemental information**

**Representing the dynamics of natural marmoset  
vocal behaviors in frontal cortex**

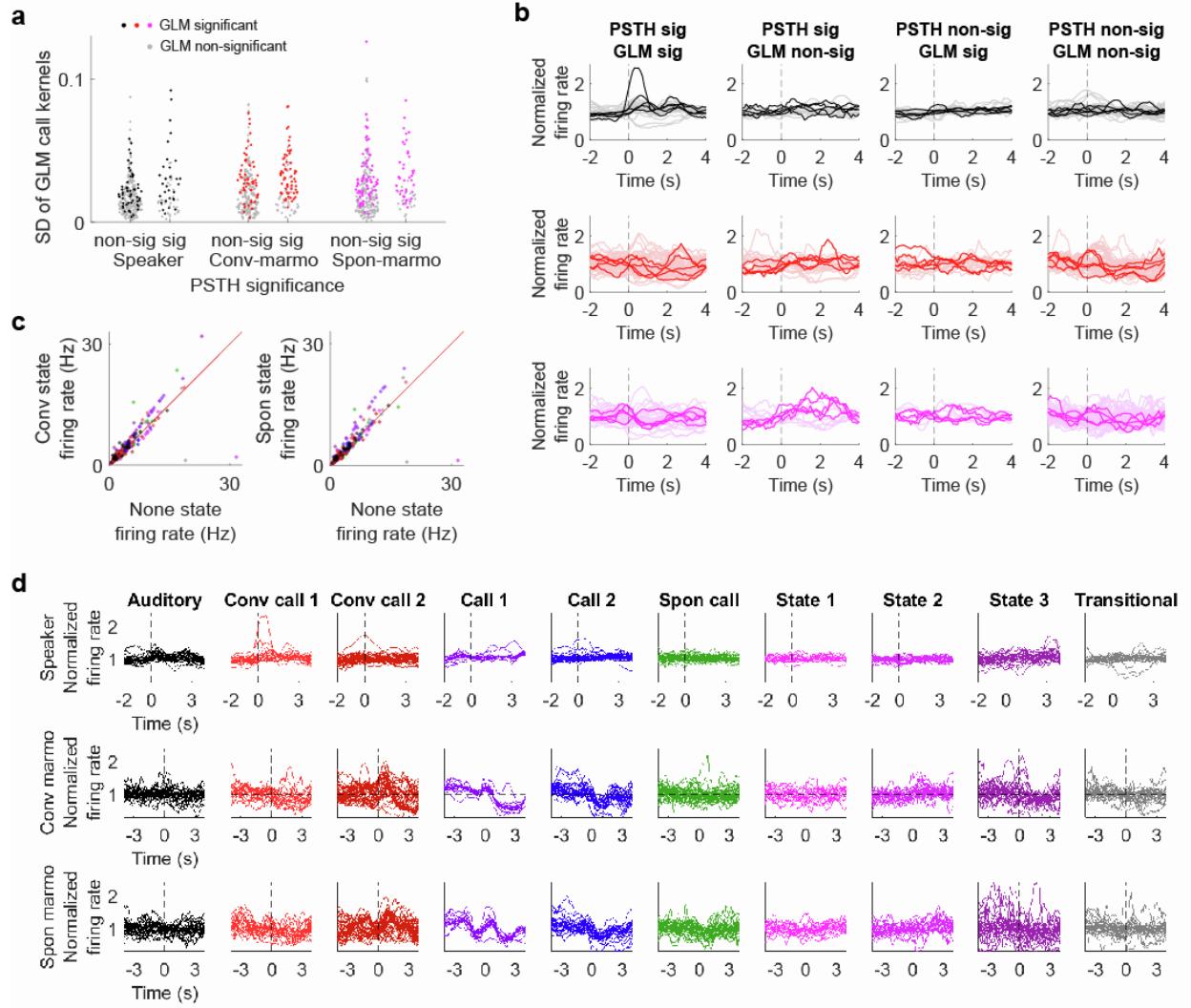
**Jingwen Li, Mikio C. Aoi, and Cory T. Miller**



**Figure S1: Justification of the GLM model design and parameters with alternative GLM models and behavioral analysis; justification of the PSTH analysis with alternative PSTH analysis based on ‘baseline’ firing rate, related to Figures 1 and 2.**

**a)** PSTH curves same as in Fig. 1d for example neuron 1 and 2 but without smoothing filter. For neuron 1, the firing rate raised sharply after the speaker call in the PSTH, while in Fig. 1d the smoothing filter makes the firing rate rising before the speaker call onset. It is noteworthy that the smoothing filter does not affect the significance test of neural response in PSTH. **b)** An alternative GLM model including 0.5 s before the speaker calls in the speaker kernel (blue) compared to the speaker kernel in the original GLM model (black). Top row: 6 example neurons where the speaker kernels are both significant in the two GLM models (52 out of 81 neurons). Middle row: 6 example neurons where the speaker kernel is significant in the original GLM model but insignificant in the alternative GLM model (29 out of 81 neurons). Bottom row: 2 example neurons where the speaker kernel is significant in the alternative GLM model but insignificant in the original GLM model (19 out of 335 neurons), plotted along with the speaker call aligned PSTH. While the speaker kernels in the two models are mostly consistent with each other after the speaker call onset, it is more often that the alternative model including extra time ahead of the speaker call results in a more moderate kernel before and at the beginning of the speaker call, and thus making the original significant kernel insignificant. This is due to the penalty for the smoothing prior (see Methods) if the neuron is not responsive to the speaker call ahead of the stimulus onset. **c)** An alternative GLM model with conversational (dashed line) and non-conversational (dotted line) speaker kernels separated, compared the speaker kernel in the

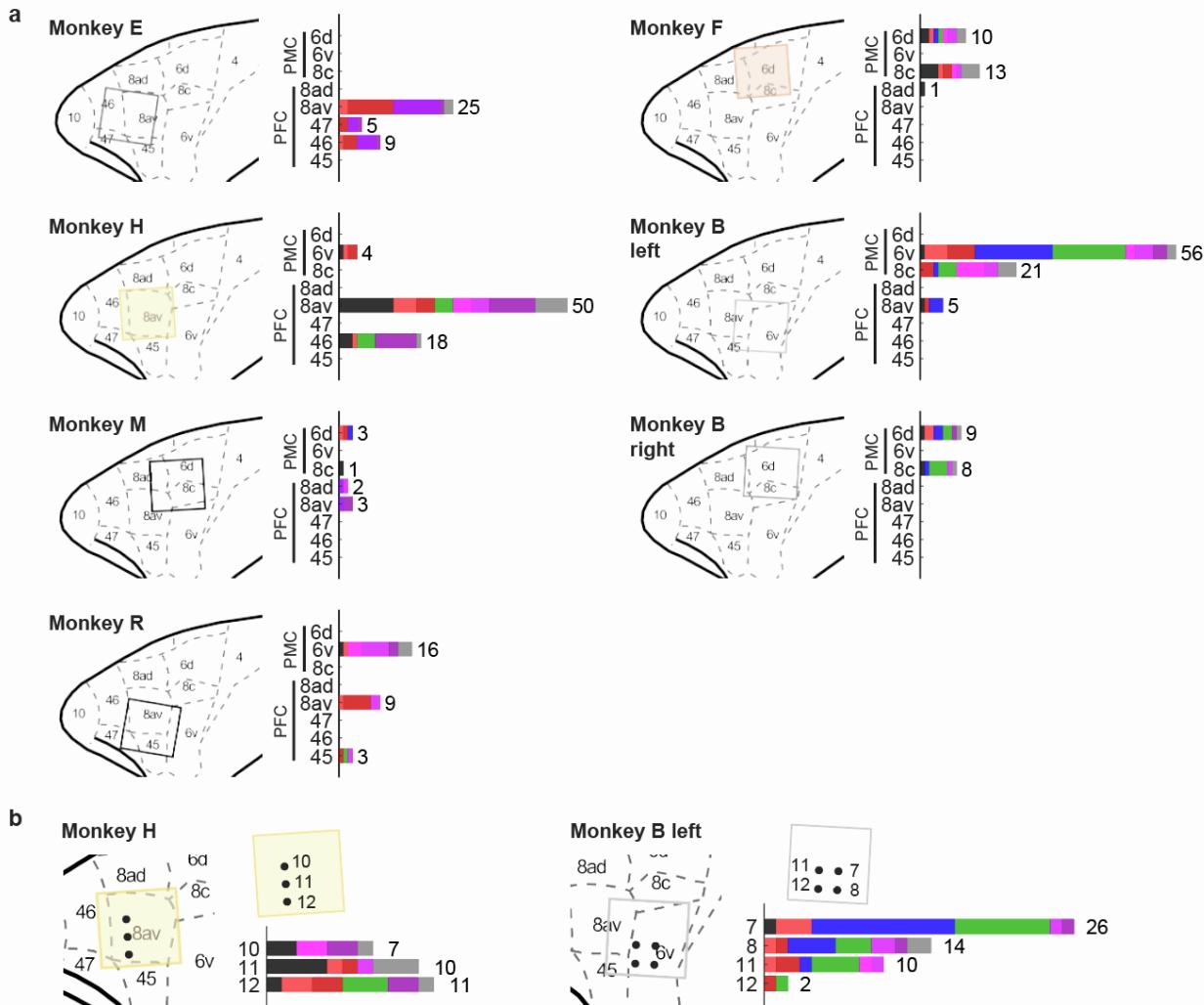
original GLM model (solid line) for the two example neurons in Fig. 1d. For neuron 1 where the speaker kernel is significant, the two types of the speaker kernels are consistent with the combined speaker kernel. For neuron 2 where the speaker kernel is insignificant, the two types of the speaker kernels may vary. Thus, we constructed the GLM model with one speaker kernel considering data efficiency. **d)** Left: the distribution of interval between a speaker call and a following marmoset call. The wide distributed range and the large variance reveals the stochastic nature of the marmoset vocal behavior. The distribution starts decaying after 8s, indicating a decreasing probability of staying in an internal state longer than 8s. Therefore, even though there is not a clear threshold to define an internal state period, we set the spontaneous state to be 8s time window before a spontaneous marmoset call in the current manuscript. **e)** The comparison between the mean firing rate across the whole session and the ‘baseline’ firing rate for each neuron. In the manuscript, we quantified the PSTH significance by comparing the pre- and post- stimulus firing rate. An alternative way of quantifying PSTH significance is to compare the post-stimulus firing rate with respect to a baseline firing rate. The difficulty lies in defining the baseline firing rate in these naturalistic experimental paradigms, since the monkeys started vocal behavior once they entered the experimental arena and there is not a firing rate ‘before trials’. Here we made an attempt of defining the ‘baseline’ firing rate as the firing rate during period when the animal is not in any external events or internal states and compared with the PSTH analysis in the manuscript. The comparison shows that the PSTH analysis in the manuscript better captures the neural response to the external events. The ‘baseline’ firing rate is highly correlated to the mean firing rate. **f)** The comparison of the two PSTH analyses for an example neuron in speaker call PSTH. Left: the difference between post- and pre- stimulus firing rate versus the difference between post stimuli and ‘baseline’ firing rate. Each dot represent data from a speaker call throughout the session. The former was tested insignificant while the latter was tested significant with the Wilcoxon signed rank test at 0.05 significance level. Right: the speaker call PSTH (thin blue line) and smoothed PSTH (thick blue line) of the neuron. Both pre- and post- stimulus firing rate are systematically higher than the ‘baseline’ firing rate (red dashed line). This might be due to the overall higher firing rate when the animal is active and aroused during conversations instead of the neuron responsive to hearing a speaker call. **g)** Same as in **f** for another example neuron in conversational marmoset call PSTH. The difference between post- and pre- stimulus firing rate yields significance, whereas the difference between ‘baseline’ and post stimuli firing rate does not. From the PSTH, a suppression after emitting a call is observed and more apparently revealed in the post- and pre- stimulus firing rate difference.



**Figure S2: Comparison of PSTH and GLM significance and general dynamics of neural activity in different clusters identified from GLM, related to Figures 2 and 3.**

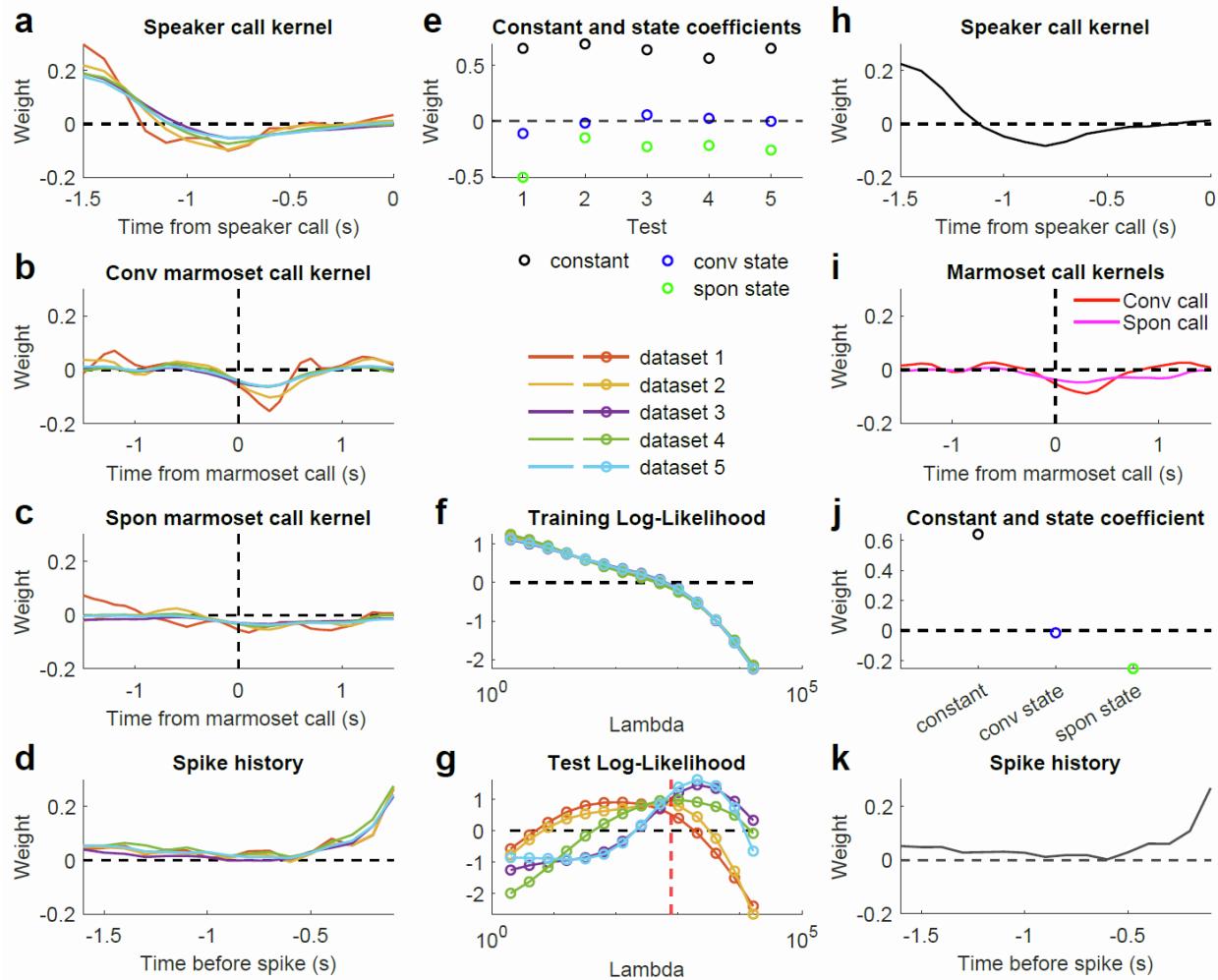
**a)** Standard deviation of GLM kernels for speaker call, conversational marmoset call, and spontaneous marmoset call in comparison with PSTH significance. GLMs identified significant neurons responding to behavioral events that are not significant in PSTHs. **b)** Normalized PSTHs in each category: significant in PSTH and significant in GLM; non-significant in PSTH and significant in GLM; significant in PSTH and non-significant in GLM; non-significant in PSTH and non-significant in GLM. Five examples in each category are highlighted for better visualization by larger linewidth. The top, middle, and bottom rows are speaker call, conversational marmoset call, and spontaneous marmoset call PSTHs, respectively. The first category, the neurons response with a certain curve after the event onset, making them both significant in PSTH analysis and GLM kernels. In the second category, the neurons tend to have an overall change across the event window; this overall change may be counted in the state coefficients instead of the event kernels in the GLM model. In the third category, the PSTHs tend to have changes that are not revealed in pre- and post- stimulus mean firing rate. The fourth category contains neurons with noise-like PSTHs, which are non-significant in both PSTH and GLM analysis. **c)** Conversational state firing rate (left) and spontaneous state firing rate (right) as a function of the none state firing rate colored by clusters corresponding to Fig. 3. Clusters are not distinguished from each other as the firing rate during these states are highly correlated.

The only cluster that stands out is a call cluster (purple) that has firing rate during conversational or spontaneous state higher than the none state firing rate. Combining with the GLM kernels, this might come from the enhancement in the marmoset call kernels before the call onset, which falls into the period of states in firing rate calculation. **d)** Normalized PSTH for speaker call (top row), conversational marmoset call (middle row), and spontaneous marmoset call (bottom row) of the neurons grouped by clusters. Neurons in the auditory cluster, two conversational call clusters, two call clusters and spontaneous call cluster show common features in corresponding PSTHs within each cluster. However, the state clusters and the transitional cluster do not show certain features in PSTHs.



**Figure S3: The location of each array and the number of neurons recorded in each brain area and functional cluster, related to Figure 3.**

**a)** plots the functional clusters for each area of frontal cortex in each subject. **b)** plots the functional clusters for single channels in two separate monkeys. Square above color bars indicate the location and number of the individual electrode channel for the array highlighted in the subfigure. **both)** The color of arrays corresponds to the same color in Fig. 3. The color of the bar plots indicates functional clusters same as in Fig. 3. Squares on each frontal cortex schematic drawing indicate the placement of the specific array. The number on the right indicates the total number of neurons in each area from the array.



**Figure S4: Cross-validation in GLM for an example neuron, related to Methods.**

**a-e)** Kernels and coefficients for fivefold cross-validation. **f)** Log-likelihood obtained from GLM in each training set as a function of the penalty parameter lambda. The Log-likelihood drops as lambda increases. **g)** Log-likelihood obtained from GLM in each test set as a function of the penalty parameter lambda. The Log-likelihood reaches an optimal value at different lambda for different test set. The optimal lambda (red dash line) is chosen at the logarithmic average of the five datasets. **h-k)** Kernels and coefficients obtained with the optimal lambda for the whole dataset.

<b>Monkey and recording site</b>	<b>Number of sessions</b>	<b>Duration of sessions</b>	<b>Number of units collected</b>	<b>Number of units with significant GLM kernels/coefficients</b>
<b>Monkey E</b>	1	80 min	42	39
<b>Monkey H</b>	20	60-90 min	104	72
<b>Monkey M</b>	6	70-120 min	11	9
<b>Monkey R</b>	14	120-160 min	34	28
<b>Monkey F</b>	12	60-120 min (1 session 46 min)	32	24
<b>Monkey B left</b>	34	120-160 min (1 session 174 min)	92	82
<b>Monkey B right</b>	7	120 min	20	17

**Table S1. Information of sessions and neurons for each recording site.** For the location of each array and the number of neurons recorded in each brain area and cluster, see Fig. S5.