Title

Invariant neural responses for sensory categories revealed by the time-varying information for communication calls

Journal

PLOS Computational Biology, 15(9)

ISSN

1553-734X

Authors

Elie, Julie E

Theunissen, Frédéric E

Publication Date

2019

DOI

10.1371/journal.pcbi.1006698

Peer reviewed

PLOS COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Invariant neural responses for sensory categories revealed by the time-varying information for communication calls

Julie E. Elie[1,2]*, **Frédéric E. Theunissen**[1,3]

1 Helen Wills Neuroscience Institute, University of California Berkeley, Berkeley, California, United States of America, **2** Department of Bioengineering, University of California Berkeley, Berkeley, California, United States of America, **3** Department of Psychology, University of California Berkeley, Berkeley, California, United States of America

* julie.elie@berkeley.edu

OPEN ACCESS

Citation: Elie JE, Theunissen FE (2019) Invariant neural responses for sensory categories revealed by the time-varying information for communication calls. PLoS Comput Biol 15(9): e1006698. https://doi.org/10.1371/journal.pcbi.1006698

Editor: Stephen V David, Oregon Health & Science University, UNITED STATES

Received: December 4, 2018

Accepted: June 8, 2019

Published: September 26, 2019

Data Availability Statement: The custom Matlab code used to calculate time varying information values is available at https://github.com/julieelie/PoissonTimeVaryingInfo, along with a tutorial on modeled data. All neural recordings are available from the CRCNS database (Elie J E and Theunissen F E (2019), Simultaneous extracellular recordings of avian auditory neurons in zebra finches presented with all the repertoire of vocalizations used by this species for vocal communication. CRCNS.org; http://dx.doi.org/10.6080/K00C4T06). Audio recordings of the vocalizations used in that study are available on http://bit.do/ZebeVocBank.

Abstract

Although information theoretic approaches have been used extensively in the analysis of the neural code, they have yet to be used to describe how information is accumulated in time while sensory systems are categorizing dynamic sensory stimuli such as speech sounds or visual objects. Here, we present a novel method to estimate the cumulative information for stimuli or categories. We further define a time-varying categorical information index that, by comparing the information obtained for stimuli versus categories of these same stimuli, quantifies invariant neural representations. We use these methods to investigate the dynamic properties of avian cortical auditory neurons recorded in zebra finches that were listening to a large set of call stimuli sampled from the complete vocal repertoire of this species. We found that the time-varying rates carry 5 times more information than the mean firing rates even in the first 100 *ms*. We also found that cumulative information has slow time constants (100–600 *ms*) relative to the typical integration time of single neurons, reflecting the fact that the behaviorally informative features of auditory objects are time-varying sound patterns. When we correlated firing rates and information values, we found that average information correlates with average firing rate but that higher-rates found at the onset response yielded similar information values as the lower-rates found in the sustained response: the onset and sustained response of avian cortical auditory neurons provide similar levels of independent information about call identity and call-type. Finally, our information measures allowed us to rigorously define categorical neurons; these categorical neurons show a high degree of invariance for vocalizations within a call-type. Peak invariance is found around 150 ms after stimulus onset. Surprisingly, call-type invariant neurons were found in both primary and secondary avian auditory areas.

Author summary

Just as the recognition of faces requires neural representations that are invariant to scale and rotation, the recognition of behaviorally relevant auditory objects, such as spoken

Funding: This study was funded by a National Institutes of Health grant to FET (NIDCD R01 0167783). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

words, requires neural representations that are invariant to the speaker uttering the word and to his or her location. Here, we used information theory to investigate the time course of the neural representation of bird communication calls and of behaviorally relevant categories of these same calls: the call-types of the bird's repertoire. We found that neurons in both the primary and secondary avian auditory cortex exhibit invariant responses to call renditions within a call-type, suggestive of a potential role for extracting the meaning of these communication calls. We also found that time plays an important role: first, neural responses carry significantly more information when represented by temporal patterns calculated at the small time scale of 10 ms than when measured as average rates and, second, this information accumulates in a non-redundant fashion up to long integration times of 600 ms. This rich temporal neural representation is matched to the temporal richness found in the communication calls of this species.

Introduction

Information theoretic analyses are well suited to the study of neural representation since this mathematical framework was developed to quantify and optimize the encoding of informative signals in communication channels [1]. In sensory systems, Information Theory (IT) has been applied extensively as a complementary approach to the estimation of stimulus-response functions such as tuning curves, spatio-temporal or spectro-temporal receptive fields or other higher-level encoding models [2]. Information theoretic approaches have been particularly powerful in explorations of the nature of the neural code and its redundancy or efficiency [3–6]. For example, IT was used in early studies in the visual system to demonstrate that spike patterns contain information beyond average rate both for static images [7] and dynamic visual stimuli [8]. IT was also used to show that spike doublets can contain synergistic information that cannot be explained by an analysis of successive single spikes [9] and that, although information can only decrease in a signal processing chain, the neural coding efficiency increases as one moves to higher levels of sensory processing [10]. Finally, IT investigations also revealed that neural efficiency is higher when sensory systems process natural stimuli versus synthetic stimuli [11–13], in support of ethological theories of optimal sensory processing [14].

In sensory systems, the mutual information between a stimulus and the neural responses has often been estimated in a stimulus reconstruction framework and for continuous dynamic stimuli in stationary conditions, where time averages can be performed. In the stimulus reconstruction framework, one attempts to estimate the information about all aspects of the stimulus; for example, in audition, the stimulus would be represented by its exact sound pressure waveform. As long as the stimulus set is rich (i.e. has very large entropy), the mutual information can be an estimate of the maximum information that can be transmitted by a neural communication channel, also known as the channel capacity [3, 4]. For instance, one can obtain the mutual information of an adapted auditory neuron processing white noise or colored noise sounds [11]. As long as the stationary assumption is valid, using continuous stimuli is also beneficial as it provides large data sets that are needed to estimate the joint probability of stimuli and neural responses, both of which can have high dimensions. Even in these conditions, it is noteworthy that a direct estimation of information is only possible when many repeats of the same stimulus can be obtained [15] or when simplifying assumptions are made [9]. Ultimately, the calculation of information based on stimulus reconstruction gives a single number corresponding to the information transmitted by a single neuron or an ensemble of neurons for a particular stimulus ensemble. By repeating the calculation for different stimulus

ensembles, one can investigate how the channel capacity of particular neurons or neural ensembles might depend on the stimulus statistics (e.g. for natural vs synthetic stimuli). Furthermore, by repeating the calculation using different symbols to represent the response, the potential nature of the neural code (e.g. time patterns vs. rate) can be revealed.

Here we are using IT in a different sensory encoding context: the accumulation of information in a recognition task, such as face recognition in the visual system [16] or word recognition in the auditory system [17]. Recognition or identification is one of the key computations performed by higher sensory areas as opposed to the task of efficient stimulus representations that is performed in lower sensory areas and that might therefore be well quantified by information values based on stimulus reconstruction. In the recognition tasks, each stimulus is described by a simple label, such as the word corresponding to a given speech sound or that is used to label a given visual object. The relevant value of information in that task is then the information about these discrete labels and the information capacity of the system in its ability to identify the stimulus as a whole. In the recognition framework, one can ask how the information about the stimulus identity or label changes as a function of time relative to the stimulus onset and to what extent that time-varying information is redundant and, thus, how it accumulates over time. For example, one could ask at what time after stimulus onset does the performance of single neurons or ensemble of neurons match a behavioral performance of word recognition. Such an IT analysis has been performed in the primate visual system using a delay-matching to sample paradigm, and using spike counts, estimated in progressively longer windows, as the neural symbols [18].

In information studies based on continuous stimulus reconstruction, the neural code can be investigated in terms of its temporal resolution (i.e. letter size) and its integration time (i.e. word length). While the same properties of the neural code can be deciphered in the recognition framework, one can also examine the relationship between spikes at different points in time and time-varying information. This analysis is meaningful because a time zero corresponding to stimulus onset can be clearly defined and is behaviorally relevant. Moreover, stimulus-response functions for such discrete stimuli are not time-invariant. Responses in sensory neurons, in vision [19, 20] and in audition [21, 22], are often characterized by an onset response (or on-response) and a sustained response, where both the precision of spikes and the information coded might be different [23, 24]. For example, a first spike latency code has been proposed as a fast encoding scheme in vision [25], audition [26] and somato-sensation [27]. Rolls et al. tested this hypothesis, by quantifying the fraction of information that is present in the first spike relative the on-going response [28].

Finally, in the recognition framework, one can also compare the information values obtained when different labelling schemes are used for identifying the stimuli as objects. For example, speech sounds could be labeled hierarchically as unique utterances, as phonemes, as syllables, as words, etc. One can then compare time-varying information about each of the levels in such hierarchical labelling scheme and gain insight on the neuro processing involved in object categorization. Although such a hierarchical representation of stimulus features has been used in encoding models for studying human processing [29], it has not yet been used in an IT analysis.

In this study, we developed a new approach for estimating time-varying information and cumulative information for sensory object identification task. Our approach assumes that time-varying neural responses can be modeled as inhomogeneous Poisson processes and generalizes well to large number of stimulus categories and to long integration times relative to the dynamics of the time-varying response. Our motivation for developing this methodology was to gain additional understanding on the neural representation of communication signals in high level auditory areas. Animal communication calls, just as speech sounds in

humans, are categorized into behaviorally meaningful units. Significant progress has been made in identifying brain regions involved in categorizing sounds, in particular in the primate brain, where neural responses that are correlated with progressively more abstracts concepts are found in primary auditory cortex, the lateral belt of the auditory cortex and the prefrontal cortex [30]. However, the neural computations involved in generating categorical responses remain poorly described [31] and only a small number of studies have examined the neural categorization of natural communication calls in non-human species [32–36]. We and others have been developing an avian model system to study the neural processing of relatively large and complex vocal repertoires [37, 38]. Our prior studies include a detailed bioacoustical analysis of the features that define each call-type of the complete vocal repertoire of the zebra finch [39] and the first characterization of neural responses to the calls from that large repertoire in primary and secondary avian auditory cortical areas [40]. In that study, we found that approximately 45% of auditory neurons encode information about call-type categories. Among those, a minority show strong selectivity for single call-type categories and invariance for calls within that category. Here, we investigated the processing in time that could lead to those observed categorical responses by comparing the time-varying information for stimuli labelled as individual utterances to the time-varying information for the same stimuli labelled by their call-type category. With that analysis, we were able to obtain values of temporal integration for stimulus identification and call-type category identification. We also analyzed the relationship between the time-varying firing rate and the time-varying information and, in particular, examined differences in selectivity in the onset versus sustained response. Finally, we used anatomical data to examine the distribution of neurons in primary and secondary avian auditory cortical areas with distinct responses properties as revealed by this IT analysis.

Results

We studied the time-varying information in a population of neurons recorded from primary and secondary regions in the avian auditory cortex of head-fixed urethane anesthetized zebra finches listening to a large set of their own species natural communication calls. Zebra finches emit various calls in different behavioral contexts and have a complete vocal repertoire composed of 11 call-types. The acoustical characteristics of each call-type have previously been described in detail [39]. Here we focused on the neural representation of 9 call-types: 1) 3 pro-social calls emitted for pair bonding and social cohesion: the Distance Call (DC), the Tet Call (Te), the Nest Call (Ne); 2) the Song (So) that is emitted as a sexual display in males; 3) 2 calls emitted in aggressive encounters: the aggressive Wsst Call (Ws) and the Distress Call (Di); 4) 1 alarm call, the Thuk Call (Th) and 5) 2 calls emitted by juveniles: the Begging Call (Be), used by young to request food, and the Long Tonal Call (LT), a contact call that is a precursor of the adult DC. These 9 call-types show distinctive acoustic features, but they are also idiosyncratically produced by each individual bird such that the caller identity is encoded in the vocalization [41]. Here we investigate how auditory neurons extract the call-type label, irrespective of the individual and rendition variability. In this context, the categorical neural responses are not only defined by their discrimination of call-types but also by their invariance for caller identity or rendition within each particular call-type. To identify categorical neurons, our stimulus set was composed of approximately 10 different exemplars of calls, train of calls or songs produced by different vocalizers for each of the 9 call-types. The call stimuli were randomly sampled from a large annotated data base of calls and songs from the complete repertoire of the male and female zebra finch. Neural responses were recorded using electrode arrays implanted in both hemispheres of 4 male and 2 female

adult zebra finches. We recorded from a total of 914 single auditory units in both primary (Field L) and secondary (CLM, CMM and NCM) avian auditory areas. In previous analysis, we showed using a decoding approach that information about call-types was found in 404 (44%) of these units [40]. Here, we further restricted our population analysis of information to neurons, from that same data set, that reached a significant level of information about the stimulus (see Methods) and analyzed the time-varying information of 337 neurons during the first 600 ms of their response. Note that many calls are shorter than 600 ms, but also that they are often produced in succession with short inter-call intervals. Thus, this analysis window could contain two or more calls or the beginning of a longer song motif composed of multiple syllables. We only analyzed the response in the first 600 ms because the estimation of the cumulative information for longer time windows became unreliable, as we will explain below. Additional details on the neurophysiological recordings can be found in the methods section and in Elie and Theunissen [40].

In the *Results*, we first describe the approach we developed to estimate instantaneous and cumulative time varying information using a Poisson statistic assumption. We illustrate these calculations with specific examples of model neurons. We then validate our assumption and use this methodology to estimate the cumulative information for neurons in the avian auditory cortex as they process natural vocalizations.

We then analyze the time-varying coding properties of the avian cortical auditory neurons with an emphasis on: the relationship between spike rate and instantaneous information, the time constants observed for the cumulative information and the relative fraction of stimulus cumulative information that is used for extracting the behaviorally relevant categories corresponding to distinct call-types.

Estimation of the time-varying information

At a given time t, the *instantaneous* mutual information between the stimulus S and the response Y_t can be written as a difference in Shannon entropies:

$$I_t = H(Y_t) - H(Y_t|S)$$

Here, $H(Y_t)$ is the response entropy for a window at time t, while $H(Y_t|S)$ corresponds to the entropy of the response given the stimulus or the conditional response entropy. $H(Y_t|S)$ can also be called the neural noise since it represents the variability in the neural response to the same stimulus. For spiking neurons, y_t represents the number of spikes in the window at time t (Note: in our notation, capitals are used for random variables and lower case for a sample from that random variable).

Similarly, the cumulative mutual information in neural responses that are discretized into time intervals is given by:

$$CI_t = H(Y_t, Y_{t-1}, Y_{t-2}, \ldots, Y_0) - H(Y_t, Y_{t-1}, Y_{t-2}, \ldots, Y_0|S)$$

The entropies now include the time course of the neural responses starting at $t = 0$ and up to time t. The reference time $t = 0$ is set to the stimulus onset in our analyses but could be any arbitrary reference point.

The conditional response entropy and the response entropy are obtained from the distribution of the conditional probability of neural responses given the stimulus, $p(y_t|s)$, and the distribution of probability of each stimulus $p(s_i)$:

$$H(Y_t|S) = \sum_i p(s_i) \sum_{y_t=0}^{R_{Max}} -p(y_t|s_i) \log_2 p(y_t|s_i)$$

$$H(Y_t) = \sum_{y_t=0}^{R_{Max}} -p(y_t)\log_2 p(y_t)$$

with

$$p(y_t) = \sum\nolimits_i p(s_i)p(y_t|s_i)$$

y_t, the neural response at time t, is measured in spike counts and takes values from zero to a maximum rate value, R_{Max}(for example, as dictated by the neuron's refractory period or numerically as $p \log p$ becomes infinitely small for high spike counts that have very small probability of occurring). The probability of the stimulus $p(s_i)$ is usually taken as $1/n_s$, where n_s is the number of stimuli, unless the study incorporates natural stimulus statistics. The probability of spike counts at time t given a stimulus, $(y_t|s_i)$, could be estimated empirically by recording hundreds or thousands of responses of the same neuron to the same stimulus. Although this approach has been shown to be possible in certain preparations [15], it severely limits the number of stimuli that can be investigated in most neurophysiological experiments. Here we propose a parametric approach where we model the distribution of neural responses to a given stimulus s_i as an inhomogeneous Poisson process. The conditional probability of response (spike count) given the stimulus is then:

$$p(y_t|s_i) = \frac{\mu_{s_i}(t)^{y_t}}{y_t!}e^{-\mu_{s_i}(t)}$$

where $\mu_{s_i}(t)$ is the mean response at time t for stimulus s_i. This mean rate was estimated empirically using the time varying kernel density estimation (KDE) proposed by Shimazaki and Shinomoto [42]. The instantaneous information estimated in this fashion is relatively straightforward, as long as the Poisson assumption is valid and a sufficient number of trials (i.e. stimulus presentations) is obtained to estimate $\mu_{s_i}(t)$ (see Methods). The validity of the Poisson assumption for our dataset is assessed and discussed further below. Note that the Poisson distribution corresponds to the distribution with maximum entropy for unbounded count probabilities given a mean rate [43]. Thus, the value of entropy calculated for the empirical data, $p(y_t|s_i)$, constitutes an upper-bound when using the Poisson distribution assumption. If the empirical data follows a different distribution, the true conditional response entropy, $H(Y_t|S)$, will be lower and, thus, the actual mutual information higher. In other words, the assumption of a Poisson distribution yields a lower bound for the estimation of the mutual information.

The estimation of the cumulative information then extends this approach to joint probabilities of responses in successive time windows, $(y_t, y_{t-1}, y_{t-2}, \ldots)$. Due to what has been labelled as the "curse of dimensionality", the numerical estimation of the unconditional probability becomes exponentially more expensive as the integration time increases. We evaluated multiple approaches based on different assumptions and found that Monte Carlo with importance sampling yielded the best results (see Methods and S3 Fig for comparison to alternative approaches).

If the Poisson assumption also holds for successive time-windows, the conditional probability of response at t is independent of the conditional response at previous times.

$$p(y_t, y_{t-1}, y_{t-2}, \ldots|s_i) = p(y_t|s_i)p(y_{t-1}|s_i)p(y_{t-2}|s_i)\ldots$$

Given this probabilistic independence, it can be shown that the joint conditional response entropy is simply the sum of the conditional response entropies at each time point

(see Methods):

$$H(Y_t, Y_{t-1}, Y_{t-2}, \ldots|S) = H(Y_t|S) + H(Y_{t-1}|S) + H(Y_{t-2}|S) + \cdots$$

Thus, the estimation of the conditional response entropy is straightforward and not affected by the integration time. The problem of dimensionality arises in the estimation of the unconditional probabilities of response and the corresponding response entropy.

The probability of the time varying response is the joint probability of observing (y_t, y_{t-1}, y_{t-2}, ...). This joint probability *cannot* be expressed as the product of the probabilities at different times because these are not independent. More intuitively observing a particular y_{t-1} will affect the probability of observing y_t. This is true because the time varying means of the Poisson distributions $\mu_{s_i}(t)$ are correlated in time; for example, if $\mu_{s_i}(t-1)$ is high we might expect $\mu_{s_i}(t)$ to also have high values. These high values could be true for one particular stimulus s_i but not for the other stimuli. Then observing a high value of y_{t-1} would predict a higher value than expected for y_t (and an increase in probability that it was caused by s_i). The joint unconditional probability distribution is:

$$p(y_t, y_{t-1}, y_{t-2}, \ldots) = \sum_{s_i} p(s_i)[p(y_t|s_i)p(y_{t-1}|s_i)p(y_{t-2}|s_i)\ldots] \neq p(y_t)p(y_{t-1})p(y_{t-2})\ldots$$

Given the lack of independence, the response entropy must then be calculated from the joint probability distribution:

$$H(Y_t, Y_{t-1}, Y_{t-2}, \ldots) = -\sum_{y_t=0}^{R_{Max}} \sum_{y_{t-1}=0}^{R_{Max}} \sum_{y_{t-2}=0}^{R_{Max}} \ldots p(y_t, y_{t-1}, y_{t-2}, \ldots) log_2\, p(y_t, y_{t-1}, y_{t-2}, \ldots)$$

The estimation of this entropy was performed using Monte Carlo with importance sampling. In Monte Carlo, random samples of a vector (y_t, y_{t-1}, y_{t-2}, ...) are drawn from a proposal distribution q(y_t, y_{t-1}, y_{t-2}, ...) and used to estimate the expected value of $\log_2 p(y_t, y_{t-1}, y_{t-2}, \ldots)$ by an algebraic average weighted by the likelihood ratio of $p(y_t, y_{t-1}, y_{t-2}, \ldots)/q(y_t, y_{t-1}, y_{t-2}, \ldots)$. The sampling stops when entropy estimations reach an equilibrium. Information estimations are also known to suffer from positive bias [44]. Here, biased-corrected estimates and errors on information values were obtained from Jackknifing the estimation of the time-varying rates and bootstrapping the Monte Carlo samples.

For these calculations, one has also to determine the size of the time window used for estimating the instantaneous information and correspondingly the steps for the cumulative information. This time window is used to estimate the spike counts and the average rate $\mu_{s_i}(t)$ and depends both on the dynamics of the stimulus and on the response properties of the neurons. By performing a coherence analysis on spike trains and a power spectral analysis on time varying rate in response to natural stimuli, we found that 10 ms (or 50 Hz) captured between 97% and 99% of the dynamics in our system (see Methods and S1 Fig).

Finally, we also estimated the information values for stimulus categories by performing the weighted sum of probabilities for stimuli belonging to each category. The information about categories at time t:

$$I_t = H(Y_t) - H(Y_t|C).$$

is obtained from the conditional probability of response given the category c_k, which is in turn calculated as the average conditional probability of response for the stimuli belonging to that

category:

$$p(y_t|c_k) = \sum_{s_k} p(s_k)p(y_t|s_k)$$

Here, $p(s_k)$ is the probability of occurrence of stimulus s_k within the category c_k. In controlled playback experiments (as here), $p(s_k)$ is $1/n_{s_k}$ where n_{s_k} is the number of stimuli used to sample the category c_k. Similarly, $p(c_k)$, the probability of occurrence of vocalizations in category c_k was taken as $1/n_c$ where n_c is the number of categories. In our system, the stimuli are individual renditions of vocal communication calls that fall into 9 call categories of the zebra finch vocal repertoire. We will contrast stimulus information to categorical information, both instantaneous and cumulative for these behaviorally relevant categories of call-types.

Time-varying information for model neurons

To validate our approach and to illustrate the behavior of time-varying information values, we calculated instantaneous and cumulative information for model neurons with simple and stereotyped response properties. Fig 1 shows the firing rates, raster plots and information values for 3 model neurons in response to 4 stimuli (S1-S4). One model neuron responds to the 4 stimuli with different mean firing rates that are constant in time (Rate Neuron). A second model neuron responds to the four stimuli with the same fixed firing rate but with different latencies (Onset Neuron). The third model neuron also responds with equal average firing rates to the four stimuli but the response occurs at different times (Temporal Neuron).

These simulations illustrate some very basic principles of neural coding. First, many different response profiles can lead to very similar rates of information: in all three simulations, the cumulative information approaches the maximum possible value (2 bits). Second, the coding capacities of neurons are a function of both the range of firing rates that can be achieved (as in the rate neuron) and the modulations in time of this neural activity. Third, the estimation of the instantaneous information gives an incomplete picture of the neural coding of a neuron as it does not incorporate the redundancy or independence of the neural representation over time. For example, on the one hand, when comparing the instantaneous information in the Rate neuron to that of the Onset neuron, one might erroneously conclude that the Rate neuron has more information while, in fact, the cumulative information shows that the Onset neuron is more informative at short time scales. On the other hand, one can also observe that the cumulative information in the Rate neuron continues to increase while the firing rate and the instantaneous information are constant; additional time points allows for a better assessment of that firing rate by time-averaging out neural noise.

These simulations also allowed us to validate our methods. The KDE for the empirical estimation of the time-varying rate based on the generated spike rasters gave very good predictions (solid lines Fig 1A–1C) of the actual model rates (dashed lines Fig 1A–1C): over all stimuli and model neurons, the average error was less than 0.02 spikes/ms. Not surprisingly then, using the actual rate versus the estimated rate yielded practically identical results in the information calculations (dashed vs solid lines, Fig 1D–1F). We also checked that the bias corrected estimates were accurate: the instantaneous information was indeed centered at zero when the response to the 4 stimuli was identical. We verified that the actual values of instantaneous and cumulative information were correct. For example, in the Temporal neuron (Fig 1C & 1F) a peak instantaneous information of 0.5 bits is expected as 1 out 4 stimuli will be almost perfectly discriminated. Finally, we also assessed the limitations of Monte Carlo with importance sampling for the estimation of the cumulative information. For neurons, with continuously high firing rates, this estimation can become unreliable at longer integration times as

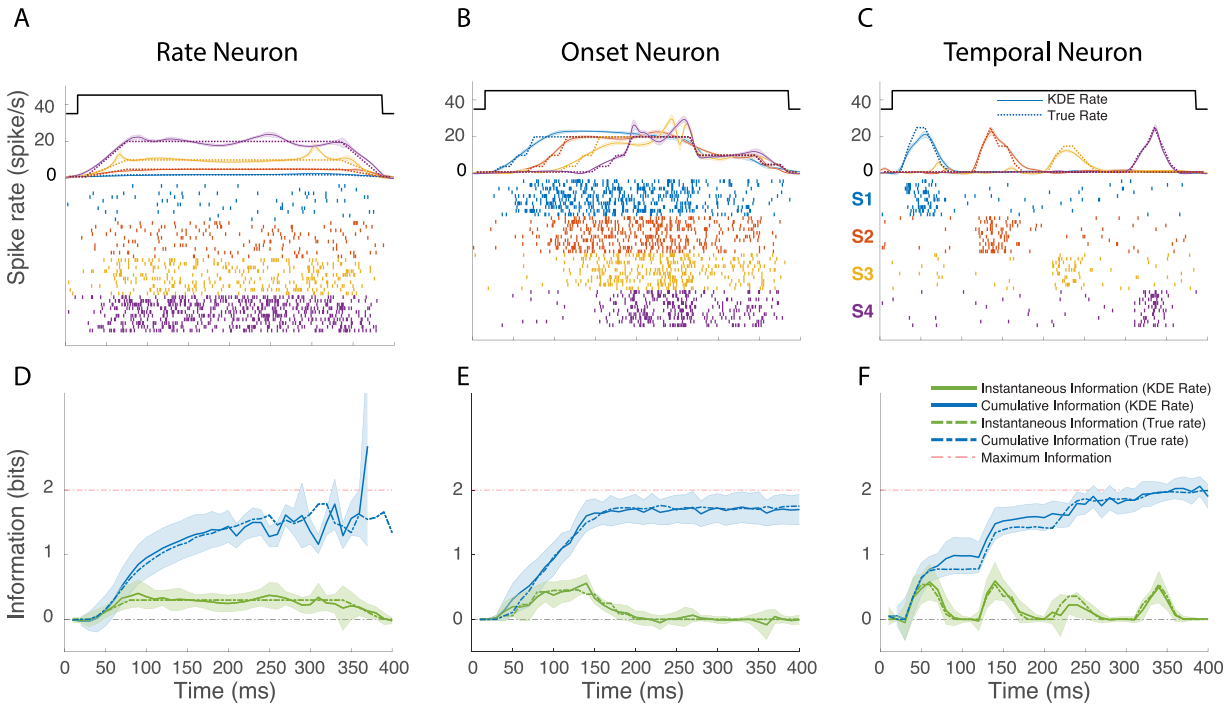

Fig 1. Instantaneous and cumulative information for 3 model poisson neurons. The *top row* (**A,B,C**) shows the firing rate and simulated spike rasters of 3 model neurons to 4 hypothetical stimuli (S1-S4). The onset and offset of all stimuli are identical and shown by the top solid black line. The model rates for the four stimuli are shown in colored dotted lines. These rates are used to generate spike patterns using an inhomogeneous Poisson process. Ten realizations for S1-S4 are shown as spike rasters. The rate recovered from those rasters using a time-varying kernel density estimate (KDE) are shown as colored solid lines. The shaded area around each line indicate the error on the KDE rate estimated by the standard error on the Jackknife estimates. The *bottom row* (**D,E,F**) shows the instantaneous and cumulative information obtained using either the true model rate or the recovered rate from KDE. The error bars in the instantaneous information calculations were obtained by jackknifing the KDE estimates. For the cumulative information calculations, the Monte Carlo simulations were repeated for each jackknife estimate. The error bars therefore include both the errors in the estimation of the true rate from spike rasters (data limitation) and the errors in the estimation of the cumulative information due to limited sampling (computational limitation). The time-varying response rates were chosen to illustrate neural representation for stimuli using different neural codes labelled here as *Rate* (**A, D**), *Onset* (**B, E**), and *Temporal* (**C, F**).

https://doi.org/10.1371/journal.pcbi.1006698.g001

illustrated by the calculation for the Rate neuron (Fig 1A & 1D). However, in those cases, the estimate of the standard error also increased drastically and allowed us to define end points for the calculation of the cumulative information.

Avian auditory neurons have approximate Poisson statistics

Our estimation of the *time*-varying information values is based on a Poisson parametrization for the distribution of spike counts in each time window of analysis and on the independence of the conditional probability of spike counts, $p(y_t|s_i)$, between successive time windows. Here, we assessed the limits of these assumptions.

First, we examined for each neuron, for each stimulus and for each time bin of 10 ms, whether the empirical distribution of spike counts across stimulus presentations was different than the one expected from a Poisson distribution whose mean was given by the kernel density estimation (resampling test; see Methods). For each neuron, we obtained the proportion of time bins where the Poisson assumption was rejected. Fig 2A shows the distribution of this proportion across the neural population. For the large majority of cells (240/404), the empirical spike count could not be distinguished from a Poisson distribution for 90% of the time bins. For a small number of neurons (5/404) non-Poisson spike count statistics could be observed in 40% of the time bins. The mean number of rejected time bins was 10.5%. Thus,

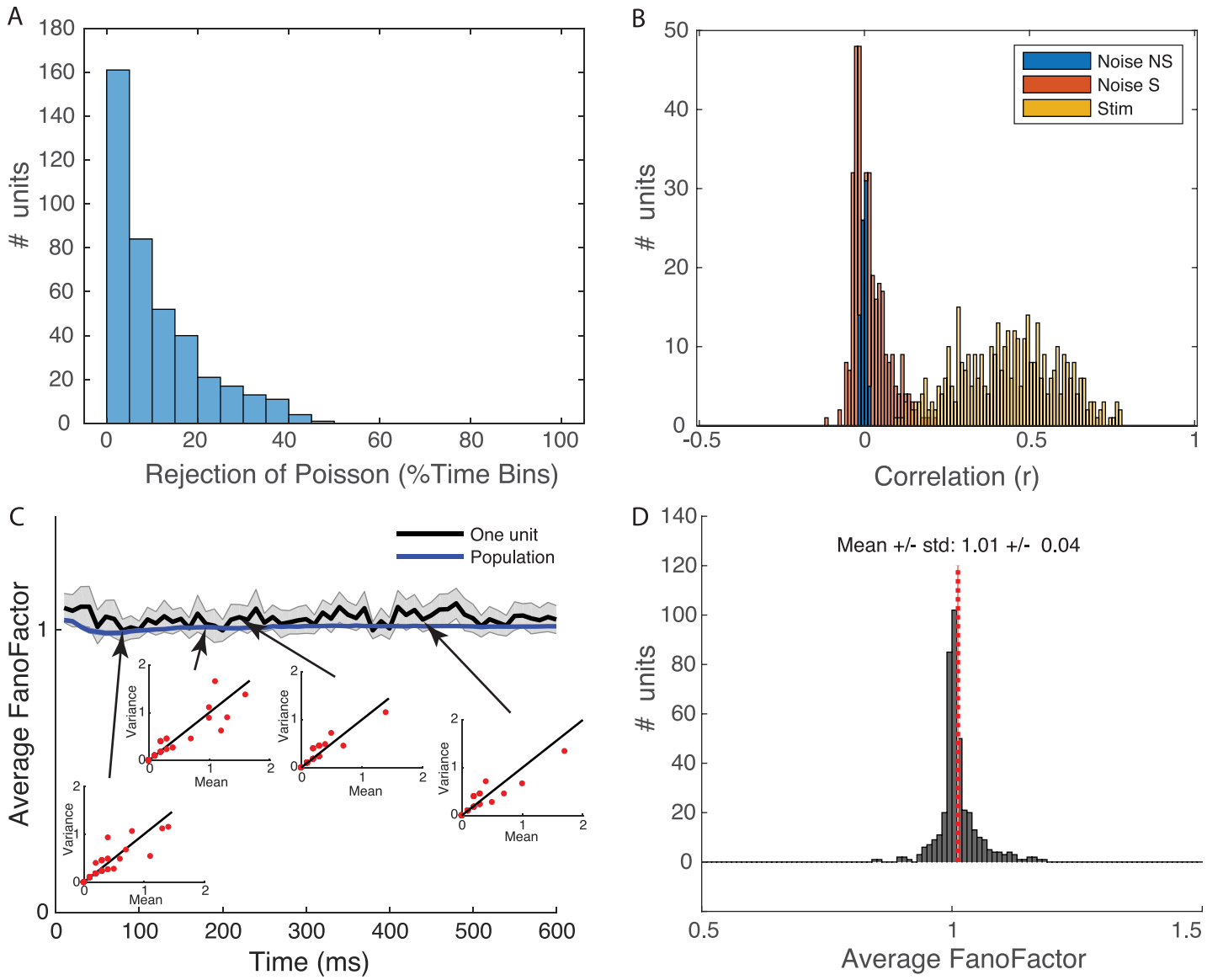

Fig 2. Poisson statistics. A. Distribution in number of neurons (n = 404) of the proportion of time bins where the Poisson assumption can be rejected. **B**. Distribution of the values of stimulus correlation (yellow) and noise correlation in pairs of successive time windows for the neural population (n = 404). Noise correlation values are represented in a stacked histogram depicting values significantly different from zero in red and non-significant values in blue. **C**. Time-varying Fano Factor (Variance/Mean) obtained from the empirical distribution of spike counts estimated in successive 10 ms windows. Spike count distributions are obtained from 10 trials and the average Fano Factor from repeating the estimation of mean and variance calculation for all stimuli presented (min = 54, max = 104). The Fano Factor is shown for one example neuron (solid black line) and also averaged across the entire population (n = 404). Error bars are ± two SEMs. The insets show the mean and variance spike count for the example neuron at a given time shown by the arrows. On those plots, each red dot corresponds to one stimulus (many points overlap). **D**. Distribution of time averaged Fano Factors for the population of neurons (n = 404). Fano Factors are not significantly different from 1.

https://doi.org/10.1371/journal.pcbi.1006698.g002

the Poisson statistics assumption for spike is warranted in more than half of our neurons. For the other neurons, the deviation is relatively small. As stated above, the Poisson assumption used for non-Poisson data yields a lower bound on information. Given the small number of presentations of each stimulus (n = 10), one can argue that the test of conformity to a Poisson distribution is relatively conservative (i.e. we might fail to reject the Poisson statistic because of small sample size). Thus, we also assessed whether the first and second moments (mean and variance) obeyed Poisson statistics. Fig 2C shows the Fano Factor estimated at each

consecutive time bins after stimulus onset, both for individual stimuli (insets) and average across stimuli (solid line). The mean Fano factor for the population was 1.04 with SEM of 0.04 (Fig 2D). This second analysis further quantifies the degree with which our neural data follow Poisson statistics and also shows that for a majority of neurons the assumption is warranted.

The second assumption used in the calculation of cumulative information is that the conditional probability of spike counts, $p(y_t|s_i)$, is independent across successive time bins. To assess the validity of this assumption, we estimated the Pearson correlation coefficient of spike counts obtained in successive time windows in responses to the same stimulus, also known as the noise correlation. Fig 2B shows the distribution of noise correlations for the 404 neurons examined. Significant noise correlations were observed (red bars versus non-significant values shown in blue), but they were small, centered around zero and their distribution showed a slight positive skew. These positive values correspond to cases were a slight increase of the spike count above average at a given stimulus presentation and time bin is somewhat giving a higher chance of greater than average spike count in the successive time bin of the same stimulus presentation. Ignoring noise correlations can lead to either under or over estimates of mutual information depending on the functions that describe the mean time-varying response to each stimulus [45, 46]. In the majority of cases, not taking into account noise correlations leads to an under-estimate of mutual information. For instance, if responses in successive time bins decrease for one stimulus and increase for another stimulus, positive or negative noise correlations will result in lower conditional response entropies, $H(Y_t, Y_{t-1}, Y_{t-2}, \ldots|S)$, and thus higher mutual information. In the general case of stimulus time-varying rates that have different and random shapes depending on the stimulus, one can also show that noise correlations increase mutual information. In some cases, however, noise correlations can increase the conditional response entropies and not taking them into account lead to an over-estimate of the mutual information. Over-estimation of mutual information commonly happens when the responses for stimuli change in the same manner in successive time bins (increase or decrease in mean rates for all sounds), then positive noise correlations can lead to an increase in conditional response entropies and thus a decrease in mutual information. In our data, the time varying rates of the different stimuli are positively correlated at the onset response, but mostly uncorrelated at other times (see Figs 3–5). We might therefore slightly overestimate mutual information at stimulus onset, but for most of our analysis we might be slightly underestimating cumulative information. Given that the noise correlations are small, this estimation bias is small.

While we can neglect noise correlations in our calculations and assume independence of conditional probability of spike counts, $p(y_t|s_i)$, across successive time bins, we are not assuming independence of the unconditional probability, $p(y_t)$, across successive time bins. Indeed, the distribution of stimulus correlation (correlation of mean rate between pairs of consecutive time bins in the trial average response of each stimulus) across the neural population show high positive values centered around 0.4 (Fig 1B). These values indicate that the change in the spike counts in successive time bins is highly driven by the stimulus identity. As such, the unconditional probability distributions in successive time windows are not independent. We do take into account these large stimulus correlations in our estimation of the cumulative mutual information.

Time-varying Information for 3 example neurons

In Figs 3–5, we show the time-varying firing rates, the time-varying instantaneous information and the cumulative information for 3 zebra finch auditory neurons with distinct response properties. The neuron in Fig 3 responded robustly to all communication calls with high and reliable firing rates. It also responded in a time locked fashion sometimes at multiple time

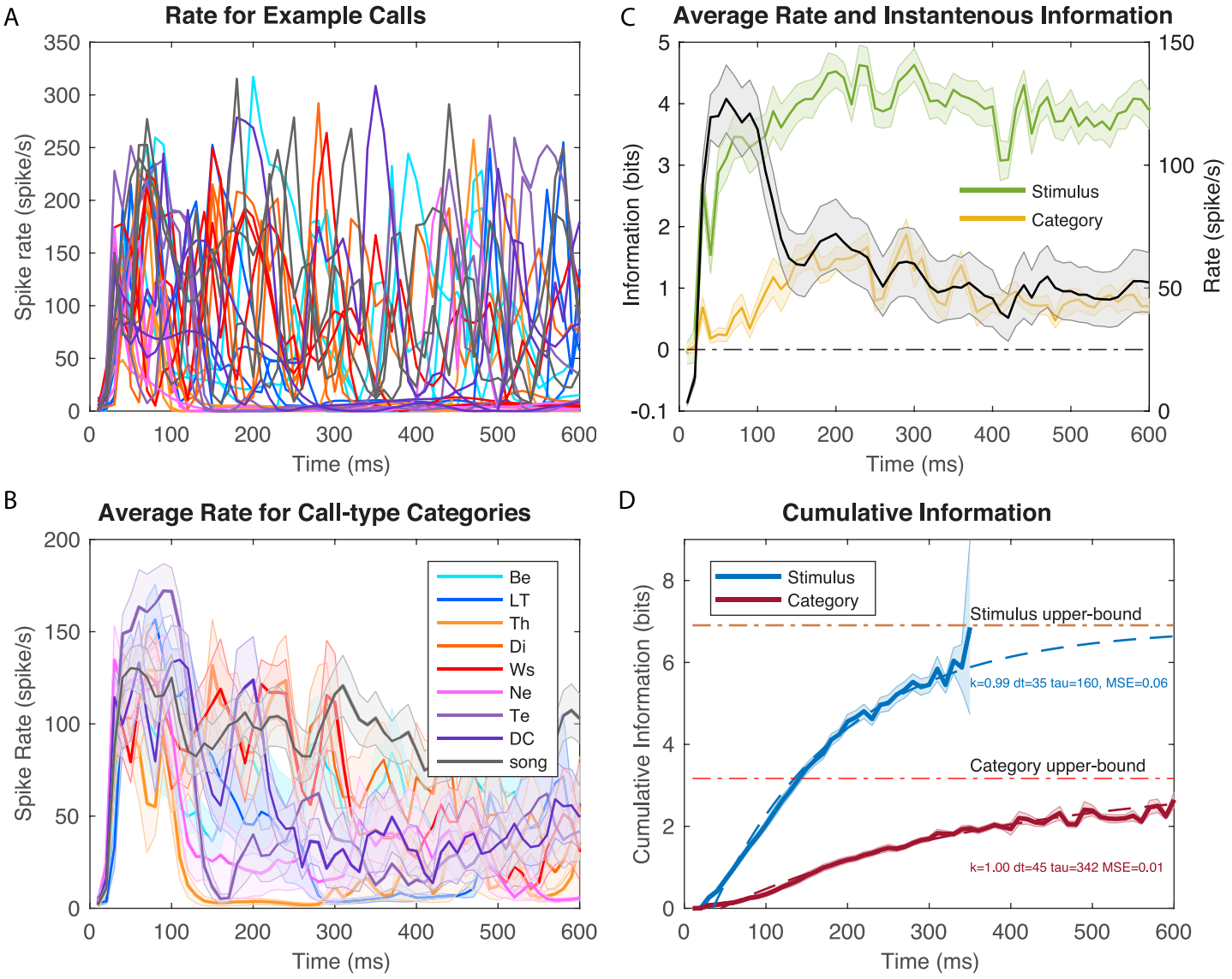

Fig 3. Time-varying firing rates and information values for example neuron 1. A. Time-varying firing rate obtained from kernel density estimation (KDE) of ten repetitions of different example stimuli. The plot shows the responses to two examples of calls from each call-type; the color scheme is as in B; 0 ms corresponds to stimulus onset. **B**. Averaged time-varying firing rate obtained for each call-type category. These average responses are obtained by averaging the KDE estimates of firing rates for each stimulus belonging to the same call-type (~ 10 example stimuli per call-type). Shaded error bars show ± one SEM. **C**. The overall (averaged over all stimuli) time-varying firing rate for the same example neuron is shown by the solid black line (right y-axis) with ± SEM as shaded error bars. The green line and yellow line (left y-axis) correspond to the instantaneous information for stimuli and call-type categories calculated in successive 10 ms windows. Shaded error bars were obtained by a jackknife procedure **D**. Cumulative information for stimuli and call-type categories for the same example neuron. The shaded error bars are ± SE obtained by jackknifing and resampling (see Fig 1 and Methods). Each cumulative information curve is also fitted using an exponential function (dashed lines) characterized by three parameters: a latency (dt, in ms), an exponential rate (tau, in ms) and the infinite time limit value expressed as a fraction of the maximum information achievable (k). The MSE is the mean square error of the fit in bits. This example neuron had a very high stimulus evoked mean firing rate with rapid and reliable time-varying dynamics leading to high information rates. Spontaneous firing was very low (~0 spikes/s). Spike rasters for this neuron are shown on S4 Fig.

https://doi.org/10.1371/journal.pcbi.1006698.g003

points for single calls. Although this neuron was not selective for a particular stimulus or call-type category, by combining rate and temporal codes it reached very high levels of instantaneous information. Moreover, this instantaneous information showed little redundancy yielding very high cumulative information. One can also observe, that for this particular neuron, the average time-varying firing rate is not correlated with instantaneous information. This neuron shows a strong onset response in its firing rate while the instantaneous information is

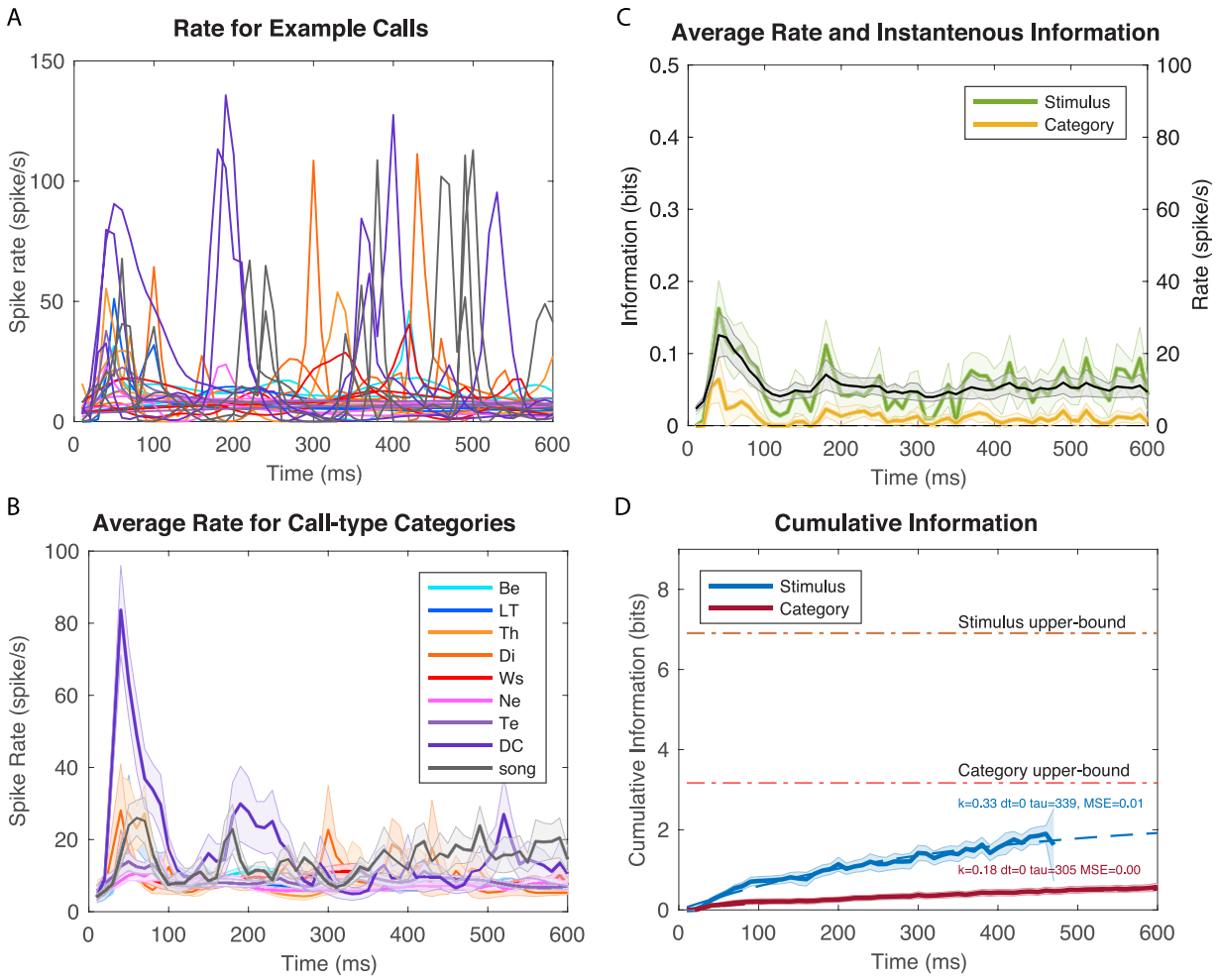

Fig 4. Time-varying firing rates and information values for example neuron 2. As in Fig 3 but for a neuron with lower firing rates and displaying selectivity for Distance Calls. Spike rasters for this neuron are shown on S5 Fig. **A**. Time-varying firing rate obtained from KDE of ten repetitions of different example stimuli. **B**. Averaged time-varying firing rate obtained for each call-type category. **C**. Overall time-varying firing rate and instantaneous information for stimuli and call-type categories. **D**. Cumulative information for stimuli and call-type categories. See details in legend of Fig 3.

https://doi.org/10.1371/journal.pcbi.1006698.g004

almost constant and even slightly lower during the onset response. The neurons in Figs 4 and 5 have much lower firing rates and exhibit selectivity for a call-type, the Distance Call (DC) and the Wsst Call (Ws) respectively. The neuron in Fig 4 exhibits both an onset and sustained response both of which are selective for DC. The neuron in Fig 5 has a much longer latency for response with correlated peaks in firing rate and instantaneous information found between 100 and 300 ms after stimulus onset (Example of spike rasters from single trials for these three neurons are shown in S4, S5 and S6 Figs).

The cumulative information curves were fitted with an exponential function that allows us to quantify the time constant of information accumulation (τ), the saturation level (k) relative to the maximum information that could be achieved (I_{Max}) and the latency (dt):

$$CI_{mod}(t) = kI_{Max}(1 - e^{-(t-dt)/\tau})$$

Most of the calls in the zebra finch repertoire have durations that are shorter than 300 ms [39]. In order to investigate the cumulative information accumulated at this behaviorally

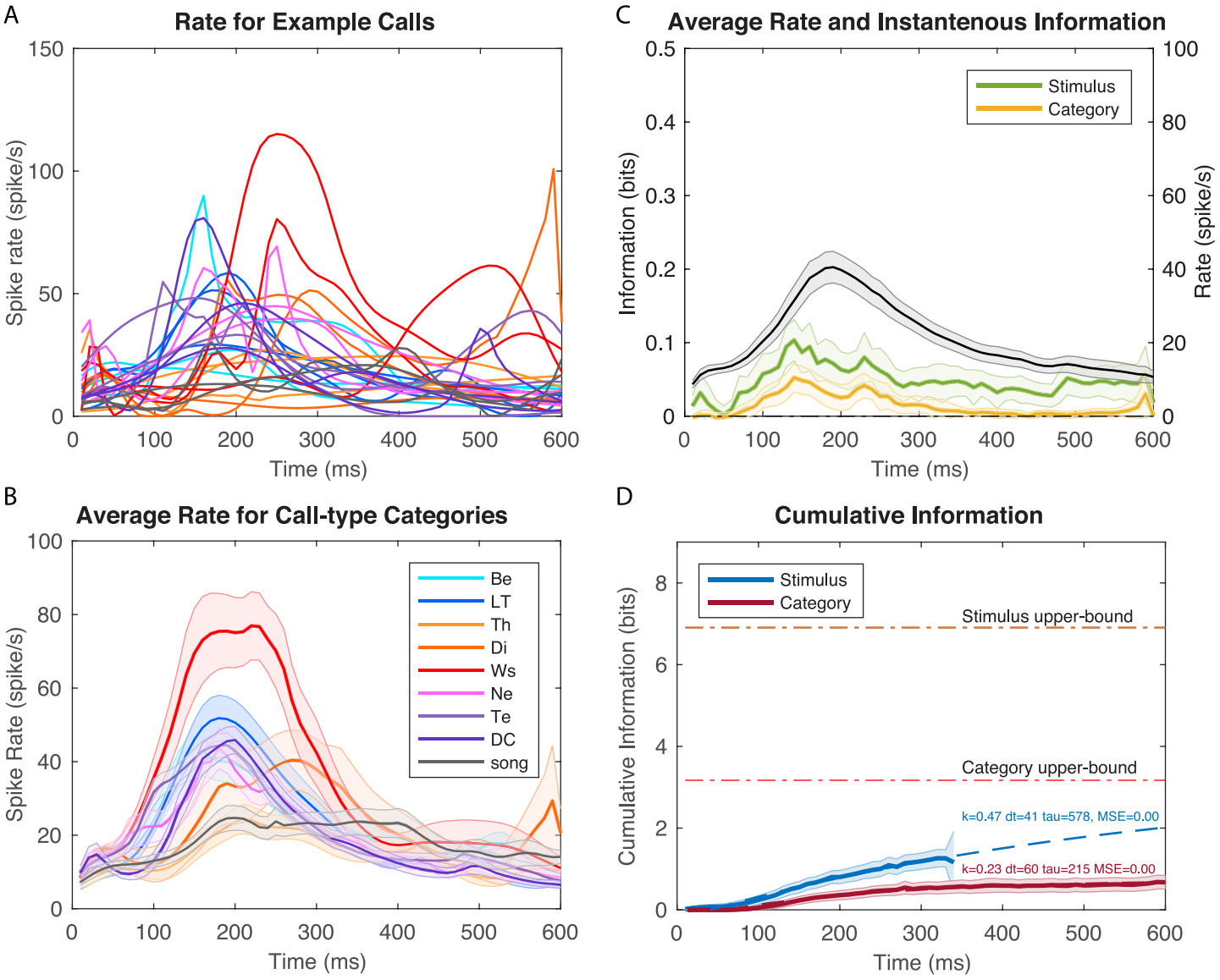

Fig 5. Time-varying firing rates and information values for example neuron 3. As in Fig 3 but for a neuron with intermediate firing rates and displaying selectivity for Wsst (Ws) or Aggressive Calls. Note also the longer latency relative to the neuron shown in 3 and 4. Spike rasters for this neuron are shown on S6 Fig. **A**. Time-varying firing rate obtained from KDE of ten repetitions of different example stimuli. **B**. Averaged time-varying firing rate obtained for each call-type category. **C**. Overall time-varying firing rate and instantaneous information for stimuli and call-type categories. **D**. Cumulative information for stimuli and call-type categories. See details in legend of Fig 3.

https://doi.org/10.1371/journal.pcbi.1006698.g005

relevant fixed point in time, we estimated the relative cumulative information as the value of the model at 300 ms relative to I_{Max}.;

$$k_{300} = CI_{mod}(300)/I_{Max}.$$

The results of the fit are shown in dashed lines in the cumulative information plot for each neuron. One can observe that the neuron in Fig 3 has exceptionally high values of saturation: with sufficiently long integrations time, single spike trials could be used to perfectly assess what stimulus (out of those used in the experiment) was heard and, thus, which call-type category it belonged to. The selective neurons in Figs 4 and 5 have much lower saturation levels as expected since they mostly respond to stimuli from a single call-type category. The neuron in

Fig 5 has longer latency in its categorical cumulative information but this is not the case for the neuron in Fig 4 that has a rapid yet selective onset response. The high firing neuron (Fig 3) also has a faster time constant τ for the stimulus information than the selective neurons shown in Figs 4 and 5. All three neurons have similar and relatively slow time constants for the categorical information in the 200–300 ms range.

Rate and time-varying information: Population analysis

Fig 6A shows the time course of the firing rate averaged across all stimuli and its relationship with the instantaneous information. On average, avian cortical auditory neurons show an onset response followed by a sustained response as observed at many stages of auditory processing and in many sensory systems. The instantaneous information, however, remains almost constant during the entire time. This is true both for the information about individual stimuli or the information about categories. Thus, the onset response is only less informative than the sustained response in bits/spike but not in bits/s; when processing natural vocalizations, the onset and sustained response are equally informative. Moreover, the information in the sustained response continues to provide new information as reflected by the continuous and relatively fast increase in cumulative information shown in Fig 6B. The non-redundancy of the information in the sustained period is also well captured by examining the derivative or slope of the cumulative information shown in Fig 6C. That plot shows how much *new* information is acquired at each time bin. The new information for categories is relatively constant (red line Fig 6C) and similar to the instantaneous information about category (yellow line Fig 6A), while the spike rate is drastically decreasing between 50 and 150ms. Thus, as time progresses, individual spikes carry more new information for categories: there is very little redundancy in the code for call-type categories up to 200 ms and during that time the coding efficiency in bits/spike increases for the new information that is acquired about categories. In contrast, the new information for individual stimuli decreases and is, on average, remarkably correlated with spike rate (Fig 6C). Thus, while, in terms of instantaneous information, the coding efficiency (in bits/spike) is greater in the sustained versus onset period, in terms of non-redundant information, the coding efficiency is remarkably constant throughout the response. This result suggests that an integrating average spike rate measure (i.e. the time running average spike count over stimuli) could serve as a relatively good proxy of cumulative information on individual stimuli. However, note that this cumulative information needs to take into account time varying spike rates (or stimulus locked spike patterns) as in our calculation.

Indeed, taking into account time varying spike rates for each stimulus as opposed to the mean firing rate across the entire time-window for each stimulus is crucial in the accurate calculation of time-varying cumulative information. As seen in Fig 6B, the increases in the cumulative information for the time-varying rate code (solid lines) is much larger than the increases obtained with the fixed rate code (dashed lines; the information still increases because of noise averaging as explained above). The time-varying rates observed in neural responses (as illustrated in Fig 3A) provide additional information. How much more? For the coding of individual stimuli (comparing the dashed blue line to the solid blue line in Fig 6B), a fixed rate code (or assumption) captures only 24% of the information at 100 ms and 21% at 300 ms. The effect for categorical information is smaller because some of the coding dynamics in time-varying responses to stimuli belonging to the same category effectively become neural noise: for categorical information, a fixed rate code captures 50% of the information at 100 ms and 41% at 300 ms.

The distributions of time constants, τ for the cumulative information for stimuli and call-type categories are shown on Fig 7A. The distributions of relative cumulative information at 300ms (k_{300}) are shown on Fig 7B. The range of time constants observed across the population

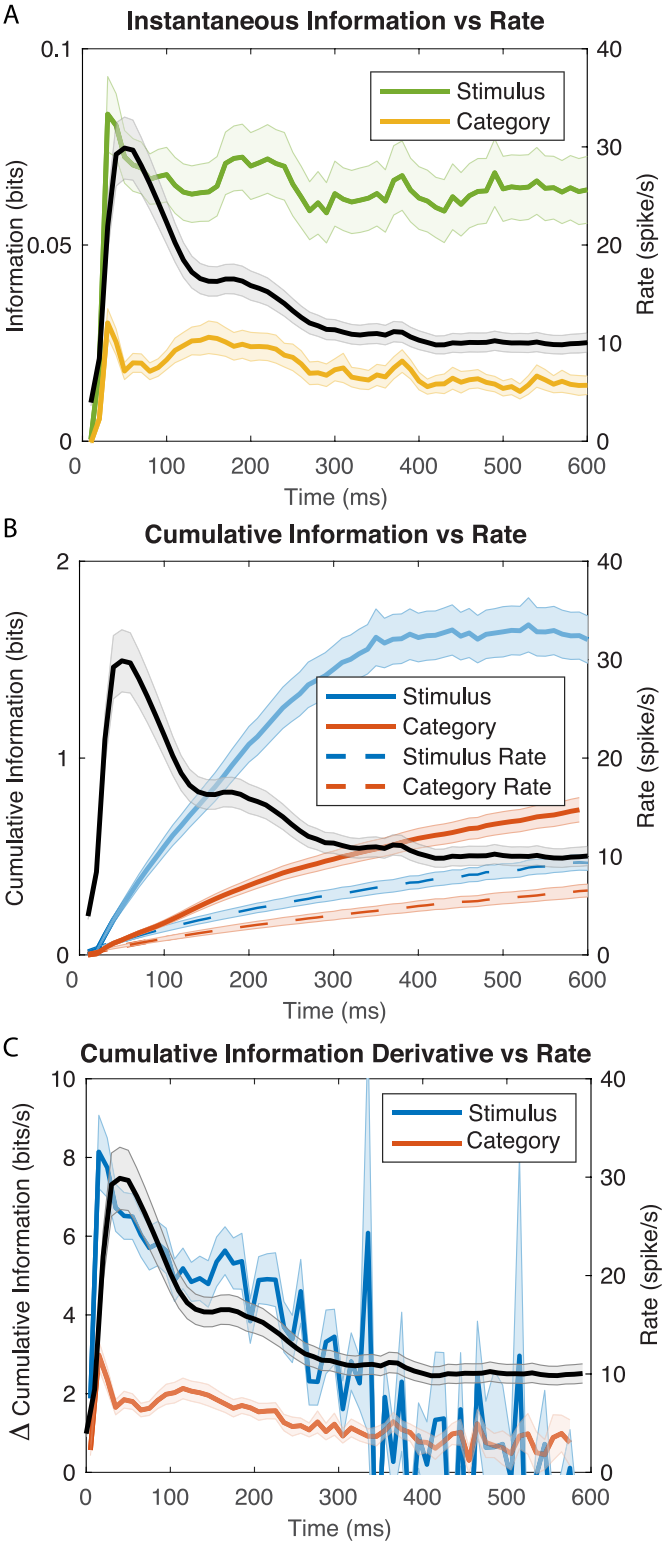

Fig 6. Time-varying firing rates and information values for the population. A. The solid black line (right y-axis) shows the time-varying firing rate averaged across all neurons (n = 337) and all stimuli played. The green and yellow lines (left y-axis) show the instantaneous information for stimuli and call-type categories respectively. **B**. The time-varying mean firing rate is shown as in A. The blue and red solid lines show the population average cumulative information for stimuli and call-type categories respectively. The blue and red dashed lines show the corresponding average cumulative information values calculated with constant averaged firing rates for each stimulus for each neuron over the first 600ms after stimulus onset (a fixed rate code). **C**. The same data as in panel B is plotted but with the derivative (or slope) of the cumulative information values. In all plots, the shaded error bars are ± 2 SEM.

https://doi.org/10.1371/journal.pcbi.1006698.g006

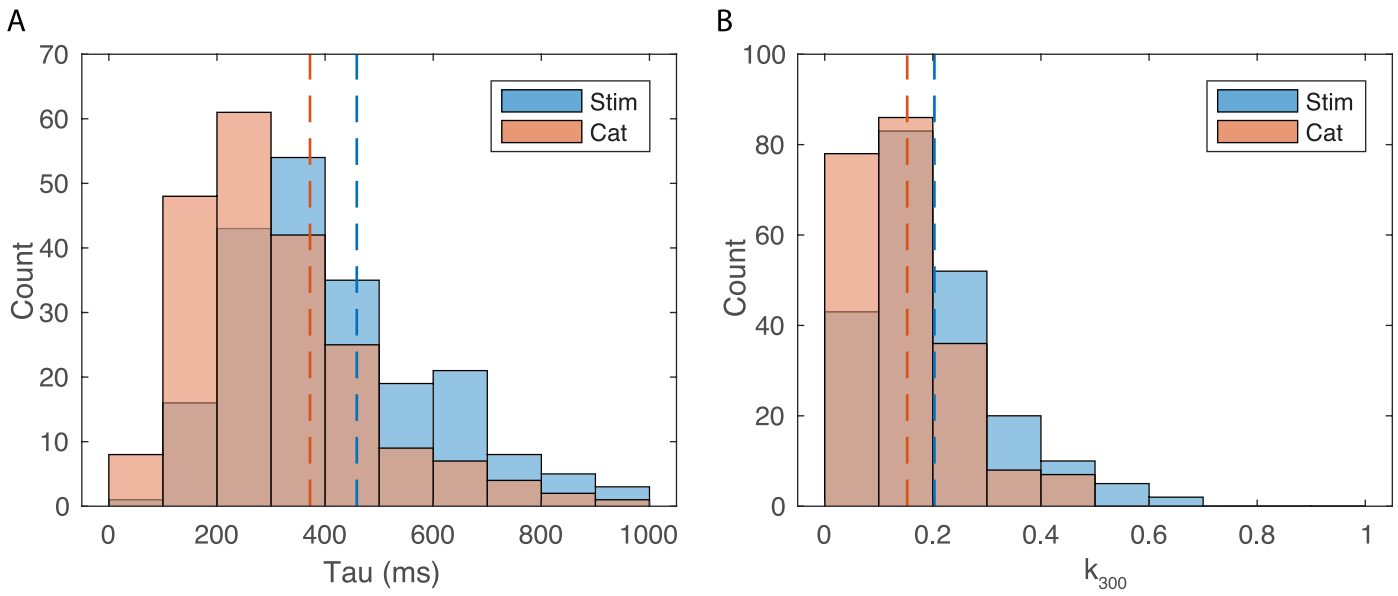

Fig 7. Time constants (tau) for cumulative information and relative cumulative information estimated at 300 ms (K_{300}). For a large fraction of neurons (214/337 neurons), cumulative information for stimuli and call-type categories were well fitted with exponential curves. The time constants of the exponential fits were used as a time integration estimate for cumulative information (A). The exponential fits were also used to estimate the cumulative information obtained at 300 ms relative to the maximum achievable information ($\log_2(n_{stim})$ or $\log_2(n_{cat})$), k_{300}. k_{300} is a number between 0 and 1 (B).**A**. Histogram of the distributions of time constants of the exponential fit for stimuli and call categories. Dashed lines indicate the mean of each distribution (Stim = 459 ms; Cat = 372 ms). **B**. Histogram of the two distributions of k_{300}. Dashed lines indicate the mean of each distribution (Stim = 0.2; Cat = 0.15).

https://doi.org/10.1371/journal.pcbi.1006698.g007

of neurons was large (100 ms-600 ms) with average time constants of 459 ms for cumulative stimulus information and 372 ms for cumulative categorical information. This difference in means of time constants is statistically significant (Paired t-test t(214) = 3.49, p = 0.00058); the ongoing time-varying rate changes continue to provide more information for decoding stimuli and less so for decoding categories of stimuli. There is also a wide distribution of relative cumulative information values (k_{300} ranging from close to zero to 0.6). On average across neurons, the k_{300} was 0.2 (or 20%) for stimuli and 0.15 (or 15%) for categories. These differences in relative information values are highly significant suggesting that, single neurons, capture more variability in stimuli than in categories. Note however that measures of relative information depend on the number of stimuli or categories. Therefore, a direct statistical comparison is not warranted. The comparison between stimulus representation and category representation requires estimations of expected values of categorical information given stimulus information which is performed below. Relative cumulative information values can, however, be used for stimulus and categories independently to assess other coding properties: although average time-varying rates and instantaneous time varying information are not well correlated within single neurons (as shown in Figs 3 and 6), the relative cumulative information is correlated with average firing rates. For cumulative stimulus information, one finds an increase in k_{300} of 1% per spike/s (Adj $R^2 = 0.36$, $F(1,213) = 120$, $p = 1.82\ 10^{-22}$) and, for cumulative categorical information, an increase in k_{300} of 0.8% per spike/s (Adj $R^2 = 0.34$, $F(1,213) = 112$, $p = 2.26\ 10^{-21}$).

Analysis of the categorical information

We are interested in identifying neurons that could play an important role in categorizing vocalizations. We had previously identified example neurons that were highly selective for particular call-types and showed a high degree of invariance such as those shown in Figs 4 and 5

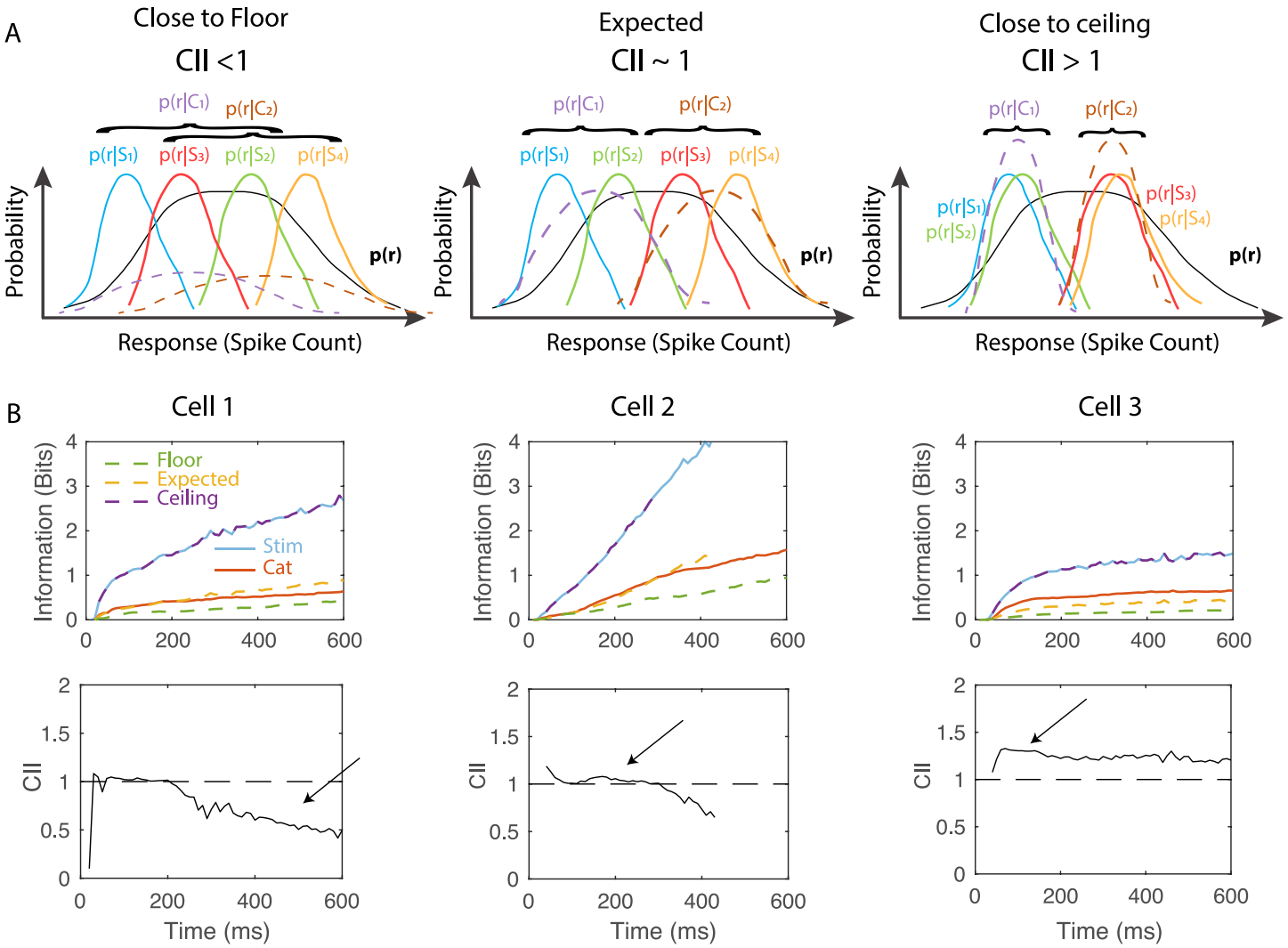

Fig 8. Cumulative information and the categorical information index. A. Schematic of hypothetical response probability distributions that yield different values of the Categorical Information Index (CII). CII is defined by comparing the cumulative categorical information to a floor (CII = 0), an expected value predicted from the stimulus information (CII = 1) and a ceiling value (CII = 2) as described in the text and methods. Each plot represents the distributions of hypothetical neural responses for different stimuli (S1 to S4) that yield different values of CII. When stimuli are grouped randomly into categories (S_1 and S_2 belong to C_1, while S_3 and S_4 belong to C_2), CII is close to 0 (left panel). When stimulus categories preserve the order of the stimulus conditioned response distributions and these distributions are equally separated, effectively coding for both stimuli and categories, CII is equal to 1 (middle panel). When stimuli belonging to the same category have identical responses, the stimulus information is equal to the category information, the response being effectively invariant within categories and CII is equal to 2 (right panel). Note that the actual response probability distributions are multidimensional (one dimension for each time lag) with Poisson marginals. **B**. Example of the cumulative information for stimuli and call-type categories and CII for 3 different representative neurons. The blue and red lines are the cumulative information for stimuli and categories respectively. Estimates of the floor (dashed-green), expected (dashed-orange) and ceiling (dashed-purple, here overlapping with blue) values for the categorical cumulative information as described in A are given as well. The CII is shown as a solid line in the second row for each example neuron. Cell 1 has long periods of time with CII < 1, Cell 2 has long periods of time with CII ~ 1 and Cell 3 has long periods of time with CII > 1 as shown by the arrows.

https://doi.org/10.1371/journal.pcbi.1006698.g008

[40]. Here, we attempted to quantify the neural invariance for call renditions within call-type categories along time. For this purpose, we computed a Categorical Information Index (CII). The CII compares the actual categorical cumulative information to three potential values (see cartoon probability distributions of neural responses in Fig 8A): a floor or minimum value (set at 0), an expected value for shared information between stimuli and categories (set at 1) and a ceiling or maximum value (set at 2). The floor is the categorical information that one would

obtain if stimuli are randomly grouped. The shared-information value is the information that one would obtain if the information about stimuli is equally shared across all stimuli and the neural responses for stimuli are perfectly sorted for each natural call-type category; for example, at a given point in time the 10 renditions of DC would elicit the 10 highest rates, the 10 renditions of LT Call the next 10 higher firing rates and so forth. This shared-information value does therefore assume that, for a given neuronal signal to noise ratio, neural responses segregate categories maximally while also preserving the maximum discrimination between stimuli within categories. The ceiling value is the categorical information value that one would obtain if a maximum amount of information about stimuli was used for the discrimination of categories and a minimum for discriminating stimuli within categories; it assumes maximum invariance to variations within a category. The ceiling value is equal to the stimulus information until it reaches $\log_2(n_c)$, where n_c is the number of categories, corresponding here to the 9 call-types. Fig 8B shows the time-varying floor (dashed-green), shared (dashed-orange) and ceiling (dashed-red) values of cumulative categorical information along with the actual stimulus (solid blue) and categorical (solid red) cumulative information for 3 neurons chosen to illustrate CIIs that are below 1, around 1 and above 1.

Fig 9A shows the distribution of CII values across the neural population (see also S7 Fig for results of this analysis in absolute information units). The thin colored lines correspond to CII curves for single neurons and they are colored according to the time average CII (see legend Fig 9A). The average CII over neurons (bold black line) is very close to the shared value of 1 as one might expect if acoustical differences across stimuli drive neural responses in a linear fashion along some acoustical feature *and* call categories segregate perfectly along that same acoustical feature. However, the average CII is also slightly (and significantly) above 1 between 120 and 260 ms showing some small degree of average invariance for call renditions within a call-type category during those times. In addition it is clear that there is a wide distribution of CII around the shared value of 1. And this distribution is positively skewed towards values higher than 1 as exemplified by the density plots calculated at 150ms. This distribution includes many neurons that exhibit a high degree of invariance for call renditions within call-type categories as also shown by the average CII for the top quartile (red solid line). Focusing on these 25% of neurons with the highest CII, we found a maximum value of CII at 175 ms, indicating that, relative to the maximum invariance achievable given their individual stimulus information values, these neurons are maximally invariant around that time. In terms of absolute value of additional categorial cumulative information relative to the expected shared information value, these same neurons reach a maximum of 0.16 bits at 320 ms (S7 Fig).

Do these high CII neurons exhibit other characteristic response properties? In the scatter plots of Fig 9B and 9C, we examined the relationship between CII (color coded) and, respectively, time constants and relative level of the cumulative information. It can be seen from Fig 9B, that neurons with high CII have relatively long stimulus time constants in comparison to their corresponding time constant observed for categorical information. This relationship is also significant for the entire population of neurons (Linear Regression explaining CII from τ_{cat}/τ_{stim}: Adj $R^2 = 0.11$, $F(1,213) = 28.5$, $p = 2.37\ 10^{-7}$). As shown in Fig 9C, neurons with high CII also have higher relative levels of categorical information in comparison to their relative values for stimulus information although this result is expected given our definition of CII (Linear Regression explaining CII from $k_{300,cat}$—$k_{300,stim}$: Adj $R^2 = 0.76$, $F(1,213) = 675$, $p = 5.5\ 10^{-68}$). Finally, one can also notice that neurons with high CII have low values of relative stimulus cumulative information (Linear Regression explaining CII from $k_{stim}(300)$: Adj $R^2 = 0.28$, $F(1,213) = 84$, $p = 3.96\ 10^{-17}$). This effect is caused by the correlation between

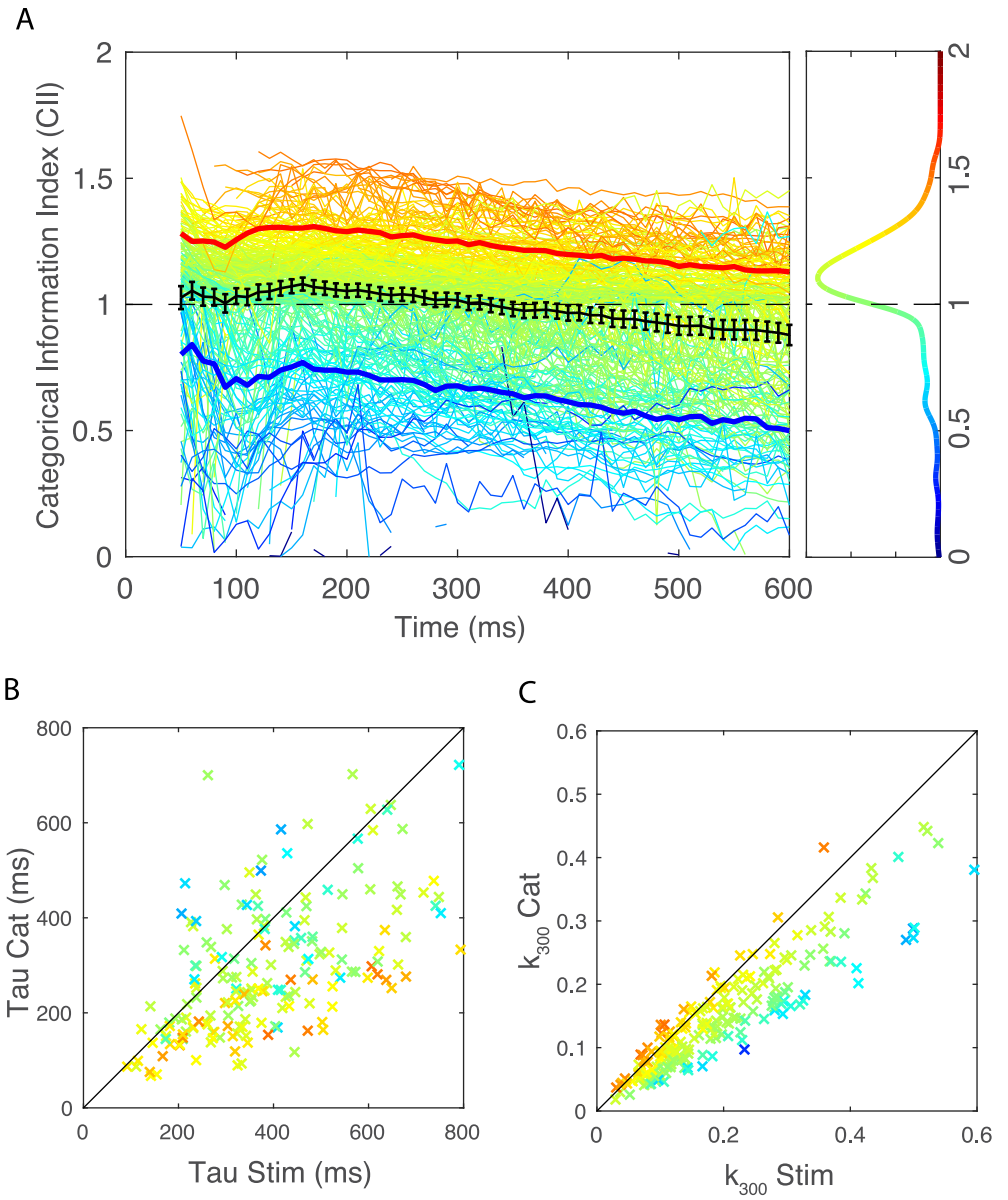

Fig 9. Categorical cumulative information index and its relationship to time constants tau and relative cumulative information estimate K_{300}. As explained in Fig 8 and in the text, the cumulative information (CI) for call-type categories can be analyzed in terms of a floor or minimum value, a shared value and a ceiling or maximum value. **A** Value of the Categorical Information Index (CII) as a function of time for all neurons in the population. Each line color corresponds to the time average CII of the neuron obtained between 50 and 300ms. The solid red line is the average obtained for the quartile of neurons with the highest CII and the blue solid line is average obtained for the quartile of neurons with the lowest CII. Error bars on the average CII (solid bold black line) are 2 SEM. The right vertical panel shows the distribution of CIIs at 150 ms. **B** Scatter plot of the time constant of the exponential fit for stimuli (Stim—x axis) versus call categories (Cat—y axis). **C** Scatter plot of k_{300} for stimuli (Stim—x axis) versus categories (Cat—y axis). Single data points in B and C are colored according to the time average CII of each neuron as in A.

https://doi.org/10.1371/journal.pcbi.1006698.g009

invariance and selectivity as we have shown previously [40]: neurons that show the highest degrees of invariance also tend to respond to a small number of call-type categories and thus a small number of stimuli. As such, high invariance and high selectivity goes hand in hand with lower values of information.

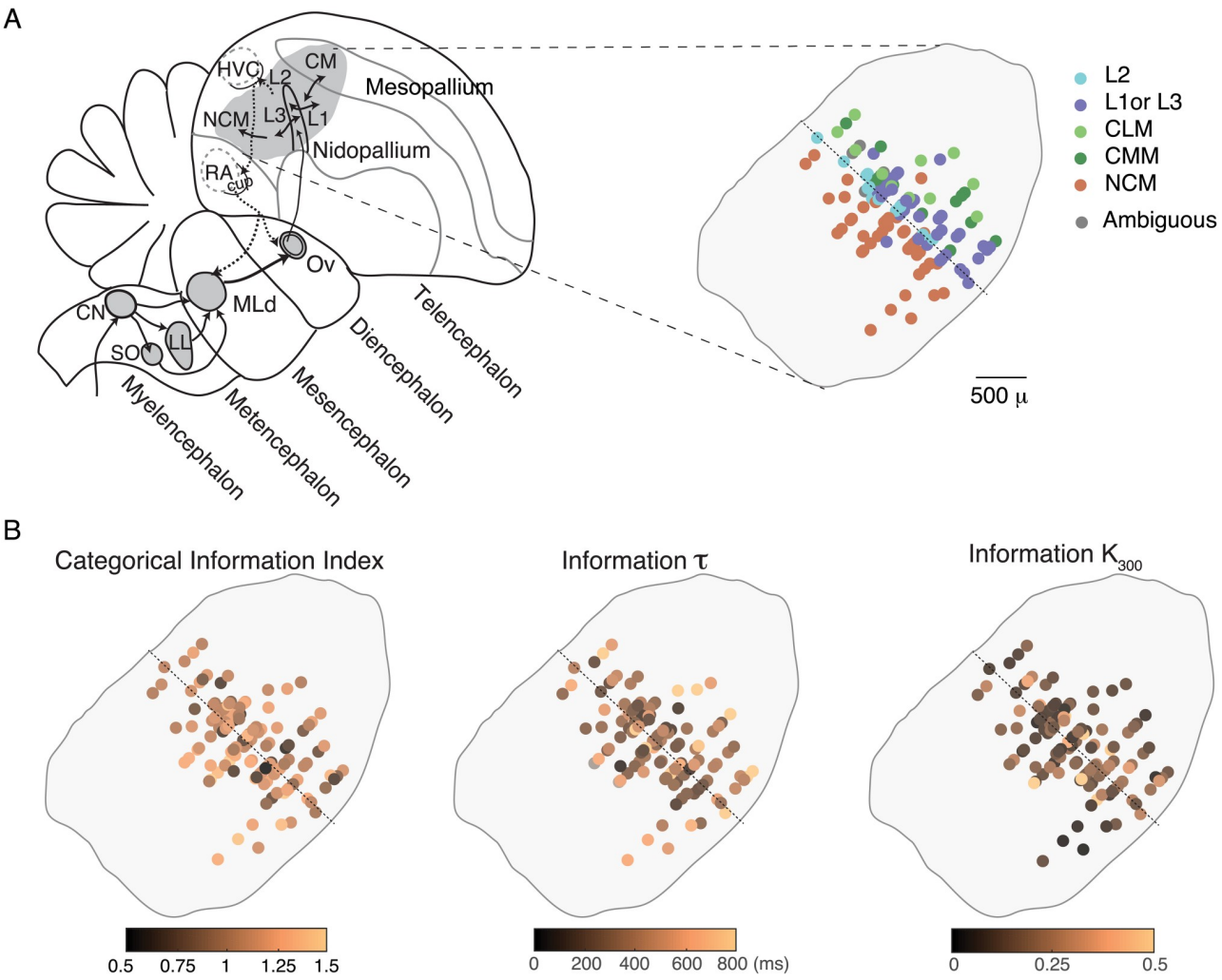

Fig 10. Anatomical organization. A. The left panel is a schematic of the different processing stages in the ascending avian auditory system (grey areas) shown on a parasagittal view (the feedfoward pathways are in solid arrows, the feedback connections are in dotted arrows and motor areas are delineated by grey dotted lines). Our recording sites were from the primary and secondary avian auditory cortical areas (the auditory pallium) and are shown on the right panel. The recording sites from different birds were aligned in the rostral-caudal dimension using the oblique rostro-caudal *long* axis of L2 (dotted line). Recording site positions also varied along the medial-lateral dimension (in and out of the page) and are all superimposed here. Abbreviations: CN, cochlear nucleus; SO, superior olive; LL, lateral lemniscus; MLd, dorsolateral nucleus of the mesencephalon; Ov, nucleus ovoidalis; RA, robust nucleus of the arcopallium; HVC, used as a proper name; NCM, caudomedial nidopallium; CM, caudal mesopallium. **B**. Functional anatomical maps of the Categorical Information Index, the time constant τ and the relative value at 300ms k_{300} of the cumulative information for stimuli for all single units with anatomical data (303/337 neurons). Each site is colored according to the value(s) of the variable being displayed. Transparency is used to reveal overlapping sites in this view as well as, potentially, combination of single units that were recorded from the same electrode and site.

https://doi.org/10.1371/journal.pcbi.1006698.g010

Anatomical organization

We examined whether neurons were organized in the avian auditory cortex based on their CII, cumulative information time constants (τ) and the relative cumulative information at 300 ms relative to the maximum (k_{300}). Most of our recording sites were identified histologically and could be assigned to avian cortical areas that had been segregated into the thalamic recipient area, L2; intermediate primary auditory regions, L1, L3, CLM and L; and secondary areas, CMM and NCM [47–50] (n = 303/337 neurons). We also obtained spatial x,y,z coordinates of the recording sites relative to the midline, the position along the rostral-caudal axis where the lamina pallio-subpallialis (LPS) is the most dorsal, and the top of the brain (Fig 10).

In all regions, we found a range of CII and thus only weak anatomical effects across areas. Using the L2 dorsal-ventral oblique axis as reference point, CII was slightly higher as one moved rostrally or caudally to higher regions of auditory processing (Linear regression: Adj $R^2 = 0.036$, $F(2,195) = 8.39$, $p = 0.0042$). However, an ANOVA also suggested that both regions NCM and L2 had slightly higher mean CII (Adj $R^2 = 0.027$, $F(4,185) = 2.77$, $p = 0.042$). We also observed an increase in the time constant (τ) for the stimulus cumulative information that parallel the increase in CII as one moved away from the L2 axis (Linear regression: Ajd $R^2 = 0.04$, $F(2,195) = 9.32$, $p = 0.00258$). The saturation constant at 300 ms (κ_{300}) for stimulus information was also higher in the central auditory region and smaller as one moved away from the L2 axis (Linear Regression: Ajd $R^2 = 0.035$, $F(2,195) = 8.18$, $p = 0.0047$). Thus, there is more invariance in higher auditory areas and longer time constants but less absolute information about the stimulus. Note again that these significant anatomical trends were characterized by very small effect sizes. We did not find, however, any effects of anatomy on the time constant, (τ), or saturation constants (κ_{300}) for the cumulative categorical information; in this data set, the subset of neurons that have high invariance have similar response properties in both lower and higher regions of the avian auditory pallium.

Discussion

We modeled responses that are observed in the auditory system as inhomogeneous Poisson processes in order to estimate the time-varying instantaneous and cumulative information for vocalizations used in communications. We showed that using Kernel Density Estimation for the time-varying firing rate and Monte Carlo with importance sampling for estimating probabilities, we were able to obtain accurate and bias-free estimates of these time-varying information values. This parametric approach is powerful because a relatively small number of trials can be used to estimate the time-varying response and thus information values can be estimated in response to a relatively large set of stimuli (here over 100 distinct vocalizations) given typical recording times. The auditory avian cortical neurons recorded here and stimulated with natural communication calls had approximate Poisson statistics with small violations that mostly resulted in our estimation of information being a tight lower bound. More generally, the same procedures could also be used with spike trains statistics that can be parametrized with probability functions that depend only on a time varying rate such as inhomogeneous gamma or inhomogeneous inverse Gaussian [12, 51] and could also be extended to ensemble of neurons. Although Poisson statistics are often observed in neural data and will correctly fit any data set obtained from pooling responses across a large number of trials [52], we realize that they make the strong assumption that the firing rate for a particular trial and for a particular neuron depends only on time. Refractory period in spiking neurons and other correlations in single neurons or in an ensemble of neurons that are not phased locked to the stimulus (i.e. noise correlations) are common violations of this assumption. In our data, we found small positive noise correlations reflecting the fact that neural noise effects extend beyond our analysis windows of 10 ms. Taking noise correlations into account can increase stimulus decoding accuracies [23]. In cases where noise correlations are informative, the method proposed here would only yield lower bound estimates of information theoretic values. Alternatively, one should try the use of spike metrics measures that can capture noise correlations [53, 54] in combination with stimulus decoding approaches to then obtain measures of information from the confusion matrix of predicted versus actual stimuli [13, 40, 55, 56]. If spike metrics can be estimated accurately (with limited data), repeating the stimulus decoding procedure for progressively longer time windows would yield cumulative information values such as those calculated here but which could consider noise correlations.

Beyond our methodological contribution, the principal goal of this analysis was to characterize the neural code in higher auditory areas for communication calls. We found that, on average, auditory cortical neurons responded to these natural stimuli with time-varying firing rates that exhibited an onset and a sustained component of the response. Although in some situations the auditory cortex appears to respond only transiently [26, 57], our data supports findings from mammalian species that have showed strong sustained responses when neurons are driven by preferred stimuli [21]. Similarly, in the human superior temporal gyrus, the sustained responses to speech have been shown to be more informative for decoding speech phonemes [58]. Natural sounds in general have also been shown to be particularly efficient at driving auditory areas [12, 59–64] and thus the presence of informative sustained responses is not surprising. Indeed, in some of our most selective neurons, such as the neuron in Fig 5, the onset response is missing and only the sustained response is observed. More generally, and on average, we found that both the transient and sustained response had information about the stimulus identity *and* that the information in these two response phases was not redundant: the stimulus space of natural vocalizations is very large and although the initial response provides some clue as to the nature of the vocalization, new and additional information is observed in the sustained response. On the one hand, one might argue that, in natural vocalizations, the sound itself is changing with time. These stimulus changes occur both within calls that are made of composite notes or in call bouts and song motifs made of multiple syllables. Sustained responses in such cases could be interpreted as multiple onset responses (but with decreasing amplitude). On the other hand, from a behavioral perspective, these communication calls correspond to a single auditory object: a message from a particular individual about a particular state. From either perspective, these observations and analyses illustrate the importance of using behaviorally relevant stimuli when analyzing the nature of the neural code.

The information about stimulus identity was shown to be approximately equal in bits/s in the onset and sustained response. Given that the average onset response (in spikes/s) is greater than the sustained response, one might conclude that the onset coding (in bit/spikes) is not as efficient. Although, this is true for instantaneous information, we also found that the coding efficiency for novel information was remarkably constant for stimulus information while increasing in the first 200 ms for categorical information.

The information efficiency in the onset versus sustained response might also be different when an ensemble of neurons is considered; the relative timing of the first spike has been shown both in audition [26] and in other sensory modalities [25, 27] to be highly informative. Moreover, in addition to stimulus identity other stimulus features are also encoded in neural responses; the timing of the stimulus is clearly marked by the onset response or transient response [22, 58, 65, 66] and other stimulus attributes such as the location, fundamental and loudness/distance of the sound source are also processed in auditory cortex [67–69]. It is therefore very likely that the onset response contains information about stimulus features other than stimulus identity or category and, potentially to a greater extent than in the sustained response (e.g. relative timing of onset). We note, however, that the opposite was found for the neural discrimination of two vowel sounds in the ferret auditory cortex where the onset response or early response was the most informative for decoding vowel identity relative to other attributes [69].

Our IT analysis also revealed the importance of stimulus locked spike patterns even in relatively short neural responses: in just 100 ms, the mean rate captured only a quarter of the information present when time-varying firing rates are considered. Thus, our analyses provide additional evidence that spiking patterns carry a significant amount of stimulus information and should therefore not be ignored in the analysis of neural responses [70, 71]. The spiking precision analyzed here was relatively coarse (10 ms windows) and matched to the time scales of the relevant dynamics in the stimulus: although zebra finch calls are much longer than 10

ms, they are complex sounds with fast spectro-temporal features [39]. Therefore, although the neural code observed here uses fast varying spike rates, it cannot be labelled as a temporal code. In a rigorous definition of a temporal code, temporal information in spike patterns must code stimulus attributes other than the stimulus dynamics [72].

The cumulative information for stimulus identity or for categories increased for sustained periods of time before saturating. These saturating curves were well fitted with exponentials and yielded relatively long-time constants of approximately 460 ms for stimuli identity and 370 ms for call categories. These *information* time constants are long in comparison to the integration times that are usually found for auditory cortical neurons; the spectro-temporal receptive fields of avian auditory neurons rarely extend beyond 50 ms [73–76] although adaptive responses on longer time scales have also been described [77–79]. Information time constants depend on the integration and adaptation time constants of neurons but *also* on the stimulus dynamics: although stimulus identity and stimulus category are fixed in time, the sound itself has time varying features that can provide additional information as time goes on. It is the triple combination of the dynamics of the natural stimulus statistics, the neuronal integration time and the neuronal signal to noise levels that are going to affect the information time constant. Natural sounds and in particular communication calls are informative objects not only because of their spectral structure but also their rich dynamical structure. The importance of time in the neural code used in the auditory system has been emphasized multiple times [80–84] and our cumulative information analysis further stresses the importance of using natural sounds or synthetic sounds that carefully match natural dynamics when probing auditory neural encoding. Ultimately, it is an information time constant that includes stimulus dynamics that should be compared to behavioral responses [18]. A behavioral assay of reaction time for all the calls in the zebra finch repertoire has not been performed but the values of a few hundred of ms correspond to the shortest time intervals between call and call back in anti-phonal calling in paired zebra finches [85]. Although, faster times could be obtained in reaction to any sounds (e.g. startle or orientation), the processing of stimulus identity to extract the information on who is calling and what is being said might require these longer processing times.

Finally, we quantified the fraction of the information about stimulus identity that could be used for extracting the call-type category. We used that analysis to characterize neurons not only in terms of their absolute coding capacities for call-type categories but also in terms of their invariance in their response to different stimuli belonging to the same call-type category. We found that, on average, information for categories is close to what one would expect if neural representation for stimulus identity is also segregated along call-type categories. Nonetheless, the invariance for categories was significantly above that expected value between 120 and 260 ms after stimulus onset. In addition, the distribution of our categorial index also had a positive skew with a long tail: many neurons had a high degree of invariance and could therefore be classified as categorical. We have shown in previous work that this highly invariant neurons could not be simply the result of sampling; when invariance metrics are calculated using stimuli that are grouped randomly but that preserve the acoustical distance found in the natural categories, one does not find the similar highly invariant neurons [40]. To better understand the response properties that give rise to this category specific invariance, one could compare the empirical distribution of Categorical Information Indices (CII) to the distribution of CII obtained either with a bank of linear modulation filters or with models that include particular non-linearities as was done in [86]. The properties of the model neurons that yield the closest CII distribution to the one empirically observed could then be compared to the actual non-linear receptive fields of the neurons that could be estimated by a multiple filter decomposition [87].

Are the categorical neurons in our study anatomically organized? On the one hand, the categorical and invariant neurons described here had even longer time constants for cumulative

stimulus information than the non-invariant neurons in our dataset suggesting that they could be higher in the auditory processing stream. On the other hand, we did not find a separate population of invariant neurons across all of our recordings of neurons informative for categories: the distribution of our categorical index was unimodal. Moreover, we found only weak anatomical correlations with very small increases in the Categorical Information Index along an anatomical axis corresponding to lower vs higher auditory areas. Higher avian auditory cortical areas have been associated with more complex spectro-temporal receptive fields [73, 76, 88], increase robustness to noise [89–91] and more specificity for processing natural sounds [12, 92]. It is possible that we failed to find a large anatomical effect because our sampling of higher auditory areas was relatively sparse; in particular we have a small number of recording sites in CLM and in the more ventral and medial regions of CMM. It is also possible that, although high-invariance neurons are found throughout the avian auditory pallium, this invariance is achieved differently and potentially for different purposes in different regions. A better description of the non-linear receptive fields of these neurons would also allow us to investigate this possibility.

However, it is also striking to see that both in our study and in previous research that investigated invariant properties of avian auditory neurons to classes of natural sounds [92, 93] there is a high degree of heterogeneity in the neural responses within each area. For example, although small differences across auditory areas were reported (and of similar effect size than the spatial effect described here), Meliza and Margoliash [92] also found neurons that were tolerant to different renditions of song types throughout the auditory pallium of the starling. Here, we found that neurons with high Categorical Information Indices could be found anatomically next to neurons with low Categorical Information Indices. Some of this functional heterogeneity could be associated with different cell types [75, 92] and a better understanding of the micro-circuitry in the avian auditory cortical areas is clearly needed [30]. It is interesting to note, however, that this mixing of low level and high-level response properties is not unique to the avian auditory system as similar heterogeneity has been found in the mouse [94] and ferret auditory cortex [95, 96]. Contrary to the visual system, the auditory system might preserve a higher mixture of low-level and high-level sensory responses properties at multiple stages of processing including the higher auditory areas involved in auditory object recognition. If this is true and universally found in vertebrates, it might be a necessary property of the computations needed for auditory object recognition, potentially related to the fact that complexity in auditory signals is in time-varying spectral patterns that quickly disappear; the fleeting nature of sounds could prevent higher processing stages from subsequently accessing lower-level representations for additional information.

Methods

Ethics statement

All animal procedures were approved by the Animal Care and Use Committee of the University of California Berkeley (AUP-2016-09-9157) and were in accordance with the NIH guidelines regarding the care and use of animals for experimental procedures. Birds subjects were euthanized after neural recordings by overdose of Isoflurane.

Animals and stimuli

Four male and two female adult zebra finches (*Taeniopygia guttata*) were used for the electrophysiological experiments. The birds were bred and raised in family cages until they reached adulthood, and then maintained in uni-sex groups. Although birds could only freely interact with their cage-mates, all cages were in the same room allowing for visual and

acoustical interactions between all birds in the colony. All birds were given seeds, water, grid and nest material ad libitum and were supplemented with eggs, lettuce and bath once a week.

Vocalizations used as stimuli during neurophysiological experiments were recorded from 15 adult birds and 8 chicks (20–30 days old). The vocalization bank obtained contains 486 vocalizations that included for each bird most of the calls in the Zebra finch repertoire: 7 call-types in adults and 2 in chicks. The adult calls included the following affiliative calls: Song (So), Distance Call (DC), Tet call (Te) and Nest call (Ne); and the following non-affiliative calls: Wsst or aggressive call (Ws), the Distress call (Di) and one of the two alarm calls, the Thuk (Th). The juvenile calls included the Begging call (Be) and the Long Tonal call (LT). Additional information about these stimuli and their behavioral meanings can be found in [39, 40, 97].

For the neurophysiological experiments, a new subset of the vocalization bank was used at each electrophysiological recording site. This subset was made from a representative number of vocalizations from the repertoire of individuals: three adult females, three adult males, two female chicks and two male chicks. From each individual caller, we randomly chose 3 call bouts from each category or fewer if fewer than 3 call bouts were obtained for that particular call-type and individual. The maximum number of stimuli that could be selected in that procedure was therefore 3x7x3 (adult males x repertoire x renditions) + 3x6x3 (adult females x repertoire x renditions) + 4x2x3 (juveniles x repertoire x renditions) = 141. The average number of stimuli played for each single unit was 114 (sd = 22, min = 34, max = 123). Ten trials were acquired for each stimulus with a few exceptions (min = 9, max = 11). Sounds were broadcasted in a random order using an RX8 processor (TDT System III, sample frequency 24414.0625 Hz) connected to a speaker (PCxt352, Blaupunkt, IL, USA) facing the bird at approximately 40cm. The sound level was calibrated on song stimuli to obtain playbacks at 75dB SPL measured at the bird's location using a sound meter (Digital Sound Level Meter, RadioShack).

Neurophysiological and histological procedures

Extra-cellular electrophysiological recordings were performed in 6 urethane anesthetized adult zebra finches. The birds were placed in a sound-attenuated chamber (Acoustic Systems, MSR West, Louisville, CO, USA) and sound presentation and neural recording were performed using custom code written in TDT software language and TDT hardware (TDT System III). Sounds were broadcasted in a random order as described above. Neural responses were recorded using the signal of two (5 subjects) or one (1 subject) electrode arrays, band-pass filtered between 300Hz and 5kHz and collected by an RZ5-2 processor (TDT System III, sample frequency 24414.0625 Hz). The electrode arrays consisted of two rows of 8 tungsten electrodes with row separation of 500 μm and inter-electrode separation within row of 250 μm (Tucker Davis Technologies, FL). Electrode impedances were approximately 2 MOhms. When two electrode arrays were used, they were placed each in one hemisphere. Spike arrival times and spike shapes of multiple units were obtained by voltage threshold. The level of the threshold was set automatically by the TDT software using the variance of the voltage trace in absence of any stimuli. Electrodes were progressively lowered and neural responses were collected as soon as auditory responses to song, white noise, Distance call or limited modulation noise could be identified on half of the electrodes in each hemisphere (the stimuli used to identify auditory neurons were different from the stimuli used in the analysis). Several recording sites were randomly selected by progressively deepening the penetration of the electrodes and ensuring at least 100 μm between two sites. This distance between recordings insured that we did not record from the same neurons at separate sites.

Table 1. Number of units per bird and per anatomical area that was included in our anatomical analyses (n = 303). The UNC column counts the units that we left as unclassified because they were found at the boundary between regions.

Bird	NCM	CMM	CLM	L2	L1/L3	UNC
BlaBro09xxF	7	0	3	0	10	0
GreBlu9508M	40	12	7	28	34	12
LblBlu2028M	4	2	0	0	2	0
WhiWhi4522M	24	11	7	3	18	0
YelBlu6903F	34	4	3	15	23	0
Sum	*109*	*29*	*20*	*46*	*87*	*12*

https://doi.org/10.1371/journal.pcbi.1006698.t001

On average 4.2±2 sites (mean ± sd) were recorded per bird and per hemisphere at a depth ranging from 400 μm to 2550 μm.

After the last recording site, the subject was euthanized by overdose of isoflurane and transcardially perfused. Coronal slices of 20μm obtained with a cryostat were then alternatively stained with Nissl staining or simply mounted in Fluoroshield medium (F-6057, Fluoroshield with DAPI, Sigma-Aldrich). While Fluoroshield slices were used to localize electrode tracks, Nissl stained slices were used to identify the position of the 6 auditory areas investigated here: the three regions of Field L (L1, L2 and L3), 2 regions of Mesopallium Caudale (CM): Mesopallium Caudomediale (CMM) and Mesopallium Caudolaterale (CLM); and Nidopallium Caudomediale (NCM). By aligning pictures, we were able to anatomically localize most of the recording sites and calculate the approximate coordinates of these sites. Since we could not localize the Y-sinus on slices, we used the position of the Lamina Pallio-Subpallialis (LPS) peak as the reference point for the rostro-caudal axis in all subjects. The surface of the brain and the midline were the reference for respectively the dorsal-ventral axis and the medial-lateral axis. The approximate coordinates of units were used to build 3-D reconstructions of all single unit positions in an hypothetic brain. We were not able to recover all the electrode traces in one of the six birds and excluded that data from the anatomical analysis. Table 1 shows the number of units from each bird for which we had an anatomical location and that had significant stimulus cumulative information (303/337 neurons; see below).

Single unit isolation was performed off-line using custom software that used a combination of supervised and unsupervised clustering algorithms. These clustering algorithms used the spike-snippets shape as described by a PCA. Sorted units where declared to be single units based on spike shape reliability across snippets. The spike shape reliability measure was a signal to noise ratio (SNR) where the signal was the difference between the maximum and the minimum of the average snippet and the noise was the standard deviation of this measure across all snippets. Single units in our data set have an SNR > 5. Additional details on these experimental procedures can be found in [40].

The data of neural responses from 404 out of 914 isolated single units were used in this study. The 404 neurons were selected based on a prior analysis that showed that this subset of units were not only auditory but also contained information about call-types, in the sense that call-types could be decoded above chance level from neural responses (see [40]). Here, we analyzed the neural response in the first 600ms after stimulus onset.

Data analysis: Time-varying spike rate estimation

Calculations of cumulative information requires the estimation of very large distributions which sizes grow exponentially with the number of time points investigated. It is therefore imperative to find the optimal width of the time bin for the analysis: as small as possible to capture information in spike patterns but as large as possible to reduce the dimensionality for the

estimation of mutual information. It is also well known that the relevant time scale can be different in different epochs of the responses (e.g. more precise spike timing at the onset phase of the response and less precise during the sustained phase) [98]. Here, we addressed this issue by first estimating the time-varying rate of our neurons using methods of density estimation based on adaptive kernels. In this methods, narrower kernels are used when spike precision is high and wider kernels at times when the spike precision is low. We then estimated the largest fixed time window that could capture most of the fastest changes in those average time-varying spike rate. Alternative methods based directly on measures of information have also been proposed [98].

For each stimulus, the 9 to 11 raw spike patterns of 600ms, sampled at 10kHz, were combined to obtain the corresponding time varying spike rate (sample frequency set at 1kHz) by applying a locally adaptive kernel bandwidth optimization method [42]. In cases where the neuron did not respond to any of the presentations of the stimulus or responded only once over all presentations, the rate was estimated as being constant for the 600ms duration of the neural response. For those two unresponsive cases, the rate was set to be 1/(2*Ntrials*NTimes) in the absence of any spike or 1/(Ntrials* NTimes) in case of one spike, with Ntrials the number of stimulus presentations (9–11) and NTimes the number of time points at which the rate was estimated (here 600, for a 600ms neural response section with a sampling rate set at 1kHz).

To investigate cumulative information values up to 600ms after stimulus onset, time-varying rates were sampled at 10ms (Nyquist limit frequency of 50Hz). To estimate, the amount of information potentially lost by this low-pass filtering, we estimated an information value based on coherence analysis of the signal to noise ratio in the raw spike train. The coherence between a single spike train (R) and the actual time-varying mean response (A) γ^2_{AR} can be derived from the coherence between the peristimulus time histogram (PSTH) obtained from half of the trials and the PSTH obtained from the other half [99].

$$\gamma^2_{\mathrm{AR}} = \left[1 - \frac{\mathrm{M}}{2} \times \left(1 - \sqrt{\frac{1}{\gamma^2_{\bar{\mathrm{R}}_{1,\mathrm{M}/2}\bar{\mathrm{R}}_{2,\mathrm{M}/2}}}}\right)\right]$$

where M the total number of trials (presentations of the stimuli) and $\gamma^2_{\bar{\mathrm{R}}_{1,\mathrm{M}/2}\bar{\mathrm{R}}_{2,\mathrm{M}/2}}$ the coherence between the two PSTHs calculated on half of the trials. The coherence between two responses is a function of frequency (ω). An estimate of the mutual information (in bits per second) between R and A responses can then be obtained by integrating over all frequencies [4, 99]:

$$I_{AR} = -\int_0^\infty \log_2[1 - \gamma^2_{AR}(\omega)]d\omega$$

For each unit, we estimated the percentage of information preserved as the ratio between I_{AR} calculated up to 50Hz and I_{AR} calculated over all frequencies. Over all units, 96.7% ±6.9% (mean ± SD) of information was conserved by a lowpass filtering at 50Hz and only 88 out of 404 cells had information losses greater than 5%. For each unit, we also calculated the proportion of cumulative power across frequencies in the time varying spike rate estimation obtained with the KDE before low-pass filtering and down-sampling (averaged periodogram in overlapping 200 ms Hanning windows and 1 kHz sampling rate). The cumulative sum of the power was calculated across frequencies and normalized by the maximum power value to obtain the proportion of cumulative power. On average across units, the cumulative power reached 98.8%±1.6% (mean ± SD) at 50Hz, further validating our choice of the temporal resolution (S1 Fig).

Testing the poisson assumptions

To estimate values of mutual information, we assumed that spike counts in our 10 ms windows had a Poisson distribution. To assess the validity of this assumption, we calculated for each neuron, for each stimulus s_i and for each time bin t, the log-likelihood function from the 10 spike counts y_j each obtained in 1 trial j assuming a Poisson distribution with the mean response $\mu_{s_{i,j}}$ given by the delete one estimate.

$$LL_{t,s_i} = \sum_{j=1}^{10} \log\left(\frac{\mu_{s_{i,j}}{}^{y_j}}{y_j!} e^{-\mu_{s_{i,j}}}\right) \quad \text{with} \quad \mu_{s_{i,j}} = \sum_{k=1,\neq j}^{10} \frac{y_k}{10-1}$$

We then generated a bootstrap distribution of 1000 of these likelihood measures each obtained for 10 samples from a Poisson with mean rate given by our KDE estimate. Finally, we calculated the number of times the log-likelihood value obtained for the data was greater or smaller than the bootstrap values to generate a p-value (two-tailed test). Given the very large number of statistical tests, a False Discovery Rate correction, the Benjamini–Hochberg procedure, was used for each neuron to estimate the number of bins for which the Poisson hypothesis could be rejected with p = 0.05.

In our estimations of mutual information, we also assumed that the conditional probability of the response given a stimulus was independent between time bins. In other words, knowing that the spike count is greater (or smaller) than the mean in one time bin provides no information about probability of spike counts in the same trial in another time bin. To test this assumption, we calculated the noise correlation. The noise correlation is simply the Pearson correlation coefficient obtained from successive time bins. A single noise correlation was obtained for each neuron by averaging across pairs of time bins and stimuli.

Information theoretic calculations

As described in the results, the time-varying *instantaneous* mutual information between the stimulus S and the response Y_t can be written as a difference in Shannon entropies:

$$I_t = H(Y_t) - H(Y_t|S)$$

and the *cumulative* mutual information for neural responses that are discretized into time intervals is given by:

$$CI_t = H(Y_t, Y_{t-1}, Y_{t-2}, \ldots, Y_0) - H(Y_t, Y_{t-1}, Y_{t-2}, \ldots, Y_0|S)$$

In the present paper we calculated 4 different types of information: the stimulus instantaneous information, the categorical instantaneous information, the stimulus cumulative information and the categorical cumulative information. Stimulus instantaneous and cumulative information were calculated for all 404 units, while categorical instantaneous and cumulative information were calculated on a restricted set of 337 neurons that presented at least one time point with a significant value of stimulus cumulative information (significant threshold set as 3 times the local error, see below for error calculations). While we verified our assumptions (minimal information loss with spike rate binning, Poisson distributions of spike counts and maximum value of spike counts) on the full set of 404 units, the population analysis of time-varying information presented in the results section only include the relevant dataset of 337 neurons.

The custom Matlab code used to calculate time-varying information values is available at https://github.com/julieelie/PoissonTimeVaryingInfo along with a tutorial on how to use the core functions.

The instantaneous information. The conditional response entropy and the response entropy are obtained from the distribution of the conditional probability of neural responses given the stimulus, $p(y_t|s)$, and the distribution of probability of each stimulus $p(s_i)$:

$$H(Y_t|S) = \sum_i p(s_i) \sum_{y_t=0}^{R_{Max}} -p(y_t|s_i) \log_2 p(y_t|s_i)$$

$$H(Y_t) = \sum_{y_t=0}^{R_{Max}} -p(y_t) \log_2 p(y_t)$$

with

$$p(y_t) = \sum\nolimits_i p(s_i) p(y_t|s_i)$$

We modeled the distribution of neural responses to a given stimulus s_i as an inhomogeneous Poisson process. The conditional probability of response (spike count) given the stimulus is then:

$$p(y_t|s_i) = \frac{\mu_{s_i}(t)^{y_t}}{y_t!} e^{-\mu_{s_i}(t)}$$

And the local entropy is:

$$H(Y_t|s_i) = \mu_{s_i}(t)\left[1/\ln(2) - \log_2(\mu_{s_i}(t))\right] + e^{-\mu_{s_i}(t)} \sum_{y_t=0}^{R_{Max}} \frac{\mu_{s_i}(t)^{y_t} \log_2(y_t!)}{y_t!}$$

Because, in our data, the probability of response is very small for high values of y_t, calculations of entropies were bounded for y_t between zero and R_{Max}. Here we set $R_{Max} = 20$ which corresponds to a rate of 2 spike/ms in the 10 ms analysis windows. The maximum rate observed in all of our neurons across all time bins was 0.8 spike/ms (S2 Fig). Note that this very high firing rate (800 Hz) was only observed once; in other words, it only occurred in one 10 ms time windows across all neurons (404) and all stimuli (114*60 = 6840) or with a p = 1/2763360. This is clearly the very upper limit of a distribution with a long tail. Our time-varying rates were well below that upper bound but we verified that the cumulative probability up to $R_{Max} = 20$ was numerically identical to 1 before estimating entropies. The error due to this sum truncation (from an analytical bound; summing up to 20 instead of up to infinity) can be estimated to a value that is below machine level precision relative to the estimated entropy: $H_{error} < 10^{-16}$ bits). As described above (see time-varying spike rate estimation), we also enforced a lower bound for $\mu_{s_i}(t)$ of 1/20.

The total conditional entropy at time t is:

$$H(Y_t|S) = \sum_{i=1}^{n_s} \sum_{y_t=0}^{R_{Max}} -p(s_i)(p(y_t|s_i) \log_2 p(y_t|s_i))$$

where $p(s_i)$ is the probability of observing stimulus s_i and n_s is the number of stimuli sampled: it is the average of the local entropies obtained for the conditional probability of response for each stimulus s_i. Assuming that our sample is representative of stimuli encountered, each

stimulus is equally probable, $p(s_i) = \frac{1}{n_s}$ or:

$$H(Y_t|S) = \frac{1}{n_s}\sum_{i=1}^{n_s} H(Y_t|s_i)$$

Alternatively, one can assume that each call category is equally probable. If k_i is the number of stimuli in the particular call category c to which s_i belongs and n_c is the number of categories, then:

$$p(s_i) = \left(\frac{1}{k_i}\right)\left(\frac{1}{n_c}\right)$$

for $s_i \in c$. Then:

$$H(Y_t|S) = \frac{1}{n_c}\sum_{i=1}^{n_s}\frac{1}{k_i} H(Y_t|s_i)$$

The unconditional probability (i.e. across all stimuli) of a response at time t is not Poisson but a Poisson mixture characterized by its set of mean rates:

$$p(y_t) = \sum_{i=1}^{n_s} p(s_i)p(y_t|s_i)$$

$$p(y_t) = \frac{1}{n_c}\sum_{i=1}^{n_s}\frac{1}{k_i}\left[\frac{\mu_{s_i}(t)^{y_t}}{y_t!}e^{-\mu_{s_i}(t)}\right]$$

The response entropy is obtained from these unconditional probabilities:

$$H(Y_t) = -\sum_{y_t=0}^{R_{Max}} p(y_t)\log_2(p(y_t))$$

$$H(Y_t) = -\sum_{y_t=0}^{R_{Max}}\left[\frac{1}{n_c}\sum_{i=1}^{n_s}\frac{1}{k_i}\left[\frac{\mu_{s_i}(t)^{y_t}}{y_t!}e^{-\mu_{s_i}(t)}\right]\right]\log_2\left[\frac{1}{n_c}\sum_{i=1}^{n_s}\frac{1}{k_i}\left[\frac{\mu_{s_i}(t)^{y_t}}{y_t!}e^{-\mu_{s_i}(t)}\right]\right]$$

The cumulative information. The conditional probability of a time varying response is the joint probability of observing (y_t, y_{t-1}, y_{t-2},. . .) given s_i. Given our Poisson assumption, the conditional probability of response at t is independent of the conditional response at previous times.

$$p(y_t, y_{t-1}, y_{t-2}, \ldots|s_i) = p(y_t|s_i)p(y_{t-1}|s_i)p(y_{t-2}|s_i)\ldots$$

We can show that the conditional entropy of the joint responses is the sum of the individual

entropies:

$$H(Y_t, Y_{t-1}, Y_{t-2}, \ldots|S) = \frac{1}{n_c}\sum_{i=1}^{n_s}\frac{1}{k_i}\sum_{y_t=0}^{R_{Max}}\sum_{y_{t-1}=0}^{R_{Max}}\sum_{y_{t-2}=0}^{R_{Max}}\ldots(p(y_t|s_i)p(y_{t-1}|s_i)p(y_{t-2}|s_i)\ldots\log_2\{p(y_t|s_i)p(y_{t-1}|s_i)p(y_{t-2}|s_i)\ldots\})$$

$$\begin{aligned}H(Y_t, Y_{t-1}, Y_{t-2}, \ldots|S) &= \frac{1}{n_c}\sum_{i=1}^{n_s}\frac{1}{k_i}\left\{\sum_{y_t=0}^{R_{Max}}(p(y_t|s_i)\log_2 p(y_t|s_i))\right.\\ &+ \sum_{y_{t-1}=0}^{R_{Max}}(p(y_{t-1}|s_i)\log_2 p(y_{t-1}|s_i)) + \sum_{y_{t-2}=0}^{R_{Max}}(p(y_{t-2}|s_i)\log_2 p(y_{t-2}|s_i))\\ &\left.+\ldots\right\}\end{aligned}$$

$$H(Y_t, Y_{t-1}, Y_{t-2}, \ldots|S) = H(Y_t|S) + H(Y_{t-1}|S) + H(Y_{t-2}|S) + \cdots$$

The probability of the time varying response is the joint probability of observing (y_t, y_{t-1}, y_{t-2},. . .). This joint probability cannot be expressed as the product of the probabilities at different times because these are not independent. The joint unconditional probability distribution is:

$$p(y_t, y_{t-1}, y_{t-2}, \ldots) = \frac{1}{n_c}\sum_{i=1}^{n_s}\frac{1}{k_i}[p(y_t|s_i)p(y_{t-1}|s_i)p(y_{t-2}|s_i)\ldots] \neq p(y_t)p(y_{t-1})p(y_{t-2})\ldots$$

This joint probability distribution could be expressed as a product of probabilities by assuming that most of the interdependence can be calculated from the previous time point, as in the Markov chain assumption at a beginning of the time series at $t = 0$. In all cases, the joint probability distribution can be written as:

$$p(y_t, y_{t-1}, y_{t-2}, \ldots) = p(y_t|y_{t-1}, y_{t-2}, \ldots)p(y_{t-1}|y_{t-2}, y_{t-3}, \ldots)\ldots p(y_0)$$

or when it is approximated by a first-order Markov chain:

$$p(y_t, y_{t-1}, y_{t-2}, \ldots) \cong p(y_t|y_{t-1})p(y_{t-1}|y_{t-2})\ldots p(y_0)$$

The conditional probability of y_t given y_{t-1} is:

$$p(y_t|y_{t-1}) = \frac{p(y_t, y_{t-1})}{p(y_{t-1})}$$

$$p(y_t|y_{t-1}) = \frac{\frac{1}{n_c}\sum_{i=1}^{n_s}\frac{1}{k_i}[p(y_t|s_i)p(y_{t-1}|s_i)]}{\frac{1}{n_c}\sum_{i=1}^{n_s}\frac{1}{k_i}[p(y_{t-1}|s_i)]}$$

Note that this joint probability distribution can have very high dimensions. Assuming the number of spikes $y_t \in [0, 19]$, the number of probabilities that must be estimated is 20^{nt} where nt is the number of windows in time. For example, calculating all the probability of all the outcomes for 10 windows (100ms) requires 20^{10}~10^{13} calculations. On the other hand, the Markov chain approximation only requires the estimation of all pair-wise joint probability distributions: for 10 windows and 20 outcomes, the number of calculations is $(10)(20^2)$~4000.

The response entropy is then calculated from the joint probability distribution:

$$H(Y_t, Y_{t-1}, Y_{t-2}, \ldots) = -\sum_{y_t=0}^{R_{Max}} \sum_{y_{t-1}=0}^{R_{Max}} \sum_{y_{t-2}=0}^{R_{Max}} \ldots p(y_t, y_{t-1}, y_{t-2}, \ldots) \log_2 p(y_t, y_{t-1}, y_{t-2}, \ldots)$$

To estimate this response entropy, we investigated various approaches: a time-running cumulative information, the Markov chain approximation and Monte Carlo with importance sampling. Monte Carlo with importance sampling gave the best results and was therefore used in our analyses. The Markov chain approximation overestimates the entropy and therefore can also be used to obtain an upper bound on the cumulative information, which could be useful in other applications. We briefly describe the three approaches. The time-running cumulative information consisted in calculating the full cumulative information (using all possible spike events) but only for a fixed number of successive time windows. We estimated that we could easily calculate all possible probabilities for 4 time-windows, corresponding to a 40 ms history. This approach gave the best estimate of the information in 40 ms windows but, in our system, grossly underestimated the cumulative information: some of the information in successive 40 ms is clearly independent (S3 Fig).

The second approximation was based on Markov chain of variable orders up to 4 (also 40 ms). With this approximation, we overestimated the cumulative information: the correlation time of the time-varying rates for different stimuli is clearly also greater than 40 ms. Using the first order Markov chain approximation:

$$\begin{aligned} &H(Y_t, Y_{t-1}, Y_{t-2}, \ldots) \\ &= -\sum_{y_t}\sum_{y_{t-1}} \ldots \sum_{y_0} \{p(y_t|y_{t-1}) p(y_{t-1}|y_{t-2}) \ldots p(y_0) \log_2 (p(y_t|y_{t-1}) p(y_{t-1}|y_{t-2}) \ldots p(y_0))\} \\ &\quad = -\sum_{y_t}\sum_{y_{t-1}} \ldots \sum_{y_0} \{(p(y_t|y_{t-1}) p(y_{t-1}|y_{t-2}) \ldots p(y_0))(\log_2 p(y_t|y_{t-1}) \\ &\qquad + \log_2 p(y_{t-1}|y_{t-2}) + \ldots + \log_2 p(y_0))\} \end{aligned}$$

Expanding that sum, the last term is:

$$\begin{aligned} L.T. &= -\textstyle\sum_{y_t}\sum_{y_{t-1}} \ldots \sum_{y_0} \{(p(y_t|y_{t-1}) p(y_{t-1}|y_{t-2}) \ldots p(y_0)) \log_2 p(y_0)\} \\ &= -\sum_{y_0}\sum_{y_1} \ldots \left\{\sum_{y_t} p(y_t|y_{t-1})\right\} (p(y_{t-1}|y_{t-2}) \ldots p(y_0)) \log_2 p(y_0) \\ &= -\sum_{y_0}\sum_{y_1} \ldots \left\{\sum_{y_{t-1}} p(y_{t-1}|y_{t-2})\right\} (p(y_{t-2}|y_{t-3}) \ldots p(y_0)) \log_2 p(y_0) \\ &\quad \ldots \\ &= -\sum_{y_0} \left\{\sum_{y_1} p(y_1|y_0)\right\} p(y_0) \log_2 p(y_0) \\ &= -\sum_{y_0} p(y_0) \log_2 p(y_0) \\ &= H(Y_0) \end{aligned}$$

The second to last term is:

$$= -\sum_{y_0}\sum_{y_1}\cdots\left\{\sum_{y_t} p(y_t|y_{t-1})\right\}(p(y_{t-1}|y_{t-2})\ldots p(y_0))\log_2 p(y_1|y_0)$$
$$= -\sum_{y_0} p(y_0)\sum_{y_1} p(y_1|y_0)\log_2 p(y_1|y_0)$$
$$= H(Y_1|Y_0)$$

The third to last term is:

$$= -\sum_{y_0}\sum_{y_1}\cdots\left\{\sum_{y_t} p(y_t|y_{t-1})\right\}\{(p(y_{t-1}|y_{t-2})\ldots p(y_0))\log_2 p(y_2|y_1)\}$$
$$= -\sum_{y_0}\sum_{y_1}\sum_{y_2}\{(p(y_2|y_1)p(y_1|y_0)p(y_0))\log_2 p(y_2|y_1)\}$$
$$= -\sum_{y_1}\sum_{y_2}\left\{\left(p(y_2|y_1)\sum_{y_0}\{p(y_1|y_0)p(y_0)\}\right)\log_2 p(y_2|y_1)\right\}$$
$$= -\sum_{y_1}\sum_{y_2}(p(y_2|y_1)p(y_1))\log_2 p(y_2|y_1)$$
$$= -\sum_{y_1} p(y_1)\sum_{y_2} p(y_2|y_1)\log_2 p(y_2|y_1)$$
$$= H(Y_2|Y_1)$$

and, similarly, for all the other terms.

Thus, the response entropy using the Markov chain approximation is:

$$H(Y_t, Y_{t-1}, Y_{t-2}, \ldots) = H(Y_t|Y_{t-1}) + H(Y_{t-1}|Y_{t-2}) + \cdots + H(Y_0)$$

with

$$H(Y_t|Y_{t-1}) = -\sum_{y_{t-1}}[p(y_{t-1})\sum_{y_t} p(y_t|y_{t-1})\log_2 p(y_t|y_{t-1}))]$$

This approximation can be extended using estimates at two previous time points, *etc.*:

$$H(Y_t, Y_{t-1}, Y_{t-2}, \ldots)$$
$$= H(Y_t|Y_{t-1}, Y_{t-2}) + H(Y_{t-1}|Y_{t-2}, Y_{t-3}) + \ldots + H(Y_2|Y_1, Y_0) + H(Y_1|Y_0) + H(Y_0)$$

where the conditional probability based on two prior measurements is:

$$p(y_t|y_{t-1}, y_{t-2}) = \frac{p(y_t, y_{t-1}, y_{t-2})}{p(y_{t-1}, y_{t-2})}$$

$$p(y_t|y_{t-1}, y_{t-2}) = \frac{\frac{1}{n_c}\sum_{i=1}^{n_s}\frac{1}{k_i}[p(y_t|s_i)p(y_{t-1}|s_i)p(y_{t-2}|s_i)]}{\frac{1}{n_c}\sum_{i=1}^{n_s}\frac{1}{k_i}[p(y_{t-1}|s_i)p(y_{t-2}|s_i)]}$$

where

$$H(Y_t|Y_{t-1}, Y_{t-2}) = -\sum_{y_{t-1}}\sum_{y_{t-2}}[p(y_{t-1}, y_{t-2})\sum_{y_t} p(y_t|y_{t-1}, y_{t-2})\log_2 p(y_t|y_{t-1}, y_{t-2})]$$

Finally, we estimated cumulative information using Monte Carlo with importance sampling. In Monte Carlo in conjunction with importance sampling, N_i samples or, here, time

varying responses $\vec{y}_i$, are taken from a proposal distribution, $q(\vec{y})$. The actual probability, $p(\vec{y}_i)$, is calculated exactly at those samples and an estimate of the expected value of the measure of interest (here $f(\vec{y}) = \log p(\vec{y})$) is obtained by the average of $f(\vec{y})$ at the sample points, $\vec{y}_i$, weighted by the likelihood ratio p/q:

$$E[f(\vec{y})] = \frac{1}{N_i}\sum_i (p(\vec{y}_i)/q(\vec{y}_i))f(\vec{y}_i)$$

Our proposal distribution was based on the distribution at each time point and assuming independence across time:

$$q(y_t, y_{t-1}, y_{t-2}, \ldots) = p(y_t)p(y_{t-1})p(y_{t-2})\ldots$$

For each sample obtained from the proposal distribution q we calculated the actual probability value using:

$$p(y_t, y_{t-1}, y_{t-2}, \ldots) = \frac{1}{n_c}\sum_{i=1}^{n_s}\frac{1}{k_i}[p(y_t|s_i)p(y_{t-1}|s_i)p(y_{t-2}|s_i)\ldots]$$

and used that probability value in the estimation of the entropy. Monte Carlo samples were chunked in groups of 100,000 samples and at each additional sample chunk, a bootstrapped/jackknife bias corrected mean and standard error were estimated (see below). Sampling stopped when the standard error was below 0.2 bits or at a max number of samples set to 5,000,000 samples. If the error at the maximum number of samples was greater than 0.6 bits, the estimation was deemed unreliable and the calculation was not performed for any successive time points.

Bias correction and standard error for information calculations. The small number of trials used to estimate spike rates is the source of bias and uncertainty in our calculation of information: a small number of stimulus presentations increase the probability of obtaining by chance estimated spike rate that are different between stimuli and so yields a positive bias on information calculations. We used a Jackknife procedure on the estimation of spike rate for each stimulus to correct for this positive bias. Jackknife kernel density estimate of the rate were obtained by applying the locally adaptive kernel bandwidth optimization method on the Ntrials possible sets of Ntrials-1 spike patterns of each stimulus (Ntrials being the number of stimulus presentation of a given stimulus). Moreover, uncertainty about information values comes from the sampling errors on spike rate and on the Monte Carlo estimation of joint spike rate probability distributions (cumulative information only). These errors on information calculations were estimated by bootstrapping the jackknife procedure:

$$Error = \sqrt[2]{\frac{\sum_{NB} var(I_{JN})}{NB}}$$

with NB the number of boostrap ($NB = 20$), I_{JN} the bias-corrected estimation of information obtained from the Jackknife procedure.

Calculation of the expected value of categorical information given the stimulus information

We computed a Categorical Information Index (CII) that compared the empirical categorical cumulative information for call-type categories, CCI, to three hypothetical values: a floor (CCI_{Floor}), an expected value (CCI_{Exp}) and a ceiling value (CCI_{Ceil}). The floor value is the categorical cumulative information obtained from random categories. The expected value is the

categorical cumulative information that would be achieved if the stimulus information was 1) equally distributed for each stimulus and 2) could be used for classifying stimuli into groups. Note that the second assumption is not necessarily true in the actual data because the categorical information is based on averaging the probabilities for stimulus from the same category and thus effectively averaging time-varying rates. If time-varying rates are not grouped by categories, then it is possible that two stimuli from two different categories are distinguishable based on their time-varying rate, but that, the average time-varying rates for the two categories are not distinguishable, or less so than expected from the average pair-wise distances. The ceiling value corresponds to the case where all the cumulative information about stimuli is used for the categorization and none to discriminate stimuli belonging to the same category: CCI_{Ceil}. The CII is a number between 0 and 2 that is then calculated as:

$$CII = \frac{CCI - CCI_{Floor}}{CCI_{Exp} - CCI_{Floor}} \; if \; CCI < CCI_{Exp}$$

$$CII = 1 + \frac{CCI - CCI_{Exp}}{CCI_{Ceil} - CCI_{Exp}} \; if \; CCI \geq CCI_{Exp}$$

The following three steps were taken to calculate the expected categorical information (CCI_{Exp}) from the stimulus cumulative information. First, the stimulus mutual information, *mi*, was expressed as the conditional probability of correct stimulus decoding, *p*, for any given stimulus (and assumed to be equal for all stimuli). Given a confusion matrix of size *nxn* obtained from a decoder for *n* stimuli, with *p* as the conditional probability given a stimulus of correct decoding (diagonal terms) and thus *(1-p)/(n-1)* as the conditional probability of error (off-diagonal terms), the mutual information is equal to:

$$mi = p\left[2\log p - \log\frac{p}{n}\right] + (1-p)\left[2\log\frac{1-p}{n-1} - \log\frac{1-p}{(n-1)n}\right]$$

The above equation was inverted numerically to solve for *p* given *mi*. Second, a new matrix was generated by grouping rows and columns of joint probabilities (and not conditional) to form a confusion matrix for categories. The number of stimulus in each category was matched to the actual values in the neurophysiological data on a unit per unit basis. Third, the expected mutual information for categories was then estimated from this new confusion matrix by subtracting the total entropy obtained from the joint probabilities, from the sum of the entropies of the marginal distributions for the rows and columns:

$$mi_{cat} = H_{Row} + H_{Col} - H_{Tot}$$

Supporting information

S1 Fig. Tests for temporal resolution. We performed two tests to assess the potential information loss from sampling the time-varying rate at 50 Hz (10 ms bins). A. The Coherence Test is based on the coherence between individual spike trains. A measure of total coherence (Information Coherence) can be obtained by integrating over frequencies (see Methods). The Information Coherence obtained by integrating from 0 to 50 Hz can then be compared to the Information Coherence obtained for the entire frequency range of 0 to 500 Hz. The histogram shows the number of cells versus the fraction of Information Coherence in 0–50 Hz. B. The Power Spectrum Test is based on the power spectrum of the time-varying rates for each neuron obtained with the Kernel Density estimation (sample frequency: 1kHz). Just as for the

Coherence Information, the fraction of the power between 0–50 Hz relative to the power between 0–500 Hz was estimated for all cells. The histogram shows the number of neurons as a function of that fraction.
(EPS)

S2 Fig. Distribution of spike counts in 10 ms bins. This distribution is shown as the number of time bins across all 404 neurons, all stimuli and all time points that had 0, 1, . . . 8 spikes. Not a single neuron had more than 8 spikes in a 10 ms bins and high spiking events were very rare (only one 10ms bin with 8 spikes over all neurons, all time bins and all stimuli). The average number of spikes per 10 ms bin was 0.108.
(EPS)

S3 Fig. Estimation of the cumulative information. Three methods were tested for the estimation of the Cumulative Information (see Methods): a Markov chain approximation of variable order up to 4 or 40 ms (purple line), the exact information in 4 bins (40 ms) evaluated in running windows (red line), and a Monte Carlo estimation (green line). For comparison, the instantaneous time-varying information in 10 ms windows (blue line) and the integral of that information (yellow line) are also shown. The ceiling value corresponds here to the $\log_2(N_{cat} = 9)$ because this example is showing the categorical cumulative information of the neuron. The Markov chain overestimates the cumulative information while the running window of 40 ms underestimates the cumulative information. The information values plotted here were obtained from the neural data of Example Neuron 1 shown in Fig 3 and also in S4 Fig. This high-firing, high-information neuron allowed us to verify that the calculations were correct around ceiling values.
(EPS)

S4 Fig. Example spike rasters for example neuron 1. The spectrogram of two (randomly chosen) stimuli from each stimulus category are shown with the corresponding spike rasters for 10 trials and a smoothed PSTH for the example neuron shown in Fig 3. This neuron had a very high stimulus driven firing rate and responded to all sound stimuli. The mlnoise stimulus is modulation limited noise: white noise that is low-pass filtered in amplitude and spectral modulations. This stimulus was used here to search for auditory regions but the responses to these synthetic sounds were not included in these analyses.
(EPS)

S5 Fig. Example spike rasters for example neuron 2. As S4 Fig but for the Example Neuron 2. Example Neuron 2, also shown in Fig 4, is selective for Distance Calls (DC).
(EPS)

S6 Fig. Example spike rasters for example neuron 3. As S4 Fig but for the Example Neuron 3. Example Neuron 3, also shown in Fig 5, is selective for aggressive calls or Wsst Calls (Ws).
(EPS)

S7 Fig. Absolute value of invariant categorical information. As explained in Figs 8 and 9 and in the text, the cumulative information (CI) for call-type categories can be analyzed in terms of a floor or minimum value, a shared value and a ceiling or maximum value. This graph represents the difference in the observed cumulative information in bits relative to the shared value. The bold black line is the average across all neurons (n = 337; error bars show 2 SEMs). The thin colored lines correspond to individual neurons. Each line color corresponds to the time average Categorical Information Index (CII) of the neuron obtained between 50 and 300ms (Fig 9A). The same color code is used in all panels of Fig 9. The solid red line in the main plot is the average obtained for the quartile of neurons with the highest CII and the blue

solid line is average obtained for the quartile of neurons with the lowest CII. The vertical plot on the right represents the density of neurons for each value of difference in information at 150 ms.
(EPS)

Acknowledgments

This research used the Savio computational cluster resource provided by the Berkeley Research Computing program at the University of California, Berkeley (supported by the UC Berkeley Chancellor, Vice Chancellor for Research, and Chief Information Officer).

Author Contributions

Conceptualization: Julie E. Elie, Frédéric E. Theunissen.

Data curation: Julie E. Elie.

Formal analysis: Julie E. Elie, Frédéric E. Theunissen.

Funding acquisition: Frédéric E. Theunissen.

Investigation: Julie E. Elie.

Methodology: Julie E. Elie, Frédéric E. Theunissen.

Project administration: Frédéric E. Theunissen.

Resources: Frédéric E. Theunissen.

Software: Julie E. Elie, Frédéric E. Theunissen.

Supervision: Frédéric E. Theunissen.

Validation: Julie E. Elie.

Visualization: Julie E. Elie, Frédéric E. Theunissen.

Writing – original draft: Frédéric E. Theunissen.

Writing – review & editing: Julie E. Elie.

References

1. Shannon CE, Weaver W. The matematical theory of communication. Chicago: University of Illinois Press; 1963.
2. Wu MCK, David SV, Gallant JL. Complete functional characterization of sensory neurons by system identification. Annual Review of Neuroscience. 29 2006. p. 477–505. https://doi.org/10.1146/annurev.neuro.29.051605.113024 PMID: 16776594
3. Rieke F, Warland D, de Ruyter van Steveninck R, Bialek W. Spikes: Exploring the Neural Code. Cambridge, MA: MIT Press; 1997.
4. Borst A, Theunissen FE. Information theory and neural coding. Nat Neurosci. 1999; 2(11):947–57. https://doi.org/10.1038/14731 PMID: 10526332.
5. Rolls ET, Treves A. The neuronal encoding of information in the brain. Progress in Neurobiology. 2011; 95(3):448–90. https://doi.org/10.1016/j.pneurobio.2011.08.002 PMID: 21907758
6. Stone JV. Principles of Neural Information Theory: Sebtel Press; 2018. 197 p.
7. Optican LM, Richmond BJ. Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretic analysis. Journal of Neurophysiology. 1987; 57:162–78. https://doi.org/10.1152/jn.1987.57.1.162 PMID: 3559670
8. Reinagel P, Reid RC. Temporal coding of visual information in the thalamus. J Neurosci. 2000; 20 (14):5392–400. PMID: 10884324.

9. Brenner N, Strong SP, Koberle R, Bialek W, de Ruyter van Steveninck RR. Synergy in a neural code. Neural Comput. 2000; 12(7):1531–52. PMID: 10935917.
10. Chechik G, Anderson M, Baryosef O, Young E, Tishby N, Nelken I. Reduction of Information Redundancy in the Ascending Auditory Pathway. Neuron. 2006; 51(3):359–68. https://doi.org/10.1016/j.neuron.2006.06.030 PMID: 16880130
11. Rieke F, Bodnar DA, Bialek W. Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. Proc R Soc Lond B Biol Sci. 1995; 262(1365):259–65.
12. Hsu A, Woolley SM, Fremouw TE, Theunissen FE. Modulation power and phase spectrum of natural sounds enhance neural encoding performed by single auditory neurons. J Neurosci. 2004; 24 (41):9201–11. Epub 2004/10/16. https://doi.org/10.1523/JNEUROSCI.2449-04.2004 PMID: 15483139.
13. Amin N, Gastpar M, Theunissen FE. Selective and efficient neural coding of communication signals depends on early acoustic and social environment. PLoS One. 2013; 8(4):e61417. Epub 2013/05/01. https://doi.org/10.1371/journal.pone.0061417 PMID: 23630587.
14. Atick J. Could information theory provide an ecological theory of sensory processing? Network. 1992; 3:213–51.
15. Strong SP, Koberle R, de Ruyter van Steveninck R, Bialek W. Entropy and information in neural spike trains. Phys Rev Letters. 1998; 80(1):197–200.
16. Kanwisher N, Yovel G. The fusiform face area: a cortical region specialized for the perception of faces. Philosophical Transactions of the Royal Society B-Biological Sciences. 2006; 361(1476):2109–28. https://doi.org/10.1098/rstb.2006.1934 PMID: 17118927
17. DeWitt I, Rauschecker JP. Phoneme and word recognition in the auditory ventral stream. Proceedings of the National Academy of Sciences of the United States of America. 2012; 109(8):E505–E14. https://doi.org/10.1073/pnas.1113427109 PMID: 22308358
18. Kuboki R, Sugase-Miyamoto Y, Matsumoto N, Richmond BJ, Shidara M. Information Accumulation over Time in Monkey Inferior Temporal Cortex Neurons Explains Pattern Recognition Reaction Time under Visual Noise. Frontiers in Integrative Neuroscience. 2017;10.
19. Bair W, Cavanaugh JR, Smith MA, Movshon JA. The timing of response onset and offset in macaque visual neurons. Journal of Neuroscience. 2002; 22(8):3189–205. PMID: 11943820
20. Bisley JW, Krishna BS, Goldberg ME. A rapid and precise on-response in posterior parietal cortex. Journal of Neuroscience. 2004; 24(8):1833–8. https://doi.org/10.1523/JNEUROSCI.5007-03.2004 PMID: 14985423
21. Wang X, Lu T, Snider RK, Liang L. Sustained firing in auditory cortex evoked by preferred stimuli. Nature. 2005; 435(7040):341–6. https://doi.org/10.1038/nature03565 PMID: 15902257
22. Zheng Y, Escabi MA. Distinct Roles for Onset and Sustained Activity in the Neuronal Code for Temporal Periodicity and Acoustic Envelope Shape. Journal of Neuroscience. 2008; 28(52):14230–44. https://doi.org/10.1523/JNEUROSCI.2882-08.2008 PMID: 19109505
23. Victor JD, Purpura KP. Nature and precision of temporal coding in visual cortex: A metric-space analysis. Journal of Neurophysiology. 1996; 76(2):1310–26. https://doi.org/10.1152/jn.1996.76.2.1310 PMID: 8871238
24. Mechler F, Victor JD, Purpura KP, Shapley R. Robust temporal coding of contrast by V1 neurons for transient but not for steady-state stimuli. Journal of Neuroscience. 1998; 18(16):6583–98. PMID: 9698345
25. Delorme A, Thorpe SJ. Face identification using one spike per neuron: resistance to image degradations. Neural Networks. 2001; 14(6–7):795–803. https://doi.org/10.1016/s0893-6080(01)00049-1 PMID: 11665771
26. Heil P. First-spike latency of auditory neurons revisited. Current Opinion in Neurobiology. 2004; 14 (4):461–7. https://doi.org/10.1016/j.conb.2004.07.002 PMID: 15321067
27. Johansson RS, Birznieks I. First spikes in ensembles of human tactile afferents code complex spatial fingertip events. Nature Neuroscience. 2004; 7(2):170–7. https://doi.org/10.1038/nn1177 PMID: 14730306
28. Rolls ET, Franco L, Aggelopoulos NC, Jerez JM. Information in the first spike, the order of spikes, and the number of spikes provided by neurons in the inferior temporal visual cortex. Vision Research. 2006; 46(25):4193–205. https://doi.org/10.1016/j.visres.2006.07.026 PMID: 17011607
29. de Heer WA, Huth AG, Griffiths TL, Gallant JL, Theunissen FE. The Hierarchical Cortical Organization of Human Speech Processing. Journal of Neuroscience. 2017; 37(27):6539–57. https://doi.org/10.1523/JNEUROSCI.3267-16.2017 PMID: 28588065
30. Tsunada J, Cohen YE. Neural mechanisms of auditory categorization: from across brain areas to within local microcircuits. Front Neurosci. 2014; 8:161. https://doi.org/10.3389/fnins.2014.00161 PMID: 24987324.

31. Sharpee TO, Atencio CA, Schreiner CE. Hierarchical representations in the auditory cortex. Current Opinion in Neurobiology. 2011; 21(5):761–7. https://doi.org/10.1016/j.conb.2011.05.027 PMID: 21704508

32. Gifford GW 3rd, MacLean KA, Hauser MD, Cohen YE. The neurophysiology of functionally meaningful categories: macaque ventrolateral prefrontal cortex plays a critical role in spontaneous categorization of species-specific vocalizations. J Cogn Neurosci. 2005; 17(9):1471–82. https://doi.org/10.1162/0898929054985464 PMID: 16197700

33. Grimsley JM, Shanbhag SJ, Palmer AR, Wallace MN. Processing of communication calls in Guinea pig auditory cortex. PLoS One. 2012; 7(12):e51646. Epub 2012/12/20. https://doi.org/10.1371/journal.pone.0051646 PMID: 23251604.

34. Plakke B, Diltz MD, Romanski LM. Coding of vocalizations by single neurons in ventrolateral prefrontal cortex. Hear Res. 2013; 305:135–43. https://doi.org/10.1016/j.heares.2013.07.011 PMID: 23895874.

35. Miller CT, Thomas AW, Nummela SU, de la Mothe LA. Responses of primate frontal cortex neurons during natural vocal communication. Journal of Neurophysiology. 2015; 114(2):1158–71. https://doi.org/10.1152/jn.01003.2014 PMID: 26084912

36. Carruthers IM, Natan RG, Geffen MN. Encoding of ultrasonic vocalizations in the auditory cortex. J Neurophysiol. 2013; 109(7):1912–27. Epub 2013/01/18. https://doi.org/10.1152/jn.00483.2012 PMID: 23324323.

37. Comins JA, Gentner TQ. Perceptual categories enable pattern generalization in songbirds. Cognition. 2013; 128(2):113–8. https://doi.org/10.1016/j.cognition.2013.03.014 PMID: 23669049.

38. Ter Maat A, Trost L, Sagunsky H, Seltmann S, Gahr M. Zebra finch mates use their forebrain song system in unlearned call communication. PLoS One. 2014; 9(10):e109334. https://doi.org/10.1371/journal.pone.0109334 PMID: 25313846.

39. Elie JE, Theunissen FE. The vocal repertoire of the domesticated zebra finch: a data-driven approach to decipher the information-bearing acoustic features of communication signals. Animal cognition. 2016; 19(2):285–315. https://doi.org/10.1007/s10071-015-0933-6 PMID: 26581377

40. Elie JE, Theunissen FE. Meaning in the avian auditory cortex: neural representation of communication calls. The European journal of neuroscience. 2015; 41(5):546–67. https://doi.org/10.1111/ejn.12812 PMID: 25728175.

41. Elie JE, Theunissen FE. Zebra finches identify individuals using vocal signatures unique to each call type. Nature communications. 2018; 9(1):4026. https://doi.org/10.1038/s41467-018-06394-9 PMID: 30279497

42. Shimazaki H, Shinomoto S. Kernel bandwidth optimization in spike rate estimation. J Comput Neurosci. 2010; 29(1–2):171–82. https://doi.org/10.1007/s10827-009-0180-4 PMID: 19655238.

43. Harremoes P. Binomial and Poisson distributions as maximum entropy distributions. Ieee Transactions on Information Theory. 2001; 47(5):2039–41. https://doi.org/10.1109/18.930936

44. Panzeri S, Treves A. Analytical estimates of limited sampling biases in different information measures. Network: Computation in Neural Systems. 1996; 7:87–107.

45. Averbeck BB, Latham PE, Pouget A. Neural correlations, population coding and computation. Nat Rev Neurosci. 2006; 7(5):358–66. https://doi.org/10.1038/nrn1888 PMID: 16760916.

46. Theunissen FE, Elie JE. Population Code, Noise Correlations, and Memory. Neuron. 2013; 78(2):209–10. https://doi.org/10.1016/j.neuron.2013.04.012 PMID: 23622058

47. Fortune ES, Margoliash D. Cytoarchitectonic organization and morphology of cells of the field L complex in male zebra finches (Taenopygia guttata). J Comp Neurol. 1992; 325(3):388–404. https://doi.org/10.1002/cne.903250306 PMID: 1447407

48. Vates GE, Broome BM, Mello CV, Nottebohm F. Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches (Taenopygia guttata). J Comp Neurol. 1996; 366:613–42. https://doi.org/10.1002/(SICI)1096-9861(19960318)366:4<613::AID-CNE5>3.0.CO;2-7 PMID: 8833113

49. Wang Y, Brzozowska-Prechtl A, Karten HJ. Laminar and columnar auditory cortex in avian brain. Proceedings of the National Academy of Sciences. 2010; 107(28):12676–81. https://doi.org/10.1073/pnas.1006645107 PMID: 20616034

50. Elliott TM, Theunissen FE. The Avian Auditory Pallium. Auditory Cortex 2011. p. 429–42.

51. Brown EN, Barbieri R, Ventura V, Kass RE, Frank LM. The time-rescaling theorem and its application to neural spike train data analysis. Neural Computation. 2002; 14(2):325–46. https://doi.org/10.1162/08997660252741149 PMID: 11802915

52. Kass RE, Ventura V, Brown EN. Statistical issues in the analysis of neuronal data. Journal of Neurophysiology. 2005; 94(1):8–25. https://doi.org/10.1152/jn.00648.2004 PMID: 15985692

53. Aldworth ZN, Miller JP, Gedeon TS, Cummins GI, Dimitrov AG. Dejittered spike-conditioned stimulus waveforms yield improved estimates of neuronal feature selectivity and spike-timing precision of sensory interneurons. Journal of Neuroscience. 2005; 25(22):5323–32. https://doi.org/10.1523/JNEUROSCI.0359-05.2005 PMID: 15930380

54. Victor JD, Purpura KP. Spike Metrics. In: Grun S, Rotter S, editors. Analysis of Parallel Spike Trains. Springer Series in Computational Neuroscience. 7. New York: Springer; 2010. p. 129–56.

55. Franco L, Rolls ET, Aggelopoulos NC, Treves A. The use of decoding to analyze the contribution to the information of the correlations between the firing of simultaneously recorded neurons. Exp Brain Res. 2004; 155(3):370–84. https://doi.org/10.1007/s00221-003-1737-5 PMID: 14722699

56. Quian Quiroga R, Panzeri S. Extracting information from neuronal populations: information theory and decoding approaches. Nat Rev Neurosci. 2009; 10(3):173–85. https://doi.org/10.1038/nrn2578 PMID: 19229240.

57. DeWeese MR, Wehr M, Zador AM. Binary spiking in auditory cortex. Journal of Neuroscience. 2003; 23 (21):7940–9. PMID: 12944525

58. Hamilton LS, Edwards E, Chang EF. A Spatial Map of Onset and Sustained Responses to Speech in the Human Superior Temporal Gyrus. Current Biology. 2018. https://doi.org/10.1016/j.cub.2018.04.033 PMID: 29861132

59. Nelken I, Rotman Y, Bar Yosef O. Responses of auditory-cortex neurons to structural features of natural sounds. Nature. 1999; 397(6715):154–7. https://doi.org/10.1038/16456 PMID: 9923676

60. Lewicki MS. Efficient coding of natural sounds. Nat Neurosci. 2002; 5(4):356–63. https://doi.org/10.1038/nn831 PMID: 11896400.

61. Woolley SM, Fremouw TE, Hsu A, Theunissen FE. Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. Nat Neurosci. 2005; 8(10):1371–9. https://doi.org/10.1038/nn1536 PMID: 16136039

62. Rodriguez FA, Chen C, Read HL, Escabi MA. Neural modulation tuning characteristics scale to efficiently encode natural sound statistics. J Neurosci. 2010; 30(47):15969–80. Epub 2010/11/26. https://doi.org/10.1523/JNEUROSCI.0966-10.2010 PMID: 21106835.

63. Santoro R, Moerel M, De Martino F, Goebel R, Ugurbil K, Yacoub E, et al. Encoding of Natural Sounds at Multiple Spectral and Temporal Resolutions in the Human Auditory Cortex. Plos Computational Biology. 2014; 10(1):14. https://doi.org/10.1371/journal.pcbi.1003412 PMID: 24391486

64. Theunissen FE, Elie JE. Neural processing of natural sounds. Nat Rev Neurosci. 2014; 15(6):355–66. https://doi.org/10.1038/nrn3731 PMID: 24840800.

65. Heil P, Irvine DRF. The posterior field P of cat auditory cortex: Coding of envelope transients. Cerebral Cortex. 1998; 8(2):125–41. https://doi.org/10.1093/cercor/8.2.125 PMID: 9542892

66. Liang L, Lu T, Wang X. Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. J Neurophysiol. 2002; 87(5):2237–61. https://doi.org/10.1152/jn.2002.87.5.2237 PMID: 11976364.

67. Furukawa S, Middlebrooks JC. Cortical representation of auditory space: information-bearing features of spike patterns. J Neurophysiol. 2002; 87(4):1749–62. https://doi.org/10.1152/jn.00491.2001 PMID: 11929896.

68. Mouterde SC, Elie JE, Mathevon N, Theunissen FE. Single Neurons in the Avian Auditory Cortex Encode Individual Identity and Propagation Distance in Naturally Degraded Communication Calls. Journal of Neuroscience. 2017; 37(13):3491–510. https://doi.org/10.1523/JNEUROSCI.2220-16.2017 PMID: 28235893

69. Town SM, Wood KC, Bizley JK. Sound identity is represented robustly in auditory cortex during perceptual constancy. Nature communications. 2018; 9(1):4786. https://doi.org/10.1038/s41467-018-07237-3 PMID: 30429465

70. VanRullen R, Guyonneau R, Thorpe SJ. Spike times make sense. Trends in Neurosciences. 2005; 28 (1):1–4. https://doi.org/10.1016/j.tins.2004.10.010 PMID: 15626490

71. Aldworth ZN, Dimitrov AG, Cummins GI, Gedeon T, Miller JP. Temporal Encoding in a Nervous System. Plos Computational Biology. 2011; 7(5). https://doi.org/10.1371/journal.pcbi.1002041 PMID: 21573206

72. Theunissen F, Miller JP. Temporal encoding in nervous systems: a rigorous definition. J Comput Neurosci. 1995; 2(2):149–62. Epub 1995/06/01. PMID: 8521284.

73. Sen K, Theunissen FE, Doupe AJ. Feature analysis of natural sounds in the songbird auditory forebrain. J Neurophysiol. 2001; 86(3):1445–58. Epub 2001/09/06. https://doi.org/10.1152/jn.2001.86.3.1445 PMID: 11535690.

74. Woolley SM, Gill PR, Fremouw T, Theunissen FE. Functional groups in the avian auditory system. J Neurosci. 2009; 29(9):2780–93. https://doi.org/10.1523/JNEUROSCI.2042-08.2009 PMID: 19261874.

75. Nagel K, Doupe A. Organizing Principles of Spectro-Temporal Encoding in the Avian Primary Auditory Area Field L. Neuron. 2008; 58(6):938–55. https://doi.org/10.1016/j.neuron.2008.04.028 PMID: 18579083

76. Kim G, Doupe A. Organized Representation of Spectrotemporal Features in Songbird Auditory Forebrain. Journal of Neuroscience. 2011; 31(47):16977–90. https://doi.org/10.1523/JNEUROSCI.2003-11.2011 PMID: 22114268

77. Chew SJ, Mello C, Nottebohm F, Jarvis E, Vicario DS. Decrements in auditory responses to a repeated conspecific song are long-lasting and require two periods of protein synthesis in the songbird forebrain. Proc Natl Acad Sci U S A. 1995; 92(8):3406–10. https://doi.org/10.1073/pnas.92.8.3406 PMID: 7724575

78. Nagel KI, Doupe AJ. Temporal Processing and Adaptation in the Songbird Auditory Forebrain. Neuron. 2006; 51(6):845–59. https://doi.org/10.1016/j.neuron.2006.08.030 PMID: 16982428

79. Bee MA, Micheyl C, Oxenham AJ, Klump GM. Neural adaptation to tone sequences in the songbird forebrain: patterns, determinants, and relation to the build-up of auditory streaming. J Comp Physiol A -Neuroethol Sens Neural Behav Physiol. 2010; 196(8):543–57.

80. Rieke F, Yamada W, Moortgat K, Lewis ER, Bialek W. Real time coding of complex sounds in the auditory cortex. Advances in Biosciences. 1992; 83:315–22.

81. Shamma S. On the role of space and time in auditory processing. Trends in Cognitive Sciences. 2001; 5(8):340–8. PMID: 11477003

82. Lu T, Wang X. Information content of auditory cortical responses to time-varying acoustic stimuli. J Neurophysiol. 2004; 91(1):301–13. Epub 2003 Oct 1. https://doi.org/10.1152/jn.00022.2003 PMID: 14523081

83. David SV, Shamma SA. Integration over Multiple Timescales in Primary Auditory Cortex. Journal of Neuroscience. 2013; 33(49):19154–66. https://doi.org/10.1523/JNEUROSCI.2270-13.2013 PMID: 24305812

84. Osman AF, Lee CM, Escabi MA, Read HL. A Hierarchy of Time Scales for Discriminating and Classifying the Temporal Shape of Sound in Three Auditory Cortical Fields. Journal of Neuroscience. 2018; 38 (31):6967–82. https://doi.org/10.1523/JNEUROSCI.2871-17.2018 PMID: 29954851

85. Ma SW, Ter Maat A, Gahr M. Power-law scaling of calling dynamics in zebra finches. Scientific reports. 2017; 7:11.

86. Bjoring MC, Meliza CD. A low-threshold potassium current enhances sparseness and reliability in a model of avian auditory cortex. Plos Computational Biology. 2019; 15(1). https://doi.org/10.1371/journal.pcbi.1006723 PMID: 30689626

87. Kaardal JT, Theunissen FE, Sharpee TO. A Low-Rank Method for Characterizing High-Level Neural Computations. Frontiers in computational neuroscience. 2017; 11(68). https://doi.org/10.3389/fncom.2017.00068 PMID: 28824408

88. Kozlov AS, Gentner TQ. Central auditory neurons have composite receptive fields. Proceedings of the National Academy of Sciences of the United States of America. 2016; 113(5):1441–6. https://doi.org/10.1073/pnas.1506903113 PMID: 26787894

89. Moore RC, Lee T, Theunissen FE. Noise-invariant Neurons in the Avian Auditory Cortex: Hearing the Song in Noise. Plos Computational Biology. 2013; 9(3).

90. Schneider DM, Woolley SMN. Sparse and Background-Invariant Coding of Vocalizations in Auditory Scenes. Neuron. 2013; 79(1):141–52. https://doi.org/10.1016/j.neuron.2013.04.038 PMID: 23849201

91. Carruthers IM, Laplagne DA, Jaegle A, Briguglio JJ, Mwilambwe-Tshilobo L, Natan RG, et al. Emergence of invariant representation of vocalizations in the auditory cortex. Journal of Neurophysiology. 2015; 114(5):2726–40. https://doi.org/10.1152/jn.00095.2015 PMID: 26311178

92. Meliza CD, Margoliash D. Emergence of selectivity and tolerance in the avian auditory cortex. J Neurosci. 2012; 32(43):15158–68. Epub 2012/10/27. https://doi.org/10.1523/JNEUROSCI.0845-12.2012 PMID: 23100437.

93. Billimoria CP, Kraus BJ, Narayan R, Maddox RK, Sen K. Invariance and sensitivity to intensity in neural discrimination of natural sounds. Journal of Neuroscience. 2008; 28(25):6304–8. https://doi.org/10.1523/JNEUROSCI.0961-08.2008 PMID: 18562600

94. Rothschild G, Nelken I, Mizrahi A. Functional organization and population dynamics in the mouse primary auditory cortex. Nature Neuroscience. 2010; 13(3):353–U21. https://doi.org/10.1038/nn.2484 PMID: 20118927

95. Bizley JK, Walker KM, Silverman BW, King AJ, Schnupp JW. Interdependent encoding of pitch, timbre, and spatial location in auditory cortex. J Neurosci. 2009; 29(7):2064–75. https://doi.org/10.1523/JNEUROSCI.4755-08.2009 PMID: 19228960.

96. Walker KMM, Bizley JK, King AJ, Schnupp JWH. Multiplexed and Robust Representations of Sound Features in Auditory Cortex. Journal of Neuroscience. 2011; 31(41):14565–76. https://doi.org/10.1523/JNEUROSCI.2074-11.2011 PMID: 21994373

97. Zann R. The Zebra Finch: A Synthesis of Field and Laboratory Studies. Oxford: Oxford University Press; 1996.

98. Dimitrov AG, Miller JP, Aldworth Z, Gedeon T. Non-uniform quantization of neural spike sequences through an information distortion measure. Neurocomputing. 2001; 38:175–81. https://doi.org/10.1016/s0925-2312(01)00441-6

99. Hsu A, Borst A, Theunissen FE. Quantifying variability in neural responses and its application for the validation of model predictions. Network. 2004; 15(2):91–109. Epub 2004/06/25. PMID: 15214701.