On the Origin of the Extracellular Field Potential in the Nucleus Laminaris of the Barn Owl (Tyto alba)

Paula T. Kuokkanen, Hermann Wagner, Go Ashida, Catherine E. Carr, and Richard Kempter

1Institute for Theoretical Biology, Department of Biology, Humboldt-Universität zu Berlin, Berlin, Germany; 2Institute for Biology II, RWTH Aachen, Aachen, Germany; 3Department of Biology, University of Maryland, College Park, Maryland; and 4Bernstein Center for Computational Neuroscience Berlin, Berlin, Germany

Submitted 28 April 2010; accepted in final form 1 August 2010

Kuokkanen PT, Wagner H, Ashida G, Carr CE, Kempter R. On the origin of the extracellular field potential in the nucleus laminaris of the barn owl (Tyto alba). J Neurophysiol 104: 2274–2290, 2010. First published August 4, 2010; doi:10.1152/jn.00395.2010. The neurophonic is a sound-evoked, frequency-following potential that can be recorded extracellularly in nucleus laminaris of the barn owl. The origin of the neurophonic, and thus the mechanisms that give rise to its exceptional temporal precision, has not yet been identified. Putative generators of the neurophonic are the activity of afferent axons, synaptic activation of laminaris neurons, or action potentials in laminaris neurons. To identify the generators, we analyzed the neurophonic in the high-frequency (>2.5 kHz) region of nucleus laminaris in response to monaural pure-tone stimulation. The amplitude of the neurophonic is typically in the millivolt range. The signal-to-noise ratio reaches values beyond 30 dB. To assess which generators could give rise to these large, synchronous extracellular potentials, we developed a computational model. Spike trains were produced by an inhomogeneous Poisson process and convolved with a spike waveform. The model explained the dependence of the simulated neurophonic on parameters such as the mean rate, the vector strength of synaptic activation of laminaris neurons, or action potentials in laminaris neurons. To find that several hundred sources are needed to reach the observed signal-to-noise ratio. The summed coherent signal from the densely packed afferent axons and activation of their synapses on laminaris neurons are alone sufficient to explain the measured properties of the neurophonic.

INTRODUCTION

Extracellular field potentials (EFPs) are important assays of neural function. Exact connections between the EFP and its neuronal generators, however, are unclear. EFPs were thought to originate exclusively from synaptic events (Buchwald et al. 1965; Mitzdorf 1985), until recent studies showed that they could also be composed of slow waveforms, including the afterpotentials of somato-dendritic spikes and membrane-potential oscillations (for review, see Logothetis and Wandell 2004). Unfiltered EFPs from many kinds of neuronal tissue show individual action potentials and slower changes in potential (e.g., Arezzo et al. 1977).

In the auditory system, an EFP that is well correlated with the acoustic stimulus is termed a neurophonic potential or “neurophonic” (Weinberger et al. 1970). The neurophonic has been observed in the auditory nerve (Snyder and Schreiner 1984; Wever and Bray 1930), in the cochlear nucleus (Marsh et al. 1970; Worden and Marsh 1968), in nucleus laminaris (NL) (Köppl and Carr 2008; Schwarz 1992; Sullivan and Konishi 1986; Wagner et al. 2005, 2009), in the lateral and medial superior olive (Bojanowski et al. 1989; Boudreau 1965; Marsh et al. 1974; Weinberger et al. 1970), and in the inferior colliculus (Marsh et al. 1974).

The origin of neurophonic potentials in the auditory system is still a matter of debate. In chicken NL as well as cat medial superior olive, the neurophonic is highest in amplitude in the vicinity of the densely packed neuron layer and therefore hypothesized to originate from synaptic currents (Guinan et al. 1972; Köppl and Carr 2008; McLaughlin et al. 2010; Schwarz 1992). In contrast, in the NL of barn owls, a prominent neurophonic is found throughout the tonotopically organized nucleus, where axon density is high and neuron density is low. Owls present an excellent model system for studying the EFP because the NL is large, homogeneous, and well organized. This nucleus has a dorso-ventral depth of ~700 μm, a medio-lateral width of ~2 mm, and an antero-posterior length of 3.5 mm. Compact NL neurons are sparsely and evenly distributed with a mean distance of ~100 μm (Carr and Konishi 1990: 75 μm is not corrected for shrinkage). In the high-frequency region (>2.5 kHz), NL neurons have very short stubby dendrites (Carr and Boudreau 1993b). Functionally, NL neurons encode interaural time differences (ITDs), which provide the main cue for azimuthal sound localization (Moiseff and Konishi 1981; Poganiatz et al. 2001). NL neurons detect the coincident arrival of phase-locked spikes from the ipsi- and contralateral nucleus magnocellularis (NM) (Carr and Konishi 1990). Despite our detailed knowledge, and the comparatively simple nature of the circuit in NL, the origin of the neurophonic in NL of the barn owl has remained elusive.

In the barn owl, the neurophonic potential in the NL is characterized by an outstanding temporal precision; the neurophonic responses to click stimuli exhibit an SD of the phase delay of ~10 ms (Wagner et al. 2005, 2009). Acoustic clicks produce robust and reliable responses; essentially identical responses emerge over hundreds of repetitions of the same click stimulus (Wagner et al. 2005, 2009). The neurophonic potential from pure tone stimulation is smooth and oscillatory (Sullivan and Konishi 1986) with amplitudes in the millivolt range. The implication of these properties is that a large number of spiking neuronal sources should contribute to the neurophonic. Even if each source has a firing rate in the range of hundreds of spikes per second, many sources are necessary in each stimulus cycle to generate a reliable and oscillatory time course. A more precise statement, however, requires a more quantitative approach.

Address for reprint requests and other correspondence: R. Kempter or P. T. Kuokkanen, Inst. for Theoretical Biology, Dept. of Biology, Humboldt-Universität zu Berlin, Invalidenstraße 43, D-10115 Berlin, Germany.
Identification of the source(s) of the neurophonic is important because they may reflect the computations performed in NL, which are essential for the localization of high-frequency sounds ≤10 kHz. Potential contributors to the neurophonic in the barn owl NL include the afferent axons from NM, synapses from NM axons onto NL neurons, and NL neurons and their afferent axons. To narrow down the source(s) of the neurophonic in owl NL, we developed and applied a new analysis technique for tone-driven responses. More specifically, we used the signal-to-noise ratio (SNR) to quantify the neurophonic potential in a way that is independent of its absolute amplitude and other characteristics, all of which may vary considerably across different recordings and different animals. We modeled the neurophonic potential as a sum of the contributions from different sources, taking into account physiological boundary conditions in NL. Such a model allowed us to numerically simulate the neurophonic potential and to analytically calculate its properties such as the SNR and its dependence on model parameters. A quantitative comparison of model and experiment led to a lower bound for the number of statistically independent sources that must contribute to the neurophonic potential. Finally, we were also able to estimate the spatial range over which an electrode collects a signal by including the different geometries related to axonal, synaptic, and somatic sources in the model.

METHODS

Experimental paradigm

Six barn owls (Tyto alba pratincola) were used to collect the physiological data presented in this study. The anatomical data in Fig. 8 are reanalyses from data published earlier (Carr and Boudreau 1993a,b) and from six additional owls used in parallel studies (Carr et al. 2010). The procedures described here conform to National Institute of Health guidelines for animal research and were approved by the Animal Care and Use Committee of the University of Maryland. Most animals were used in two or three separate physiology experiments, spaced ~1 wk apart.

Anesthesia was induced by injections of ketamine hydrochloride (3 mg/kg, im, Ketavet, Phoenix, St. Joseph, MO) and xylazine (2 mg/kg, im, Xyla-ject, Phoenix). Supplementary doses of ketamine and xylazine were administered to maintain a suitable plane of anesthesia. Body temperature was measured with a cloacal probe inserted and maintained at 39°C by a feedback-controlled heating blanket (Harvard Instruments, Braintree, MA). Buprenorphine hydrochloride (0.3 mg/kg, im, Buprenex, Reckitt and Colman Products, Richmond, VA) was administered at the end of each recovery experiment.

Surgery and stereotaxics

Initially, the owl’s head was placed in a custom-designed stereotaxic frame and stabilized using ear bars and a beak holder. A metal headplate and a short metal pin marking a standardized zero point were permanently glued to the skull. After this, the ear bars and the beak holder were removed, and the head was held by the headplate alone. An opening was made in the skull around the desired area relative to the zero point, and the dura was cut open. Each electrode was moved in defined amounts in the rostrocaudal and mediolateral axes before being driven down into the brain. In some cases, the electrode was angled to facilitate access to extremely medial or lateral regions of the brain stem.

Electrodes and recording setup

Owls were placed on a vibration-isolated table within a sound-attenuating chamber (IAC, New York, NY). Commercial Epoxyylite-coated tungsten electrodes (Frederick Haer) were used, with impedances between 2 and 20 MΩ. A silver chloride pellet, placed under the animal’s skin around the incision, served as the reference electrode (WPI, Sarasota, FL). Electrode signals were amplified and band-pass filtered (100–13,000 Hz) by a custom-built headstage and amplifier. The noise floor of the equipment was in the range of 1–10 μV. The recording was passed in parallel to an oscilloscope, a threshold discriminator [SD1, Tucker-Davis Technologies (TDT), Gainesville, FL], and an A/D converter (DD1, TDT) connected to a personal computer via an optical interface (OI, TDT). A continuously refreshed, software-generated display of the waveforms that triggered TTL pulses aided in trigger judgment. Analog waveforms were saved for off-line analysis.

Stimulation and recording

While lowering an electrode in the brain toward NL, noise bursts were presented as search stimuli. Once auditory responses were discernible, tonal stimuli were applied from both the ipsi- and the contralateral side to measure best frequency and to estimate the position of the electrode. To record the neurophonic, tone bursts of different frequencies (500 Hz to 10 kHz) were presented monaurally. Tone bursts were repeated three to five times. The duration of a tone burst was 100 ms, with 5-ms rise/fall times and a constant starting phase. In total, 200 ms of the response were saved per trial (single trial data shown in Fig. 1A). The stimulus began 5 ms after the recording started and ended at 105 ms. The level of the tones was generally 20–30 dB above threshold. The interstimulus interval was either 500 or 700 ms.

Data were recorded from 171 locations in the NL of six anesthetized barn owls. We acquired data in one to three recording sessions in each owl and in each session from only one of the NLs. There were one to seven penetrations in each session, and each penetration contained 1–28 dorso-ventrally distributed recording locations. In a single penetration, the distance between the recording locations was generally ≥100 μm (~2/3 of the locations). In four penetrations, we used distances of 10 or 15 μm. Both ipsi- and contralateral ear were stimulated for almost all locations, sometimes repeatedly. In what follows, one “recording site” means a recording of the response to monaural stimulation. We analyzed 378 recording sites altogether.

Anatomical analyses

Golgi data and the plastic sections through NL were published in Carr and Boudreau (1993b) and reanalyzed for this study, e.g., to correct for the shrinkage of the tissue. The dorsal brain stem was reconstructed to determine the size and volume of NL. For normal light microscopy, three barn owls were anesthetized and perfused as above, and the brains were sectioned on a freezing microtome at 40 μm thickness. Two brains were cut in the coronal plane and one in the horizontal plane. Sections were collected in order in phosphate buffer, mounted, Nissl stained, and coverslipped. Every third section was traced on a computer connected to a microscope. We used an Olympus BX-60 microscope equipped with a motorized stage drive (LEP Mac 5000), and coupled to a PC containing the NeuroLucida software (Microbrightfield, Colchester, VT) through the Lucivid system for morphometry or through a digital camera (DVC, Austin, TX) for photomicrograph acquisition. Cytoarchitectonic boundaries were determined in Nissl-stained sections and
reconciled for 10% shrinkage, determined by comparing with Araldite-embedded sections, as follows. The dorso-ventral depth of the central region of the nucleus laminaris was measured from both Nissl sections [635 ± 36 (SD) μm, 15 sections, 3 owls] and from Araldite-embedded sections (634 ± 38 μm, 15 sections, 2 owls). The nucleus laminaris is ~100 μm thicker rostrally and ~100 μm thinner caudally, and sections throughout the rostrocaudal extent of NL had a mean depth of 640 ± 190 μm (26 sections, 1 owl). With 10% shrinkage for Araldite (Kushida 1962), this produces an estimated depth of 705 μm, consistent with physiological measurements (Carr and Konishi 1990).

Axon and neuron density were measured in NL, using osmium treated Araldite embedded material stained with Toluidine blue, and prepared for Carr and Boudreau (1993a,b). All brains were sectioned in the transverse plane, orthogonal to the major axis of the brain stem (Carr and Boudreau 1993a,b; Carr and Konishi 1990). Some of the prepared 200 μm plastic sections were sectioned orthogonal to the delay line axons, i.e., orthogonal to the transverse section plane and also parallel to the dorsal and ventral borders of the NL, to generate the most accurate estimate of maximum axon density. To measure axon diameters in and around the NL, a detailed morphometric analysis was performed. Axon diameters were measured from myelinated profiles cut in cross-section, whereas the mean Feret’s diameter was used for fibers cut on a tangent. Feret’s diameter takes the value of the distance measured across minimum and maximum tangents across a particle and avoids choosing between maximum and minimum diameters.

Images of NL axons were captured at 2,000

*real number* diameters were measured from myelinated profiles cut in cross-section plane and also parallel to the dorsal and ventral borders of the NL, to generate the most accurate estimate of maximum axon density. To measure axon diameters in and around the NL, a detailed morphometric analysis was performed. Axon diameters were measured from myelinated profiles cut in cross-section, whereas the mean Feret’s diameter was used for fibers cut on a tangent. Feret’s diameter takes the value of the distance measured across minimum and maximum tangents across a particle and avoids choosing between maximum and minimum diameters. Images of NL axons were captured at 2,000× (Neurulida) and analyzed using interactive profile recognition. All measurements were corrected for 10% tissue shrinkage for aldehyde fixation and plastic embedding (Kushida 1962).

**Analysis of the measured neurophonic**

**FOURIER TRANSFORM AND POWER SPECTRAL DENSITY.** To characterize the neurophonic, we calculated the Fourier transform for time-discrete data points $x_i$ where $i = 1, \ldots, n$ is a sample index and $n$ is the total number of samples. For a time resolution $\Delta_t$ and $T = n \Delta_t$ being the width of the analyzed time window, the Fourier transform of $x_i$ is

$$\tilde{x}_i = \sqrt{T} \sum_{n=1}^{n} x_i \exp(-i \pi f / \Delta_t)$$

(1)

where the frequency $f$ is an integer multiple of $1/T$. The frequency resolution therefore depends on the length $T$ of the analyzed time window; we always took an 80 ms time interval, which corresponds to a frequency resolution of 12.5 Hz. Also the stimulus frequencies $f_{stim}$ were always multiples of 12.5 Hz. The temporal resolution of the recorded data was $\Delta_t = 1/48,077$ s or 20.80 μs. The data were resampled to 50,000 Hz (MATLAB-function “interp1” with default settings, MathWorks, Natick, MA) to have an integer number of samples in the 80 ms segment.

The power spectral density (PSD) at frequency $f$ is the square of the absolute value of the Fourier transform

$$P(f) = |\tilde{x}_i|^2$$

(2)

The PSD has units of the data squared, divided by frequency; for example, mV²/kHz for the EFP. We calculated the PSD using the MATLAB function “periodogram.”

In our recordings of the neurophonic in the barn owl NL, the largest response amplitude was typically obtained when the stimulus frequency was close to the best frequency (BF, see the definition in RESULTS) at the recording site. Only those recording sites at which the PSD showed a clear, distinct peak at that particular stimulus frequency were accepted for further analysis. A response peak in the PSD of $\geq0.16$ mV²/kHz was needed to fulfill this criterion.

**SIGNAL-TO-NOISE RATIO.** The signal-to-noise ratio (SNR) is the ratio of the signal PSD to that of the noise PSD. For tonal stimuli, we assumed to find the signal PSD at the stimulus frequency, $f = f_{stim}$ and the noise PSD at neighboring frequencies, $f \neq f_{stim}$. Thus

$$\text{SNR}(f_{stim}) = \frac{P(f_{stim})}{P(f)}$$

(3)

To estimate the noise level, we averaged the noise PSD amplitudes over an interval of 1 kHz around the stimulus frequency, excluding the peak at the stimulus frequency. The signal level is defined as the PSD level at the stimulus frequency minus the noise level.

We defined signal and noise in the frequency domain, but separation of the two components is also possible in the time domain. The common way to extract the signal component from a response in the temporal domain is to average over many trials obtained with an identical stimulus. This procedure reduced the noise. However, here we recorded only three to five trials for each stimulus frequency at each site. An average over such a low number of trials was insufficient to considerably reduce the noise. Therefore we took advantage of having periodic stimuli (tones with frequency $f_{stim}$): each trial contained many stimulus cycles of length $1/f_{stim}$ typically hundreds.

To estimate the signal, we selected an 80 ms interval of the response to be analyzed, for example from 10 to 90 ms after stimulus onset. The signal component was assumed to be cyclic with period $1/f_{stim}$. To average across cycles of a waveform that had a sampling frequency of 50,000 Hz, we resampled the 80-ms interval to 10, 20, or 40 sampling points per stimulus period; for $f_{stim} \geq 5$ kHz, we used only 10 or 20 sampling points per period. Resampling enabled us to average the response cycle-by-cycle (as well as across trials), which yielded the “cyclic-mean” waveform (Fig. 4B). Using the cyclic-mean waveform, we generated a signal waveform of 80 ms length by concatenating identical copies of the cyclic-mean and restoring the original sampling frequency (Fig. 4C). To extract the noise, we subtracted the 80 ms signal waveform from the original response (Fig. 4D).

**Computational model of the neurophonic in NL**

One aim of modeling the neurophonic potential was to numerically simulate a voltage signal that resembles the experimentally measured one. A similar model has been used by Ashida et al. (2007) to analyze the intracellular potential in NL neurons. Here we describe the time course of the extracellular field potential $V(t)$ as the sum of waveforms from $N$ sources

$$V(t) = \sum_{n=1}^{N} A_n \sum_{i=1}^{k_n} k_n(t - t_{n,i})$$

(4)

where $t$ is time and $A_n$ is the relative amplitude of source $n$ for $n = 1, \ldots, N$. The symbol $k_n$ denotes the time-dependent kernel describing the spike waveform associated with the source $n$, and $t_{n,i}$ is the time of the $i$th spike of source $n$. In what follows, we drop the index $n$ of the kernel and consider only one source type, i.e., one spike waveform, at a time.

The distributions of amplitudes $A_n$ were chosen to simulate two cases. In the first and simplest case, amplitudes $A_n$ were identical for all sources, i.e., there was no distance dependence decay. In such an idealized scenario, all sources were located at a similar distance from the tip of the electrode. In a second case, we assumed a spatially uniform distribution of sources. Furthermore, the amplitude of a source decayed with its distance $r$ from the electrode tip. For a neuronal dipole, the generic geometric configuration where current sink and source are spatially separated, the amplitude $A_n(r)$ was proportional to $1/r^2$ (Logothetis et al. 2007) for large $r$.

The kernel function $k(t)$ was chosen such that its PSD roughly matched the measured PSD in Fig. 3D. The kernel’s phase spectrum was estimated from typical extracellular spike waveforms as pre-
sented, for example, by Gold et al. (2006). To fulfill these criteria in a simple model, we approximated the shape of the spike waveform \( k \) with a Gabor function

\[
  k(t) = \exp\left(-\frac{t^2}{2\rho^2}\right)\exp\left(2\pi f_g t + \varphi\right)
\]

with width \( \rho = 0.09 \) ms, oscillation frequency \( f_g = 3.9 \) kHz, and phase \( \varphi = 0.8 \) rad (Fig. 6B). Absolute peak amplitudes of the kernels were in the range of 100 \( \mu \)V (Gold et al. 2006). We assumed that the shapes of the kernels \( k \) are equal for all sources of one specific type (axon, synapse, or neuron). Any filtering resulting from the experimental setup was assumed to be included in the form of the kernel, i.e., the shape of the kernel was assumed to undergo the same filtering as the data. Furthermore, we neglected the small equipment noise included in the spectrum.

The time \( t_{n,i} \) denotes spike number \( i \) in source \( n \). Spikes were generated through an inhomogeneous Poisson process with a time-dependent rate \( p(t) \). To describe the sustained phase of a neurophonic driven by a monaural tone at frequency \( f_{\text{stim}} \), we accounted for three constraints as motivated by the physiology of NL. First, the time-dependent rate \( p(t) \) oscillated with the stimulus frequency \( f_{\text{stim}} \). Second, spikes were phase locked to the stimulus with some vector strength. Because different sources could be locked to a tonal stimulus at different absolute phases, the resulting vector strength of spikes from a population of sources might be smaller than the individual vector strengths of single sources. To take such a phase jitter into account, we described phase locking by a population vector strength \( v \) (also called synchronization index), which is taken as the vector strength of each source, where all sources were assumed to be locked to the same stimulus phase. Third, we assumed that spikes in all sources were mutually independent, i.e., each individual source generated spikes according to an inhomogeneous Poisson process.

To describe the spiking probability for a tonal stimulus at frequency \( f_{\text{stim}} \) we used an \( \nu_{\text{stim}} \)– periodic sum of Gaussian functions (wrapped Gaussian function as in Fig. 6A; see also Jammalamadaka and SenGupta 2001) forming the Poisson rate

\[
p(t) = \lambda / f_{\text{stim}} \sum_{m=-\infty}^{\infty} G_{\lambda}(t - m/f_{\text{stim}})
\]

where \( \lambda \) is the mean rate of the Poisson process and \( G_{\lambda}(x) = 1/\sqrt{2\pi}\sigma \exp\left[-x^2/(2\sigma^2)\right] \) denotes a normalized Gaussian with zero mean and SD \( \sigma > 0 \). The phase locking of spikes, as quantified by the vector strength \( v \), was related to the width \( \sigma \) of the Gaussian through (Kempter et al. 1998b)

\[
v = \exp\left[-\frac{1}{2}(2\pi f_{\text{stim}})^2\sigma^2\right]
\]

To describe spontaneous activity in the absence of a tonal stimulus, we took \( v = 0 \), that is, the Poisson firing rate is constant: \( p(t) = \lambda \). Numerical simulations of these computational models of the neurophonic in NL are shown in Figs. 6, 7, 9, and 10.

Mathematical analysis of the computational model

To interpret the measured neurophonic data, we compared it to the simulated neurophonic. It was important to know how, for example, the cyclic-mean amplitude and the SNR of the simulated neurophonic depended on simulation parameters, i.e., the spike waveform \( k \), the population vector strength \( v \), the stimulus frequency \( f_{\text{stim}} \), the number \( N \) of sources, and their mean firing rate \( \lambda \). Model spikes were generated by an inhomogeneous Poisson process. To understand the behavior of the numerical simulations of the model, we analyzed the model mathematically.

The analysis was simplest for kernels being equal for all sources. The expected amplitude of the cyclic-mean response was approximated by the amplitude \( a \) of the first harmonic, that is, the Fourier transform at \( f = f_{\text{stim}} \) of the simulated neurophonic. This approximation was reasonable for \( v < 0.5 \), because higher harmonics could be neglected. With the definition of the Fourier transform in Eq. 1, we found that the amplitude \( a \) is (Kempter et al. 1998b)

\[
a = 2\pi v \left| \tilde{k}(f_{\text{stim}}) \right|
\]

indicating its linear dependence on \( N, \lambda, \) and \( v \) (Fig. 7B). Note that the Fourier transform of a square-integrable function, for example, the spike waveform \( k \), is defined as

\[
\tilde{k}(f) = \int_{-\infty}^{+\infty} dt\ k(t) \exp(-i2\pi ft)
\]

In the following equations, the absolute value of this Fourier transform is important to derive PSDs. For a kernel as in Fig. 6B with a peak amplitude of 100 \( \mu \)V, we obtained \( \left|\tilde{k}(5\text{kHz})\right| = 1.46 \times 10^{-2} \text{mV/kHz} \).

Because the signal power at the stimulus frequency \( f_{\text{stim}} \) was equal to \( Ta^2/4 \) (derived from Eqs. 1 and 2 for a sinusoidal signal with an amplitude \( a \)), we found from Eq. 7 that the PSD at the stimulus frequency, \( f = f_{\text{stim}} \), is

\[
P(f_{\text{stim}}) = N^2 \lambda^2 v^2 |\tilde{k}(f_{\text{stim}})|^2 T
\]

where \( T \) is the duration of the analyzed response, which is 80 ms throughout this paper. For \( f \neq f_{\text{stim}} \), the PSD could be calculated similarly (see also Snyder and Miller 1991) but had a different dependence on the parameters. In summary, we found

\[
P(f) = \begin{cases} N^2 \lambda^2 v^2 |\tilde{k}(f)|^2 T & \text{for } f = f_{\text{stim}} \\ N \lambda |\tilde{k}(f)|^2 & \text{for } f \neq f_{\text{stim}} \end{cases}
\]

From the definition of the SNR in Eq. 3 we found

\[
\text{SNR} = N\lambda^2 T
\]

Note that the SNR is independent of the kernel \( k \). The SNR linearly depends on the number \( N \) of independent sources (Fig. 7A) and the mean population firing rate \( \lambda \), and quadratically depends on the vector strength \( v \). Similar analytical expressions of the SNR for inhomogeneous Poisson spike trains have been derived previously in the context of stochastic resonance (Gammaitoni et al. 1998; Hohn and Burkitt 2001; Lindner et al. 2009; McNamara and Wiesenfeld 1989; Wiesenfeld et al. 1994; Shimokawa et al. 1999). Because the rate \( \lambda \) and the vector strength \( v \) may depend on the stimulus frequency \( f_{\text{stim}} \), the SNR may also depend on \( f_{\text{stim}} \), as seen in SNR tuning curves (Fig. 4F).

Solving Eq. 11 for \( N \), we obtained

\[
N = \frac{\text{SNR}}{\lambda^2 T v^2}
\]

which allowed us to estimate the number \( N \) of sources that contribute to the neurophonic from experimentally accessible quantities. We note that \( N \) in Eq. 12 provided a lower bound; weakening the model assumptions, for example, by allowing that the kernels are different for different sources or that sources are not independent, only increased the necessary number of sources to reach a certain SNR, even if \( \lambda, v, \) and \( T \) are unchanged.

To show the effect of nonidentical kernels for different sources, we analytically calculated the cyclic-mean amplitude \( a \) and the SNR for the case of distance-dependent kernel amplitudes \( A_r(r) \). We therefore assumed a uniform distribution of sources in space. The density is \( 3/(4\pi r_0^3) \), where \( 2r_0 \) is a measure for the mean distance between neighbors. We further assumed that the kernel of a source at distance \( r_0 \) from the electrode had the form \( k(t) \) and relative amplitude \( A_r(r_0) = 1 \); cf. Eq. 4. We finally assumed that the kernel amplitudes decreased with distance \( r \) to the electrode via \( 1/r^2 \) (Logothetis et al. 2007), that is, \( A_r(r) = (r_0/r)^2 \), but the waveform \( k(t) \) remained unchanged.

\[
J\ Neurophysiol\ •\ VOL\ 104\ •\ OCTOBER\ 2010\ •\ www.jn.org
\]
We derived the resulting cyclic-mean amplitude \( a \) in two steps. First, we calculated the amplitude resulting from a thin spherical shell around the electrode and, second, we added contributions from shells. A shell at radius \( r \) and with infinitesimal thickness \( dr \) had a volume \( 4\pi r^2 dr \). Multiplying this volume by the density \( 3/(4\pi r^3) \) of sources resulted in \( 3\pi dr/r^3 \) sources. The sources’ amplitude was \((r/s)^2\), and amplitudes of all sources were equal because they were at equal distance \( r \) from the electrode. The resulting contribution \( da \) to the total amplitude was

\[
da = 6\pi aR^2 \int f_{\text{stim}}(r) \frac{dr}{r_s^3}
\]

This equation is analogous to Eq. 7, in which we replaced the number \( N \) of sources by \( 3\pi dr/r^3 \) and the amplitude factor 1 by \((r/s)^2\). We integrated this equation to sum up contributions of sources within a sphere of radius \( R \), which yielded the expected amplitude \( a \) of the first harmonic of the summed signal

\[
a(R) = 6\pi \int f_{\text{stim}}(r) \frac{dr}{r_s^3} \int_0^R \frac{dr}{r_s^3}
\]

The upper integration limit was \( R \), which typically was much larger than \( r_s \). As the lower integration limit we chose \( r_s \) and not 0. The difference is negligible for large numbers of sources; however, this choice was necessary to avoid divergent integrals in the calculation of the noise power at distances \( r < r_s \), at which the \( 1/r^2 \) dependence of amplitudes does not hold. Therefore the expected cyclic-mean amplitude is

\[
a(R) = 6\pi \int f_{\text{stim}}(r) \frac{dr}{r_s^3} \int_0^R \frac{dr}{r_s^3} = 6\pi \left( \frac{R}{r_s} - 1 \right) \quad (13)
\]

where \( R/r_s = N^{1/3} \) connects the radius \( R \) of the sphere to the number \( N \) of sources within this sphere (Fig. 7B). The signal power at the stimulus frequency \( f_{\text{stim}} \), was, as before, equal to \( T_a(R)N^{2/3} \), and thus

\[
P(f_{\text{stim}}) = 9\pi a^2 \int f_{\text{stim}}(r) \frac{dr}{r_s^3} \int_0^R \frac{dr}{r_s^3} = 9\pi \left( \frac{R}{r_s} - 1 \right)^2 T
\]

It increased almost quadratically with increasing \( R \), and therefore was global.

To calculate the noise power, we summed noise contributions from thin spherical shells, similar to above approach for the amplitude. By analogy to Eq. 10 for \( f \neq f_{\text{stim}} \), we replaced the number \( N \) of sources by \( 3\pi dr/r^3 \) and the amplitude factor 1 by \((r/s)^2\), where the fourth power accounts for the fact that we summed up noise powers (and not amplitudes). An integration yielded the noise power

\[
P(f) = 3\pi \int f_{\text{stim}}(r) \frac{dr}{r_s^3} \int_0^R \frac{dr}{r_s^3} = 3\pi \left( \frac{R}{r_s} - 1 \right) \quad (14)
\]

which saturated with increasing \( R \) and was therefore local. The SNR was

\[
\text{SNR} = 3\pi \left( \frac{R}{r_s} - 1 \right) \quad (14)
\]

Because \( N = (R/r_s)^3 \), we found from Eq. 14 the dependence of the SNR on \( N \)

\[
\text{SNR} = 3\pi a^2 T N^{1/3} (N^{1/3} - 1) \quad (15)
\]

for \( N > 1 \) (Fig. 7A), or, equivalently

\[
N = \left( \frac{1}{8} + \sqrt{1 + \frac{4 \cdot \text{SNR}}{3\pi a^2 T}} \right)^3 \quad (16)
\]

RESULTS

In this study, we characterize the general structure of the neurophonic potential in the NL of owls, using newly developed analysis tools. We derive a computational model that explains its most salient properties. Finally, we focus on the possible origin of the neural responses.

Properties of the neurophonic

We analyzed EFP recordings from 378 sites in the auditory brain stem of six anesthetized barn owls. Neurophonic responses to tones were recorded in the mid-to-high BF region \( (>2.5 \text{ kHz}) \) of the tonotopically organized NL. When the frequency of the tonal stimulus was close to the BF at the recording site, the response typically contained a strong oscillatory component (Fig. 1, A and B). For the quantitative analysis, we only accepted recording sites that exhibited a response peak in the PSD of \( \approx 0.16 \text{ mV}^2/\text{kHz} \) (see METHODS). Responses at 47 sites did not fulfill this criterion, leaving responses at 331 sites. All but one of the recording sites with nonoscillatory responses were from the same owl (owl 120).

GENERAL TEMPORAL STRUCTURE OF THE NEUROPHONIC POTENTIAL. After a short transient response at the onset (Fig. 1B), the driven response reached a tonic level that lasted for the duration of the stimulus (Fig. 1C). The neurophonic potential was notable for both its coherence and large amplitude and displayed a smooth time course without noticeable unitary events; single units could not be isolated. After the end of the tonal stimulus, the spontaneous activity was reduced for \( \approx 10 \) ms (Fig. 1D) before it returned to the prestimulus level. The spontaneous activity was smaller in amplitude than the driven activity (Fig. 1E). In what follows, we focus on the sustained responses and analyze 80 ms intervals from 10 to 90 ms after stimulus onset (black line in Fig. 1A) and 10 to 90 ms after the stimulus offset at 100 ms.

AMPLITUDE OF THE NEUROPHONIC POTENTIAL IN THE TIME DOMAIN. To quantify the magnitude of the EFP amplitude, we used the SD. In the example shown in Fig. 1, the SD of the sustained component of the neurophonic potential, averaged over the 80 ms interval, was 0.34 mV. For the whole sample, when the stimulus frequency was equal to the BF of each recording site, the SD varied between 0.02 and 2.66 mV, with a median of 0.37 mV (Fig. 2A).

Likewise, the spontaneous activity was characterized by its SD (average over all repetitions of 80 ms intervals: 0.15 mV in Fig. 1). For the whole sample, the SD of the spontaneous activity varied between 0.01 and 0.97 mV, with a median of 0.17 mV (Fig. 2B). The SDs for driven (at BF) and spontaneous activity (Fig. 2, A and B) were highly correlated \((r = 0.946, P < 10^{-100})\). Both distributions of the SD had several modes because values in different owls covered different ranges, possibly also because of variation in electrode impedance. This high variability of the SD impedes a direct comparison of the EFP across animals. The analysis tools developed later in this study overcome this problem.

BEST FREQUENCIES. Iso-intensity tuning curves were assembled to further characterize recording sites: We derived the SD of the neurophonic in response to stimulation with tones at different frequencies (Fig. 2C). In general, the SD as a function of the stimulus frequency showed a clear maximum and a monotonic decay on both sides of the maximum until the spontaneous level was reached. At some recording sites, like the one shown in Fig. 2C, the decrease was nonmonotonic and had a second, smaller peak.
Iso-intensity tuning curves defined the BF at a recording site as follows: a line at half height of a tuning curve was derived from its peak value and the mean value of the spontaneous levels. The midpoint of the line at half height yielded the BF, which was 5.0 kHz in the example shown in Fig. 2C. The BFs in the whole sample ranged from 2.8 to 6.3 kHz (4.7 ± 0.7 kHz; Fig. 2D). Furthermore, there was a low but significant correlation between the BF and SD at BF ($r = 0.26, P < 10^{-5}$); however, the distribution of BFs was unimodal. In the following, further properties of the neurophonic will be derived for tonal stimulation at the BF if not stated otherwise.

AMPLITUDE OF THE NEUROPHONIC POTENTIAL IN THE FREQUENCY DOMAIN. In response to tones at BF, the time course of the neurophonic was similar across trials. In particular, responses were locked to the periodic stimulus (Fig. 3A), whereas spontaneous activity was not coherent (Fig. 3C).

To characterize the phase locking of the EFP, we computed the power spectral density (PSD, in units of mV^2/kHz). In the example depicted in Fig. 3B, the PSD exhibited a large peak at the stimulus frequency (signal level). The level of the PSD in this frequency bin was 2.181 mV^2/kHz, which is much larger than the average level in surrounding frequency bins (noise level, ~0.01 mV^2/kHz).

Let us briefly discuss the shape of the PSD as a function of frequency, disregarding the signal peak. In the example shown in Fig. 3B, the level increased from 0.5 to 3 kHz by about one order of magnitude, had a shallow maximum between 3 and 5 kHz, and decreased by about two orders of magnitude from 5 to 10 kHz, which was the highest frequency considered in this analysis. The PSD was noisy and did not display any further salient features. Spectra obtained at other recording sites had similar properties.

The PSD of the spontaneous activity (Fig. 3D) closely resembled the shape of the PSD of the driven activity (Fig. 3B), apart from the absence of the peak at the stimulus frequency. The average level was lower in the spontaneous PSD: ~2.5 × 10^{-3} mV^2/kHz at 5 kHz. Similar shapes but different levels of
driven (without the signal peak) and spontaneous PSDs were also typical for the whole sample.

**Separation of Signal and Noise.** We separated the signal and the noise in the time domain, even though the separation in the frequency domain would be equivalent for responses in which the signal is sinusoidal (see Methods for details). The separation yielded the cyclic-mean response in the temporal domain, as well as the level of the signal and the noise PSDs in the frequency domain. In the example shown in Fig. 4, A–C, the amplitude of the sinusoidal cyclic-mean response was 0.32 mV. In the population, the cyclic-mean amplitude ranged from 0.01 to 2.80 mV, with a median of 0.31 mV and an interquartile range of 0.44 mV.

The PSD of the reconstructed sinusoidal signal showed a dominant peak at the stimulus frequency $f_{\text{stim}}$ (Fig. 4C, right). The height of this peak was, by definition, identical to the peak of the response PSD minus the noise level (signal level, $2.17 \text{ mV}^2/\text{kHz}$ = $2.18 - 0.01 \text{ mV}^2/\text{kHz}$). There was also a smaller peak at the 2nd harmonic frequency $2f_{\text{stim}}$ (arrow and black dashed line at 2.18 mV$^2$/kHz) and the average noise level (gray dashed line at 0.01 mV$^2$/kHz). C: spontaneous activity at the same recording site as in A, 25–30 ms after stimulus onset. D: mean PSD of the spontaneous response. The dashed line shows the same level as in B (noise). The shown example is from recording site 17.10 in owl 502. In A and B, a tonal stimulus at 5.0 kHz at intensity of 60 dB SPL was applied.

**Fig. 3.** Trial-to-trial variability in driven and spontaneous activity and power spectral density (PSD) of the neurophonic. A: 3 repetitions of the driven response, 25–30 ms after stimulus onset. B: mean PSD of the driven response in A, based on activity 15–95 ms after the onset of the tone; note the large signal peak at the stimulus frequency $f_{\text{stim}}$ (arrow and black dashed line at 2.18 mV$^2$/kHz) and the average noise level (gray dashed line at 0.01 mV$^2$/kHz). C: spontaneous activity at the same recording site as in A, 25–30 ms after stimulus offset.

**Fig. 4.** Signal-to-noise ratio (SNR) of responses to tonal stimuli. A: response: driven activity (left) and its PSD (right) for a tonal stimulus at frequency $f_{\text{stim}} = 5.0 \text{ kHz}$. The two horizontal dashed lines indicate the levels of the peak of the PSD at $f_{\text{stim}}$ (signal) and the mean PSD around $f_{\text{stim}}$ (noise). B: cyclic-mean: a response is wrapped to 1 cycle (width $1/f_{\text{stim}}$). The wrapped segments (gray lines) are averaged (cyclic-mean, black line). C: signal: several cycles of the same cyclic-mean are shown in the time trace (left); the PSD (right) has a peak at $f_{\text{stim}}$ and a smaller peak at the 2nd harmonic frequency $2f_{\text{stim}}$ (arrow). D: noise (left) is derived by subtracting the signal in C from the response in A. Noise is calculated for each trial separately. The PSD of the noise (right) is an average over all trials. E: signal and noise levels as a function of stimulus frequency. The vertical gray line indicates the stimulus frequency $f_{\text{stim}} = 5.0 \text{ kHz}$ used in A. Stimulus frequencies range from $f_{\text{stim}} = 3.0 - 7.0 \text{ kHz}$ with 50 Hz steps. F: SNR (in units of dB) as a function of stimulus frequency. The example shown is from the recording site 17.10 in owl 502.
peak at the frequency of the second harmonic \((1 \times 10^{-4} \text{ mV}^2/\text{kHz})\) at \(2 f_{\text{stim}} = 10 \text{ kHz}\). The level of the second harmonic was only \(-45 \text{ dB}\) (or 0.003%) compared with the signal level at \(f_{\text{stim}}\), indicating a highly sinusoidal signal (Fig. 4C, right). In the whole population of 331 recording sites, the cyclic-mean response at BF was almost always sinusoidal: the level of the second harmonic was less than \(-20 \text{ dB}\) with respect to the fundamental in all but 10 cases. The median level of the second harmonic was \(-47.0 \text{ dB}\) in comparison to the fundamental, with an interquartile range of 12.3 \text{ dB}.

The spectrum of the noise (Fig. 4D, right) showed all components of the response’s PSD except the large peak at \(f_{\text{stim}}\). The noise level was obtained from the average over 41 data points (corresponding to 1 kHz) of the PSD around the stimulus frequency and an average over all trials \((9.61 \times 10^{-3} \text{ mV}^2/\text{kHz}; \text{dashed line in Fig. 4D, right})\).

To further characterize the neurophonic at a recording site, tuning curves of signal and noise levels were assembled. To this end, we measured signal levels and noise levels at one recording site for various values of the stimulus frequency \(f_{\text{stim}}\) (Fig. 4E). In general, the signal level was larger than the noise level for stimulus frequencies around the BF of the recording site: the signal level showed a maximum near the BF and declined (usually monotonically) at both sides of the maximum. The shape of the tuning curve in Fig. 4E was similar to the tuning curve derived from the SD (Fig. 2C, note the logarithmic ordinate scale in Fig. 4E).

The separation of signal and noise enabled us to calculate the SNR. The SNR provided the best measure to quantify the neurophonic. For the recording site described in Fig. 4, A–D, with responses driven by a tonal stimulus at frequency \(f_{\text{stim}} = 5.0 \text{ kHz}\), the SNR was 412.7, corresponding to 26.2 \text{ dB} (Fig. 4F). The frequency at which the highest SNR occurred was lower than the BF (maximum SNR of 28.1 \text{ dB} achieved at \(f_{\text{stim}} = 4.4 \text{ kHz}\) in this example.

**Signal Level, Noise Level, and SNR in the Whole Sample.** In our population of 331 recording sites, tonal stimuli at the BF led to PSDs with signal and noise levels that covered wide ranges of values. The signal levels varied over about seven orders of magnitude and ranged from \(1.4 \times 10^{-5}\) to \(3.0 \times 10^{-2}\) by \(2.7 \times 10^{2} \text{ mV}^2/\text{kHz}\). The noise levels varied over about four orders of magnitude and ranged from \(3.5 \times 10^{-5}\) to \(5.0 \times 10^{-1}\) by \(5.3 \times 10^{-1}\) \text{ mV}^2/\text{kHz} \) (Fig. 5A). On the other hand, the SNR at the BF had a range of only three orders of magnitude: from 1 to 1,071, with a median of 317. On a logarithmic scale, the range was from 0 to 31.2 \text{ dB}, with a median of 25.4 \text{ dB} (Fig. 5B) and an interquartile range of 5.8 \text{ dB}.

Figure 5B also shows the distribution of the recording sites’ maximal SNR. The highest value was 32.2 \text{ dB} (range, 10.9–32.2 \text{ dB}; median, 26.6 \text{ dB}; interquartile range, 5.6 \text{ dB}). The median of the maximum SNR was slightly larger (1.2 \text{ dB}) than the median of the SNR at BF. Nevertheless, both distributions were similar. We conclude that the neurophonic potential in barn owl NL shows an exceptionally high SNR for tonal stimulation in the high-frequency (>2.5 kHz) range.

**Modeling the neurophonic**

In this section, we address the question of how a neurophonic with an SNR of >30 \text{ dB} and an amplitude in the range of 1 \text{ mV} may be generated. To this end, we set up a computational model of the neurophonic, which was numerically simulated, and quantified in the same way as the experimental data. This comparison, together with a mathematical analysis of the model, allowed us to constrain model parameters.

We will begin with a generic version of the model to show its main features before we turn to a more detailed one. The detailed implementation of the model is described in Methods. For the simplest possible model, we considered only one kind of source. We also assumed that all sources were mutually independent and that they contributed equally, i.e., with the same waveform and amplitude to the neurophonic. Furthermore, we modeled neuronal activity as an inhomogeneous Poisson processes; for the tone-driven neurophonic, we assumed a periodic firing rate (Fig. 6A). The probability that a homogeneous population of \(N\) sources produces a spike in a small time interval was proportional to this time-dependent firing rate. This rate was characterized by the stimulus frequency \(f_{\text{stim}}\), the vector strength \(\nu\) of population phase-locking, and the mean firing rate \(\lambda\) for each of the \(N\) sources. Spike trains generated in this way are shown in Fig. 6, C–E (top left), for three different numbers \(N\) of sources. A convolution of the spikes with some kernel, for example, the spike waveform in Fig. 6B, yielded continuous voltage traces (Fig. 6, C–E, bottom left).

Typical values of model parameters were motivated by physiological data from owl NL and NM. The population-mean vector strength was set to \(\nu = 0.4\) for a stimulus frequency \(f_{\text{stim}} = 5 \text{ kHz}\) (Köppl 1997b), and the population-
mean firing rate was set to $\lambda = 400$ Hz (Carr and Konishi 1990; Peña et al. 1996). We did not consider spontaneous firing in the simplified model. Extracellular spike waveforms of sources within NL were not available. Therefore the spike waveform, or kernel, was approximated by a Gabor function. The parameters of the Gabor were chosen so that its PSD roughly matched the shape of the measured noise spectrum in NL (Fig. 3D). Furthermore, the kernel was required to match generic extracellular spike waveforms (e.g., Gold et al. 2006), and its amplitude was set to a typical value of 100 mV. The spike waveform in Fig. 6B satisfied all criteria. Thus all but one of the degrees of freedom were fixed in the simplest model: the number $N$ of statistically independent sources was the only parameter that was varied in the simulations outlined in Fig. 6.

DEPENDENCE OF THE SIMULATED NEUROPHONIC ON THE NUMBER OF SOURCES. The temporal structure of the simulated neurophonic strongly depended on the number $N$ of sources. When only few sources were used, the singular spike waveforms could be identified ($N = 2$; Fig. 6C), and the neurophonic was not oscillatory at the stimulus frequency $f_{\text{stim}} = 5$ kHz. Even $N = 20$ sources, as in Fig. 6D, were insufficient to achieve the smooth oscillatory temporal course in the voltage trace seen in experiments (Fig. 3A). The simulated neurophonic was similar to owl data only for a large number of sources, for example, $N = 200$ as in Fig. 6E.

How does the number $N$ of independent sources affect the PSD? Clearly, the average level of the PSD increased with $N$ (Fig. 6, C–E, right column). The height of the spectrum’s peak at the stimulus frequency and the SNR also increased with $N$.

For two independent sources, the SNR was 10 dB; for 20 independent sources, the SNR was 20 dB; and for 200 independent sources, the SNR increased to 30 dB, suggesting a linear relationship.

Further numerical simulations in Fig. 7A (gray squares) and a mathematical analysis of the model (gray line; see also METHODS) prove that the SNR depends linearly on the number $N$ of independent sources. Equation 12 tells us that $N = 324$ independent sources were needed to reproduce the highest SNR observed in our recordings (32.2 dB). Further simulations (data not shown) and analytical calculations verified that the SNR also linearly depended on the firing rate $\lambda$ because the sum of $N$ Poisson sources with rate $\lambda$ was equivalent to one source with rate $N \lambda$. Finally, SNR depended quadratically on the vector strength $v$ and was independent of the shape and amplitude of the spike waveform (see also Eq. 11), which supports the importance of the SNR as a useful measure.

Figure 7B shows that the cyclic-mean amplitude depended linearly on $N$ (gray line and squares); see also Eq. 7. The cyclic-mean amplitude, however, also strongly depended on the spike waveform, in contrast to the SNR. With the waveform as shown in Fig. 6B (amplitude, 100 mV) and for 200 sources, the resulting cyclic-mean amplitude was in the range of 1 mV, as observed in experiments.

In this first simple model, we assumed that all sources contributed equally to the neurophonic; that is, all kernels were identical. Further simulations and mathematical analyses (data not shown) indicated that any deviation from this special case of identical kernels only increased the neces-
SNR and cyclic-mean amplitude in the Poisson model of the neurophonic. Gray lines and squares depict results for identical sources; black lines and circles indicate results for spatially evenly distributed sources with distance-dependent kernel amplitudes proportional to 1/r^2. The solid lines show analytical results. The symbols indicate results obtained through numerical simulations. A: the SNR increased as a function of the number of sources; see also Eqs. 11 and 13. To reach an SNR of 30 dB, 200 identical sources are required. For the distance-dependent kernel amplitudes, 633 sources are required. B: the cyclic-mean amplitude increased as a function of N; see also Eqs. 7 and 13. The increase was linear in N for identical sources but proportional to (N^0.3 \sim 1) for distance-dependent kernel amplitudes. Simulation parameters: vector strength, v = 0.4; mean firing rate, \lambda = 400 Hz; duration of the analyzed response, T = 80 ms. Kernels were as in Fig. 6B, either identical for all sources with amplitude 100 \mu V or distance-dependent with amplitude proportional to 1/r^2; in the latter case, the amplitude was 100 \mu V at a distance of 50 \mu m, and the mean distance of neighboring sources was 100 \mu m.

Values of the numbers of sources in biologically motivated scenarios

Thus far we outlined generic features of the model. We now explore two specific scenarios of distributions of sources in greater depth. These scenarios are motivated by the anatomy and neurophysiology of the NL and consider only one type of source at a time.

ANATOMY OF NL. In the NL, three possible types of sources may contribute to the extracellular potential: the nodes of Ranvier of the afferent NM axons, which conduct the phase-locked spikes from NM to NL; the chemical synapses between NM axons and NL neurons; and the NL neuronal spikes. GABAergic inputs were neglected because they do not seem to phase lock to high-frequency tones (Yang et al. 1999).

NL is a large, dorsoventrally flattened oval nucleus in the dorsal brainstem, below the fourth ventricle (Fig. 8). NL is \sim 700 \mu m deep (dorso-ventral), \sim 2 mm wide ( medio-lateral), and \sim 3.5 mm long (rostro-caudal; Fig. 8, H and I). All values are corrected for shrinkage if not stated otherwise. NL has a total volume of \sim 6.4 mm^3 (4.8 mm^3 if not corrected for 10% shrinkage) and is surrounded by a glial envelope (Cheng and Carr 2007), penetrated dorsally by axons from the ipsilateral NM and ventrally by axons from the contralateral NM (Fig. 8, A and J). NM axons interdigitate and traverse the dorso-ventral dimension of NL (Fig. 8, A and B). These axons are densely packed, with a mean axon diameter of 3.3 \pm 1.5 \mu m (n = 406; 3.0 \pm 1.3 \mu m, not corrected for shrinkage; cf. Carr and Boudreau 1993b: 3.18 \pm 0.74 \mu m, not corrected for shrinkage), and a density of \sim 72,000 axons per square millimeter (Fig. 8D). The mean distance between two neighboring nodes of Ranvier is \sim 60 \mu m (Carr and Boudreau 1993b; Carr and Konishi 1990: 58 \pm 5.3 \mu m, not corrected for shrinkage). NL neurons are medium-sized ovals with maximum diameters of 15–20 \mu m (not corrected for shrinkage; Carr and Konishi 1990) and large numbers of short (5.0 \pm 1.9 \mu m) stubby dendrites (Figs. 8, D–F; not corrected for shrinkage. Carr and Boudreau 1993b). NL neurons are sparsely distributed; the mean distance between neighboring NL cells is \sim 100 \mu m (Fig. 8D; Carr and Boudreau 1993b: 75 \pm 8 \mu m not corrected for 10% shrinkage and for the projection of the slice volume to a plane). In total, there are \sim 13,000 neurons in NL (Kubke et al. 2004; Winter and Schwartzkopf 1961). If the volume of NL is divided by the total number of neurons, the density of the neurons in the 5 kHz region is \sim 2,000 mm^-3. This density also corresponds approximately to the mean distance of 100 \mu m. This simple density measure slightly overestimated the cell density in the rostral two thirds of NL because cell density in NL was not even. The highest
cell density was found in the caudal, low BF region. Each NL neuron was contacted by ~100 (range, 45–150) afferent NM axons from each side, and each NL neuron has only one axon (Carr and Boudreau 1993b). The NL spike is probably generated in the first node of Ranvier of the NL axon (Ashida et al. 2007; Carr and Boudreau 1993a; Kuba et al. 2006).

Taking into account these characteristics, in what follows, we use the computational model to estimate how many sources of each type are needed to reproduce properties of the measured neurophonic. Such estimates will allow us to identify the putative generators(s). Here we divide the generators into two classes: 1) independent sources, which is a lower bound for the true number of sources. We note that assumption of identical kernel amplitudes for all neurons. Thus if NL neuronal spikes are the only source of the neurophonic, a significant fraction of the total number of NL neurons to reach an SNR of 30 dB (above we obtained N = 633 for a slightly different set of parameters). Thus if NL neuronal spikes are the only source of the neurophonic, a significant fraction of the total number of ~13,000 neurons in the NL must contribute. The low density (2,000 mm\(^{-3}\)) of NL neurons implies that the electrode picks up signals from a sphere with a radius of ~440 \(\mu m\).

A numerical simulation that illustrates this “NL-output” scenario is presented in Fig. 9. This simulation resembles the sustained response of the data shown in Figs. 1, 3, and 4 as closely as possible. To match the data, we fixed basic parameters (\(\lambda_{\text{driven}} = 240\), \(\lambda_{\text{spont}} = 50\), \(f_{\text{stim}} = 5\), \(v = 0.5\)) and chose N = 206 to reach the observed SNR = 26.2 dB (Eq. 12), which will be presented in the next section.

1) CONTRIBUTION OF THE OUTPUT OF NL TO THE NEUROPHONIC. We first considered how the output of NL, that is the action potentials of NL neurons, contributed to the neurophonic. For monaural acoustic stimulation with tones near BF, as in our recordings, the firing rate of NL neurons was typically about \(\lambda_{\text{driven}} = 210–240\) Hz, and the vector strength was \(v = 0.2–0.5\) for \(f_{\text{stim}} = 5\) kHz (Carr and Konishi 1990; their Table 2; Peña et al. 1996). The spontaneous firing rate was set to \(\lambda_{\text{spont}} = 50\) Hz (Carr and Konishi 1990; Viete et al. 1997).

Using these constraints in our model, Eq. 12 with SNR = 30 dB, \(\lambda = 240\) Hz, \(v = 0.5\), and \(T = 80\) ms predicts \(N = 208\) independent sources, which is a lower bound for the true number of sources. We note that Eq. 12 was based on the assumption of identical kernel amplitudes for all neurons.

Refining the model through a distance-dependent decay of amplitudes according to \(1/r^2\), Eq. 16 indicates that we need \(N = 690\) NL neurons to reach an SNR of 30 dB (above we obtained \(N = 633\) for a slightly different set of parameters). Thus if NL neuronal spikes are the only source of the neurophonic, a significant fraction of the total number of ~13,000 neurons in the NL must contribute. The low density (2,000 mm\(^{-3}\)) of NL neurons implies that the electrode picks up signals from a sphere with a radius of ~440 \(\mu m\).
This number of NL neurons can be found within a radius of 295 μm around the tip of the electrode. Furthermore, a rather large kernel amplitude of 620 μV (at a distance 50 μm) had to be chosen to match the observed cyclic-mean amplitude of 0.32 mV (Eq. 13). Moreover, the large kernel amplitude resulted in a less smooth voltage trace in comparison to the physiological data (Figs. 9A and 1A).

2) CONTRIBUTION OF THE INPUT TO NL TO THE NEUROPHONIC. To estimate how the neuronal structures providing the input to the NL may contribute to the neurophonic, we evaluated the activity of the contribution from NM. The phase-locked input to NL originates exclusively from the NM, that is, via NM axons within the NL and via synapses between terminals of NM axons and somata of NL neurons.

Tones presented monaurally at a frequency of 5 kHz activate NM neurons with BFs near 5 kHz and tonic firing rates of about λ_{driven} = 450 Hz (Carr and Konishi 1990; Peña et al. 1996) and a vector strength of about v = 0.4 (Köppl 1997b). To estimate the number N of independent inputs contributing to the neurophonic for the case of monaural stimulation, we considered that half of the N inputs are spontaneously active because they originate from the nonstimulated side. For the nonstimulated fibers, we took a spontaneous firing rate of λ_{spont} = 200 Hz (Köppl 1997a) and set their vector strength to zero.

To estimate the expected SNR for the case of identical kernels for all sources, the noise level was caused by N/2 driven sources and N/2 spontaneously active sources. Extending Eq. 10, we found that the noise level was proportional to [(N/2) λ_{driven} + (N/2) λ_{spont}]. The N/2 driven sources gave rise to the signal level, which was proportional to [(N/2) λ_{driven} v]^2 T; see also Eq. 9. The SNR is

\[ \text{SNR} = \frac{N \lambda_{driven}^2 v^2 T}{2(\lambda_{driven} + \lambda_{spont})} \]  

To reach an SNR of 30 dB for N = 500 axons can be found in the vicinity of the tip of an electrode within NL. For a density 72,000 mm^{-2} of axons, the required 500 axons fit in a small halo with a radius of ~27 μm. However, different axons might not be statistically independent because an axonal tree originating from a single NM neuron has many branches (Fig. 8, A and B), and each axonal tree represents only one independent source in our model. The minimum halo for N = 500 independent sources was therefore larger. However, even for a 10-fold lower local density of independent axons, the electrode halo will increase only by a factor of 10 to ~210 μm.

A numerical simulation that shows this “NL-input” scenario is presented in Fig. 10. This simulation closely resembles the sustained response of the data shown in Figs. 1, 3, and 4. To match the data, we fixed basic parameters (λ_{driven} = 450 Hz, λ_{spont} = 200 Hz, f_{stim} = 5 kHz, v = 0.4) and chose the total number N of 210 axons to reach the observed SNR = 26.2 dB (Eq. 17). Furthermore, the kernel amplitude 58 μV was used to match the

---

FIG. 9. Numerical simulation of the neurophonic—output scenario. All kernels were identical in shape, but their amplitude depended on the distance with 1/r^2. A: stimulus (top) and response (bottom). The driven response (10–110 ms) and the spontaneous activity (0–10 and 110–200 ms) were sums of the activity of N = 206 sources. The mean distance between sources was 100 μm, and they were distributed uniformly. B: 5 ms intervals of the neurophonic, 25–30 ms after stimulus onset, as labeled in A, at an expanded time scale. The 3 traces were from statistically independent trials. C: mean PSD of the driven response in B, based on activity 10–90 ms after stimulus onset. D: spontaneous activity, 25–30 ms after stimulus offset, as labeled in A. E: mean PSD of the spontaneous activity. The dashed line shows the same level as in C (noise), for reference. Simulation parameters: N = 206 sources, kernel amplitude: 620 μV. Driven activity: vector strength v = 0.5, mean firing rate λ_{driven} = 240 Hz, stimulus frequency f_{stim} = 5 kHz. Spontaneous activity: v = 0, λ_{spont} = 50 Hz. Resulting properties of the driven neurophonic: cyclic-mean amplitude: 0.38 mV, SD: 0.36 mV, signal level (peak): 4.753 mV^2/kHz, noise level: 9 · 10^{-3} mV^2/kHz, SNR: 27.5 dB. Resulting properties of the spontaneous neurophonic: SD: 0.07 mV, noise level around 5 kHz: 1 · 10^{-3} mV^2/kHz.
observed cyclic-mean amplitude of 0.32 mV (Eq. 7 for \(N/2 = 105\) coherent sources).

The example simulation in Fig. 10 is generic in the sense that all model parameters were fixed by experimental constraints—there was no free parameter that needed to be tuned to reproduce the recorded neurophonic. In particular, the SNR and cyclic mean amplitude were reproduced in the numerical simulations by selecting corresponding numbers of sources and kernel amplitudes. Thus, further simulations (data not shown) of the input scenario also matched the data very well. The more important results of the model were the values of the parameters such as number of independent sources and the peak amplitude of the kernel. They could be obtained directly from the mathematical analysis of the computational model in METHODS, in particular Eqs. 7 and 17 for the input scenario (Fig. 11, gray) and Eqs. 13 and 15 for the output scenario (Fig. 11, black). Only for this input scenario were the obtained values of the parameters consistent with the anatomy and physiology of the NL. In summary, our analysis indicates that the contributions of the input to NL can explain the measured properties of the neurophonic.

![Fig. 10. Numerical simulation of the neurophonic—input scenario. All kernels were identical, i.e., each source had the same shape and amplitude. A: stimulus (top) and response (bottom). The driven response (10–110 ms) is a sum of spontaneous activity and monaurally driven activity, each having an equal number of \(N/2 = 105\) sources. The spontaneous activity (0–10 and 110–200 ms) is from \(N = 210\) spontaneously active sources. B–E: see the caption of Fig. 9. Simulation parameters: \(N = 210\), kernel amplitude: 58 \(\mu V\). Driven activity: vector strength \(v = 0.4\), mean firing rate \(\lambda_{\text{mean}} = 450\) Hz, stimulus frequency \(f_{\text{stim}} = 5\) kHz. Spontaneous activity: \(v = 0\), \(\lambda_{\text{spont}} = 200\) Hz. Resulting properties of the driven neurophonic: cyclic-mean amplitude: 0.30 mV, SD: 0.26 mV, signal level (peak): 2.92 mV/\(\text{kHz}\), noise level: \(4 \times 10^{-3}\) mV/\(\text{kHz}\), SNR: 28.5 dB. Resulting properties of the spontaneous neurophonic: SD: 0.10 mV, noise level around 5 kHz: \(2 \times 10^{-3}\) mV/\(\text{kHz}\).](http://jn.physiology.org/)

**DISCUSSION**

A combined computational, anatomical, and electrophysiological study of the neurophonic potential observed in the high-frequency region of the barn owl NL was used to connect the EFPs to their neuronal generators. We shall first compare our results with earlier observations of the neurophonic, then discuss the relative measure of SNR as an indicator for the number of contributing sources, and finally draw conclusions about the nature of sources.

**Neurophonic in the barn owl and in other animals**

Our data are consistent with those of Sullivan and Konishi (1986), in that the time courses of the frequency-following neurophonic potentials in response to monaural tone bursts were typically oscillatory \(\leq 7\) kHz, with amplitudes in the millivolt range. Such a response behavior is indicative of the volley principle (Wever and Bray 1930), because each single neuron cannot produce an action potential at such a high rate.

A strong neurophonic is also present in chicken NL (Köppl and Carr 2008; Schwarz 1992). The observed increases of the PSD in the vicinity of the cell body layer led to the conclusion...
that the neurophonic potential is created by synaptic currents (Schwarz 1992). In contrast, in the barn owl, when the electrode is in NL, there are no local peaks (within a range of 100 μm) in the amplitude, PSD, or SNR of the neurophonic (H.W., unpublished observations). Instead, the neurophonic increases steadily from the edges of NL to the center. Because the structure of NL and the shape of neurons in NL in chick and owl are different, it is not clear whether chick and owl neurophonic in NL share the same sources. In the chick, NL neurons are asymmetric (bipolar), and they are arranged in a densely packed layer of 1–3 neurons (Smith and Rubel 1979). The extracellular neurophonic has an amplitude of about 100 μV or ~10–20 times smaller than in owl, and BFs are lower (Carr and Konishi 1990; Köppl and Carr 2008; Schwarz 1992). According to the model presented here, the extracellular neurophonic in chick may be explained by signals from NM synapses, with additional contributions from NL action potentials.

In cat auditory nerve, the neurophonic may represent a spatial summation of the coherent, phase-locked activity (Snyder and Schreiner 1984). In cat medial superior olive, with a structure similar to the chick NL in that both have a monolayer-like array coincidence detector neurons, the neurophonic is strongest close to this cell layer (Bojanowski et al. 1989; Guinan et al. 1972; Wernick and Starr 1968). There, the neurophonic has been proposed to be generated by a dipole field originating from the bipolar neurons with oriented dendrites, which generate a phase shift of ~0.5 cycles between the sink and the source of the dipole (McLaughlin et al. 2010). To our understanding, the high-frequency neurophonic in the barn owl is not produced by such a dipole field because the phase shift observed when moving the electrode through NL is generally much greater than 0.5 cycles (Sullivan and Konishi 1986).

Derivation of SNR and SNR in other neuronal systems

The SNR, in general, can be used to quantify how much information the response carries about the stimulus (Borst and Theunissen 1999). Here we used the SNR to quantify the measured neurophonic and to compare it to the computational model. For tones, the SNR is a relative measure in a narrow frequency band. To derive the SNR, the phase locking allowed us to average the responses to tone bursts not only across trials (for review, see Borst and Theunissen 1999) but also across the many cycles, typically hundreds. The resulting SNRs in the owl’s NL reached extraordinary high values of ~32.2 dB (range, 10.9–32.2 dB; median, 26.6 dB); the lowest values were probably from locations near the border of NL. Maximum values around 32 dB for extracellular signals were much larger than the estimates for intracellular signals (Ashida et al. 2007). We calculated 27.6 dB (Eq. 11 for monaural input: \( N = 100, \lambda = 450 \text{ Hz}, \nu = 0.4, T = 0.08 \text{ s} \)) which is near our median SNR.

For comparison, SNRs have also been measured in other neural systems. For example, in the saccular nerve of a teleost fish, the SNR reached 30 dB for an 85 ms time window (Tomchik and Lu 2006). For crayfish photoreceptors, SNRs of ~22 dB for 120 s recording time were reported (Bahar et al. 2002). These SNRs obtained for periodically driven responses are much higher than SNRs obtained from intrinsically oscillating systems (Hurtado et al. 2004), although the two cases are difficult to compare because, for the latter, energy is distributed across a wider frequency band.

Modeling results and implications for the origin of the neurophonic potential

We set up a novel model of the neurophonic potential in the auditory brain stem to show its sources. Other, similar approaches have modeled intracellular voltages (Ashida et al. 2007; Gerstner et al. 1996; Kempter et al. 1996, 1998a,b; Kuhlmann et al. 2002). The assumptions underlying our model were straightforward: an inhomogeneous Poisson process with a periodically modulated rate and a kernel that matched typical waveforms as recorded in other neuronal systems. In general, the amplitude of a spike depends on the distance \( r \) of the source to the tip of the electrode. We considered two special cases. First, a distance-dependent decay where the amplitude was proportional to \( 1/r^2 \) (Logothetis et al. 2007; Rall 1962), which is reasonable for neuronal dipoles whose spatial extent is smaller than \( r \). This “far field” case was used to simulate the contribution to the neurophonic because spikes of NL neurons, which have a diameter of 15–20 μm. For distances \( r \) comparable to the size of the neuronal source, the decay may be weaker (Nunez and Srinivasan 2006). In a second case, we assumed that kernel amplitudes were constant. This “near field” case was used to simulate the neurophonic caused by NM axonal arbor, which are elongated structures with a length in the range of 1 mm within NL.

An important model parameter was the peak amplitude of a spike, which was not available for NL of owls because single units could not be isolated from the EFP recorded with conventional metal or glass electrodes. Spontaneous activity in slices of very young owl NL (Lautemann et al. 2008) indicated peak amplitudes of ~100 μV. A kernel amplitude 100 μV may also be an upper bound for amplitudes of extracellular spikes of NL neurons when measured at a reference distance of 50 μm, which is one half the distance between nearest neighbors (Carr and Boudreau 1993b). The value 100 μV is typical for the much larger hippocampal pyramidal neurons, which exhibit action potentials with intracellular spike amplitude in the range of 100 mV (Gold et al. 2006; Henze et al. 2000). In contrast, the intracellular spike amplitude of adult NL neurons has been hypothesized to be only ~10 mV (Ashida et al. 2007; Funabiki and Konishi 2005; see also Golding et al. 1995; Kuba et al. 2005; Oertel et al. 2000; Scott et al. 2005). Spike amplitudes may also be estimated as follows: at the peak of a spike, membrane currents are in the range of 2 nA (Ashida et al. 2007; see also Scott et al. 2010), which corresponds to an input resistance of 5 MΩ for spike amplitudes of 10 mV (K. Funabiki, personal communication). For a spike generated in the first node of Ranvier of the NL afferent axon, ~60 μm away from the soma (Carr and Boudreau 1993a), and for a specific resistance of ~2 Ω m of NL tissue, we can estimate the spike amplitude resulting from such a dipole (Logothetis et al. 2007; p.815) to be at most 2 μA at a distance \( r = 50 \text{ μm} \) from the center of the dipole. In summary, the estimated extracellular kernel amplitude of 19 μV for NL spikes is well below 620 μV, which is the amplitude required to explain the measured cyclic-mean amplitude of the neurophonic with NL spikes only (Fig. 9).

Identical kernel amplitudes. In the preceding approaches, with identical kernel amplitudes and with distance-dependent, dipole-like decay of kernel amplitudes, the mathematical analysis (results in Eqs. 11, 15, and 17) proved that the SNR of the modeled neurophonic was independent of the form and the amplitude of the extracellular action potentials (i.e., the kernels). The
only unknown parameter that remained was \( N \), which is the number of independent sources that contribute to a given SNR.

The model with identical kernels for all sources sets a lower bound for the number \( N \) of sources to reach a certain SNR. Any kind of variability, for example, in the amplitudes or shapes of the kernels, only increases the \( N \) needed to create a certain SNR. Moreover, weakening the assumption of independent sources increases the necessary number of sources. Including in the model further noise sources (e.g., an invariant background noise or GABAergic synapses) also increases \( N \). We found that at least hundreds of independent neuronal sources are needed to reach an SNR of 30 dB. Equations 11 and 17 show how \( N \) depends on the experimentally accessible quantities.

A natural upper bound for the number of independent sources in NL may be the number of NM neurons, because each axonal tree of an NM neuron represents one independent source in NL. The required number of independent axonal branches (to generate the measured properties of the neurophonic) can already be found within a small radius of \( \sim 150 \mu m \) around the electrode. The attenuation of kernels may be neglected for \( r \) that are smaller than the spatial extent of an action potential propagating along an NM axon. Its spatial width in NL is \( \sim 1 \text{ mm} \) for a conduction velocity of 5 m/s and for an assumed temporal width as low as 0.2 ms (Carr and Konishi 1990; their Fig. 5). Therefore within the estimated electrode halo, the distance dependence of the kernel amplitude is weak and may be weaker than \( 1/r \) (Nunez and Srinivasan 2006). However, amplitudes of axonal action potentials are quadrupolar at large distances, i.e., they show a \( 1/r^3 \) dependence of the amplitude for large \( r \) (Nunez and Srinivasan 2006). It is clear that such a strong decay at large \( r \) is necessary to avoid arbitrary large SNRs and cyclic mean amplitudes in models of the neurophonic. We indeed included a strong decay in our model of the input scenario (with identical kernel amplitudes) simply by restricting contributions of sources to a finite halo around the electrode: inside the halo the amplitudes were constant, outside they were zero. A halo in the range 150 \( \mu m \) also guarantees that, for each axon, there are at least two nodes of Ranvier, which are the source of axonal currents. Finally, for cyclic-mean amplitudes in the millivolt range, kernel amplitudes in the range of 100 \( \mu V \) were sufficient (Fig. 11). Numerical simulations based on such identical kernels matched the data well (Fig. 10).

Each of the \( \sim 13,000 \) NL neurons is contacted by \( \sim 100 \) afferent NM axons from each side (Carr and Boudreau 1993b). The contributions of the estimated 2,600,000 synapses within NL can simply be included in the kernel that describes the waveform caused by a spike of an NM neuron because the synapses’ activity is highly correlated to the activity of the axons feeding them. Thus including synapses does not change the above arguments.

Although currents originating from synapses between NM axons and NL neurons might be larger than currents from nodes of Ranvier of NM axons, it is unclear how synaptic currents affect the extracellular field. Synaptic currents may contribute little to the extracellular field because neurons in the high-frequency region of NL are nearly spherical and because synapses are distributed uniformly on the neuron’s surface (Carr and Boudreau 1993b). This “closed field” configuration (Lorente de Nô 1947) predicts that synaptic currents in NL neurons might have a relatively small impact on the extracellular field, in contrast to the elongated NM axons. The “closed field” configuration is broken only by the outgoing NL axon, which is notably large and oriented orthogonal to the NM axons (Carr and Boudreau 1993a).

We propose that contributions from the input to NL, i.e., axons from NM and synapses between NM axons and NL somata, yield the necessary number of inputs to explain both the high SNR and large amplitude of the neurophonic. In this case, contributions to the neurophonic originate within \( \sim 150 \mu m \) of the recording electrode, which is in line with results by Katzner et al. (2009) who showed that 95% of the local field potential in visual cortex originate within 250 \( \mu m \) of the recording electrode. Furthermore, Nelson and Pouget (2010) argued that the electrode impedance and geometry do not appreciably affect recordings of the field potential. That the neurophonic reflects the input was already suggested from models for the development of temporal feature maps in NL (Kempter et al. 2001; Leibold et al. 2001, 2002).

DISTANCE-DEPENDENT DECAY OF KERNEL AMPLITUDES. In addition to the above case of identical kernels for all sources, we considered a dipole-like decay of amplitudes, where sources were spatially uniformly distributed within a spherical halo of radius \( R \) around the electrode tip. Although the kernel amplitudes decreased with distance, the PSD of the signal increased with the radius \( R \) of the halo because the number of sources in a thin spherical shell of radius \( r \) is proportional to \( r^2 \), which compensates the \( 1/r^2 \) dependence of the kernel amplitude. The PSD “signal” level is thus determined by the size of a region in which neurons are synchronously active. On the other hand, the PSD of the noise saturates with increasing \( R \). The noise is therefore “local,” in contrast to the “global” signal. In effect, the SNR in Eqs. 14 and 15 also increases with increasing \( R \) and \( N \), respectively.

The \( 1/r^2 \) dependence of kernel amplitudes gave rise to two predictions. First, for a measured SNR of 30 dB, at least \( N = 690 \) NL neurons are required, which can be found within a sphere of radius \( N^{1/3} R_S = 440 \mu m \) (the minimum halo of an electrode). Second, extracellular spike amplitudes should be in the same range as the cyclic-mean amplitude, i.e., in the millivolt range. Such large extracellular spike amplitudes are difficult to reconcile with the properties of NL neurons and difficulties in isolation of single units in NL.

The hypothesis that NL neurons within a sphere of radius 440 \( \mu m \) contribute to a tone-driven neurophonic is inconsistent with what we know about the tuning and the synchrony of the activity of NL neurons. The width of the short axis of NL parallel to the axonal delay lines is only \( \sim 700 \mu m \) (Carr and Konishi 1988). Moreover, spikes of neurons along this dorso-ventral axis are coherent but not necessarily synchronous, because the preferred phase of firing depends on the neuron’s position, which reflects their tuning to the interaural time difference (Carr and Konishi 1990). For example, for an axonal conduction velocity of 5 m/s and a stimulus frequency of 5 kHz, neurons that are 500 \( \mu m \) apart from each other (along this axis) fire 180° out of phase (Sullivan and Konishi 1986). Their summed contribution to the neurophonic would not increase the cyclic-mean amplitude and the SNR because of negative interference. Positive interference is possible only within a distance of \( \sim 250 \mu m \) from the tip of an electrode. Inside a sphere with radius 250 \( \mu m \), there are, however, only about \( (250/50)^{3/2} = 125 \) NL neurons. Such a small number of NL neurons cannot explain an SNR of 30 dB.

What is the range of synchrony in the two remaining dimensions of NL? The tonotopic axis in NL runs from caudomedial to...
rostrolateral (owl, Carr and Konishi 1990; chicken, Köppl and Carr 2008; Ruel and Parks 1975). We assume that a frequency lamina in the 5 kHz range would extend only ~600 μm in the caudomedial to rostrolateral dimension, given the width at half height of tuning curves of ~1.2 kHz (Péna et al. 2001) and the frequency gradient in this region of ~0.5 mm/kHz (Carr and Konishi 1990, their Fig. 3A). Tuning curves and SNR tuning curves (e.g., Fig. 4E) had a similar width, which confirms the estimated spatial range of synchrony. In our estimate of the range of synchrony, we neglected any change or variability in phase of estimated spatial range of synchrony. In our estimate of the range of synchrony is smaller than the 880 μm upper bound for the range of synchrony is smaller than the 880 μm upper bound for the range of synchrony is smaller than the 880 μm upper bound for the range of synchrony is smaller than the 880 μm upper bound for the range of synchrony is smaller than the 880 μm upper bound for the range of synchrony is smaller than the 880 μm upper bound for the range of synchrony is smaller than the 880 μm upper bound.

We conclude that it is difficult to explain both the high SNR and the large cyclic-mean amplitude of the tone-driven neurophonic with the contributions from NL action potentials only. NL axons also should not contribute significantly because their number is approximately a factor 100 less than the NM axons. The initial segment of the NL axon is, however, very large (≤10 μm in diameter when it leaves the cell body) and may make a significant contribution, although it should be noted that it is strongly myelinated with its first node at 60 μm distance from the soma and that it is orthogonal to the NM axons (Carr and Boudreau 1993a).

To clarify the origin of the neurophonic in future experiments, drugs could be applied into NL to block synaptic responses. For example, pressure injection of GABA in NL in vivo reduced responses (Takahashi and Konishi 2002). However, analogue waveforms were not analyzed in this study. To avoid possible artifacts caused by pressure injection, drugs should ideally be applied iontophoretically. In addition to GABA, both AMPA and N-methyl-D-aspartate (NMDA) receptor antagonists could be applied to separate contributions of axons from contributions of the NL cell bodies.

In summary, the high cyclic-mean amplitude and the large SNR of the neurophonic in response to high-frequency tones provide strong constraints for the origin of the neurophonic potential in barn owl NL. We can exclude large contributions from the spikes of NL neurons. Our analyses and simulations show that a large number of sources, i.e., hundreds, are needed to form the neurophonic potential as observed in barn owl NL. Many arguments support the input from NM as the origin, although the contributions from NM axons and their synapses cannot be separated. Further experimental and theoretical work is needed to resolve this issue.

Acknowledgments

This work profited from the advice of S. Brill, N. Chenkov, J. Donoso, Y. Fukuda, J. Jaramillo, N. Lautemann, R. Schaette and R. Schmidt. We gratefully acknowledge the technical assistance of T. Mauget (University Maryland Electron Microscopy center), S. Shah, and K. Yan.

Grants

This work was supported by the Bundesministerium fuer Bildung und Forschung (Bernstein Collaboration in Computational Neuroscience: Temporal Precision O1GQ07101 to H. Wagner and O1GQ07102 to R. Kempter), the Deutsche Forschungsgemeinschaft (WA 606/12-1; Emmy Noether KE 7881/4 Sonderforschungsbereich (SFB) 618 “Theoretical Biology”), and National Institute of Deafness and Other Communication Disorders Grants DC-00436 to C. E. Carr and IH P30 DC-0466 to the University of Maryland Center for the Evolutionary Biology of Hearing.

Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

References


