Spectral Determination of Responses to Species-Specific Calls in the Dorsal Nucleus of the Lateral Lemniscus

ERIC E. BAUER, ACHIM KLUG, AND GEORGE D. POLLAK

Section of Neurobiology, University of Texas, Austin, Texas 78712

Received 10 April 2002; accepted in final form 11 July 2002

Bauer, Eric E., Achim Klug, and George D. Pollak. Spectral determination of responses to species-specific calls in the dorsal nucleus of the lateral lemniscus. J Neurophysiol 88: 1955–1967, 2002; 10.1152/jn.00261.2002. This study evaluated how neurons in the dorsal nucleus of the lateral lemniscus (DNLL) in Mexican free-tailed bats respond to both tone bursts and species-specific calls. Up to 20 calls were presented to each neuron, of which 18 were social communication and 2 were echolocation calls. We also measured excitatory response regions (ERRs): the range of tone burst frequencies that evoked discharges at a fixed intensity. Neurons were unselctive for one or another call in that each neuron responded to any call so long as the call had energy that encroached on its ERR. Additionally, responses were evoked by the same set of calls, and with similar spike counts, when they were presented normally or reversed. By convolving activity in the ERRs with the spectrogram of each call, we showed that responses to tones accurately predicted discharge patterns evoked by species-specific calls. DNLL cells are remarkably homogeneous in that neurons having similar BF responses to each of the species-specific calls with similar response profiles. The homogeneity was further illustrated by the ability to accurately predict the response profiles of a particular DNLL cell to species-specific calls from the ERR of another similarly tuned DNLL cell. Thus DNLL neurons tuned to the same or similar frequencies responded to species-specific calls with latencies and temporal discharge patterns that were so similar as to be virtually interchangeable. What this suggests is that DNLL responses evoked by complex sounds can be largely explained by a simple summation of the excitation in each neuron’s ERR. Finally, superimposing the spectrograms of each call on the responses evoked by that call revealed that the DNLL population response re-creates both the spectral and the temporal features of each signal.

INTRODUCTION

The dorsal nucleus of the lateral lemniscus (DNLL) is an auditory nucleus located just below the inferior colliculus (IC) and is characterized by two prominent features. The first is that it is a purely GABAergic nucleus that provides strong inhibitory innervation to both the ipsi- and contralateral ICs (Adams and Mugnaini 1984; Burger and Pollak 2001; Covey 1993; Faingold et al. 1993; Li and Kelly 1992; Oliver and Shneiderman 1989; Ross et al. 1988; Shneiderman and Oliver 1989; Shneiderman et al. 1988; Winer et al. 1995; Zhang et al. 1998). The second is that its neuronal population is dominated by binaural cells, where its high-frequency neurons are excited by stimulation of the contralateral ear and inhibited by stimulation of the ipsilateral ear (Brugge et al. 1970; Covey 1993; Kelly et al. 1998; Markovitz and Pollak 1994; Yang and Pollak 1994a,b). These neurons are sensitive to interaural intensity disparities (IIDs), the cues animals use to localize high-frequency sounds (Erulkar 1972).

Because of the strong inhibitory projections the DNLL sends to the IC, it has been the subject of numerous anatomical, biochemical, and neurophysiological studies. The focus of those studies, especially the neurophysiological studies, has been on the processing of interaural disparities and how the processing of those disparities impacts the responses of binaural neurons in the opposite IC (Burger and Pollak 2001; Faingold et al. 1993; Ito et al. 1996; Kelly and Kidd 2000; Kelly et al. 1996; Kidd and Kelly 1996; Shneiderman et al. 1988, 1999; Yang and Pollak 1994b). Only a few studies, however, evaluated what types of temporal discharge patterns are evoked in DNLL neurons in response to monaural stimulation of the contralateral ear (Aitkin et al. 1970; Kelly et al. 1998; Markovitz and Pollak 1993; Yang and Pollak 1994a, 1997). Because the DNLL sends inhibitory projections to the IC on the same side, the DNLL must have a pronounced influence on the processing in the IC and thus on the way that IC neurons respond to both simple and complex signals. It follows, therefore, that an understanding of how the DNLL impacts the IC depends on knowing what form of information the DNLL presents to the IC.

We address this question in this report by showing how DNLL neurons in Mexican free-tailed bats respond to both tones and a wide range of both species-specific calls and other complex signals. We chose to evaluate these features in Mexican free-tailed bats because they are highly social animals that employ a rich repertoire of spectrally and temporally complex communication calls for a variety of social interactions, including mother-infant interactions, courting, agonistic encounters, and territoriality (Balcombe and McCracken 1992; French and Lollar 2000; Gelfand and McCracken 1986; McCracken 1984). Here we show that DNLL neurons process and respond to these complex signals in a simple fashion. In the discussion, we compare DNLL responses to species-specific calls with the responses of IC neurons to the same species-specific calls as described in the companion report. We suggest that the differences in processing between the two nuclei provide insights into some of principal transformations that occur between the IC and the lower nuclei whose efferent projections converge on the IC.

Address reprint requests to: G. D. Pollak (E-mail: gpollak@mail.utexas.edu).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
METHODS

Surgical procedures

Fifteen Mexican free-tailed bats (Tadarida brasiliensis mexicana) were used in this study. Surgical procedures were performed under general anesthesia (isoflurane, IsoFlo, Abbott Labs) as described in the companion paper (Klug et al. 2002). Recordings were begun after the bats had fully recovered from the anesthetic, typically about 1 h. Bats were offered water ad libitum, and lidocaine cream was applied to wounds several times during the experiment. The bats typically lay quietly during the remainder of the experiments, although if they showed signs of discomfort, doses of the neuroleptic ketamine hydrochloride (Vetamine, Mallinckrodt Veterinary, 1/40 dilution, 0.01 ml injection) were administered.

We recorded from each bat in one or two experimental sessions, each of which lasted from 8 to 12 h. At the end of a session, the electrode position was marked with a dye (fluorescent dextran). The bat was either sutured for use in the next day’s experiment or killed for histological processing. If kept for a second experimental session on the next day, the bat was removed from the stereotaxic setup, keeping the front head bar attached to the skull to allow accurate repositioning the next day. The bat was removed from the cushion and re-anesthetized using IsoFlo. Silicone grease was reapplied to the hole, and antibiotic cream was applied to the skull. The reflected skin was then sutured over the exposed skull. Lidocaine cream was then applied to all wounds. The bat was returned to its cage and allowed to recover.

If killed, the bat was removed from the stereotaxic apparatus and from the cushion. It was then overdosed with the inhaled anesthetic IsoFlo and perfused with 4% formaldehyde. After removal from the skull, the brain was placed in a 30% sucrose solution for 24–48 h for the purpose of cryoprotection. The brain was then sectioned at 50-μm thickness on a freezing microtome, and the brain slices were counterstained with cresyl violet. Electrode positions within the DNLL were then confirmed through visualization of the fluorescent marker. All experimental procedures were in accordance with the protocol approved by the University of Texas Institutional Animal Care Committee.

Recordings

Action potentials were isolated using glass micropipettes filled with buffered 1 M NaCl and 2% Fast Green (8–20 mΩ). The electrode was advanced through the brain by a Burleigh microdrive controlled remotely from outside the recording chamber. The signal from the electrode was sent to a preamplifier and then to a Dagan AC amplifier (model 2400). The output of the Dagan amplifier was sent through a second band-pass filter (300–3,000 Hz) and then to a window discriminator, which was fed to a custom made real-time clock. The signals from the clock were fed to a computer (Macintosh G3) for real-time analysis, storage, and later retrieval.

Selection and preparation of complex sounds

A library of 20 natural species-specific calls, as well as their time-reversed versions, was created and stored for later playback. Eighteen social communication calls (SC1–18) and 2 echolocation calls (EC9–10) were selected. Examples are shown in the top panel of Fig. 1. The social communication calls were selected from a sound database of thousands of individual bat calls. The social communication calls selected were representative examples of the different call types. Each call was digitized off of analog tape at an equivalent sampling rate of 200 kHz and converted to AIFF audio format on an Apple computer with SoundEdit software (Macromedia). A sampling rate of 200 kHz allowed encoding of sound frequencies of ≤100 kHz, which is above the reported hearing range of Mexican free-tailed bats.

The species-specific calls were then high-pass filtered (2 kHz and above) and adjusted so that they were all at the same peak intensity that corresponded to the peak intensity of tone bursts at 70 dB SPL. Different intensities were obtained by attenuating those signals, usually in steps of 10 dB. Computer-generated variants, namely time-reversed versions of the sounds, were created in SoundEdit.

Other complex sounds used were naturally occurring while others completely synthetic. A set of 10 sounds recorded in a human environment consisted of two examples of human speech, two examples of human infant vocalizations, three examples of music, two examples of telephone ringers, and a toy squeak. Examples are shown in Fig. 1. These sounds were recorded at sonic frequencies, and then time and frequency were shifted by a factor of 8 to bring the frequencies into the bat’s ultrasonic hearing range. Another set of complex sounds consisted of five different animal vocalizations (Fig. 1). Four of the sounds were from four different species of birds, while the fifth was a primate call. These sounds, acquired from the Cornell Bioacoustics Laboratory through purchase of their software package, Canary, were also shifted up by a factor of 8. The final set of sounds were completely synthetic (Synthesized Sounds) and were designed to mimic the structure of the social communication calls (Fig. 1). The sounds consisted of either pure tones or sinusoidal FM sweeps of durations similar to the maximum duration of the communication calls. The sounds contained several frequency components, either harmonically related or of random frequencies. There were either 5 or 10 components. The energy distribution of the components either followed the pattern of the social communication calls, with the most energy in the fundamental and each subsequent harmonic having 10 dB less energy, or the energy of each component was randomly assigned.

Stimulus playback and data acquisition

Stimuli were presented monaurally to the contralateral (excitatory) ear. Binaural tones were used only to determine the cell’s binaural characteristics. All stimuli were presented using custom-built software and hardware. The tone bursts were generated by a Macintosh computer with custom-built software. The complex sounds were created and stored ahead of time as AIFF sound files which the program could access. All stimuli were then uploaded from the Macintosh G3 into a custom made Downloadable Arbitrary Waveform Generator through a 24-bit digital interface (National Instruments DIO-24) and a digital distributor just prior to that particular sound’s presentation.

All stimuli were presented with a delay of 10 ms relative to the start of data acquisition by using a custom-made pulse delay and at a rate of four per second, controlled by a real-time clock. The stimuli were presented, after the addition of a 200-V DC bias, through calibrated custom earphones (Schuller 1997) whose output was flat, within ±5 dB, from about 8 to 70 kHz. At the start of each experiment, the earphones were inserted into the ear canals and situated a few millimeters from the tympanic membrane. The flexible pinnae were folded onto the housing of the microphones and wrapped with Scotch tape. The acoustic isolation between each ear in this arrangement is 40 dB or greater (Wenstrup et al. 1986).

Tone bursts of 20-ms duration were used to search for neurons. After a cell was isolated, its best frequency (BF), or frequency to which it was most sensitive, and threshold at BF were obtained, followed by rate-level functions, binaural properties, and temporal discharge patterns to tones. We next obtained the cell’s excitatory response region (ERR), or range of frequencies that evoked discharges at a fixed intensity. ERRs were generated with 2.0-ms tones and were obtained at two to three intensities for each neuron. The frequency steps were always 1.0 kHz and the intensity steps were always 10 dB.

The reason for choosing 2-ms tones involves the mathematics of convolution. To perform an accurate convolution with a spectrogram matrix, ideally one would want to use an impulse response. To use a...
longer response could blur the temporal information of the resulting convolution. For reasons of practicality, a 2.0-ms tone was the shortest stimulus available in this study and proved sufficiently short to provide sharp convolution results.

After acquiring ERRs, the species-specific calls were presented. Each call was played 20 times and at multiple intensities. Not all calls were played to each cell because of the variable amount of time that single cells could be held in an awake animal. For all cells, at least 10 of the species-specific calls were played. For most, reversed versions of those 10 calls were also played. For more than half the cells, an additional 10 calls and their reverses were also played. Playback of the first 10 calls was then repeated, for reasons described in the following text.

As this study was intended to investigate how the DNLL encodes complex sounds in general, and not just species-specific sounds, several sets of nonspecies-specific sounds were then played. Nearly all cells (26 of 32) were tested with the set of complex sounds recorded from the human environment. Most cells (19 of 32) were tested with the animal vocalizations and about half with the synthesized sounds.

**Convolution**

After generating the excitatory response regions, predictions of how the cell should respond to the complex sounds were created by convolving the ERR with the spectrograms of the sounds. Convolution is a mathematical process in which two matrices are slid past each other, and the amount of overlap between the two is measured at each time point by multiplication. One matrix represented the neuron’s ERR and was derived from the tone-evoked responses presented at a fixed intensity, while the other matrix was a spectrotemporal representation of one call, where we computed a separate matrix for each call. The convolution not only returns the total magnitude of the overlap but also an envelope of the predicted PST histogram, i.e., the envelope of the predicted firing pattern to that call (how the envelope is generated is described in the following text). Because the response envelope generated for each call had a latency, overall shape, and magnitude, we refer to these features as the neuron’s predicted “response profile” for that call. Envelopes of PST histograms (predicted response profiles)
were calculated for each of the calls and were normalized to the maximum peak response for any one of the 10 calls.

The matrices of the sounds and the matrix of the ERR had the same parameters. The frequency resolution was 780 Hz, and the temporal resolution was 2.56 ms. The time and frequency resolutions were set by Canary. Because ERRs were generated in 1-kHz increments, we interpolated the data obtained at 1-kHz increments into the 0.78-kHz matrix, thereby generating the matrix constructed from the ERR. Only data from an ERR at one intensity were used at any one time. In other words, for comparison with complex sound-response data obtained at an intensity of 20 dB above BF threshold, the ERR matrix was constructed from tone evoked responses obtained 20 dB above BF threshold. For comparison with responses to calls obtained at a higher intensity (30–40 dB above BF threshold), the ERR matrix was constructed from tone responses evoked at the same peak intensity.

The result of a convolution was a prediction of how that cell should have responded to that sound if the response to the complex sound were a linear summation of the activity evoked by pure tones. The accuracy of the prediction was determined by comparing the predicted response profile to the actual response profile evoked by that sound. The comparisons were based on the correlation coefficient for each predicted and obtained response. To create the envelope of the evoked response profile, the values for each bin of the PST histograms were retrieved from the database and exported to MS Excel with the same 2.56-ms time resolution as used for the predicted responses. The obtained responses were normalized to the highest spike count evoked by any of the 10 calls, in the same way as we did for the predictions. Also, values that were smaller than 10% of the maximum peak value were set to zero to eliminate noise. The result was the envelope of the PST histogram (with a binwidth of 2.56 ms) of the response actually evoked by each of the 10 calls, in the same way as we did for the predictions. Also, values that were smaller than 10% of the maximum peak value were set to zero to eliminate noise. The result was the envelope of the PST histogram (with a binwidth of 2.56 ms) of the response actually evoked by each of the 10 calls, in the same way as we did for the predictions.

Any predictive model of how a cell should respond cannot account for random variability and thus has a ceiling of correct predictions limited by the actual repeatability/randomness of responses of the cell. In other words, if the prediction is found to match the actual response with a correlation of 0.5, but the cell’s own responses were variable, then a correlation of 0.5 is as good as the model could possibly be.

A “ceiling correction” on the correlations between the evoked and predicted responses was therefore performed. The ceiling was calculated as the correlation between the two separate sets of evoked responses to 10 SCs. The two data collections were separated in time by about 10 min during which time other data sets using other sounds were collected. The response profiles of the two data sets were correlated following the same procedure used in the correlation of evoked and predicted responses detailed in the preceding text. The average correlation for the 10 SCs for one cell was then considered the ceiling of predictability for that cell. The correlation values for the 32 cells had a median of 0.85. The correlations between the evoked and predicted responses of a cell were then normalized to the ceiling value of that cell. In this way, it was possible to determine how much of the nonrandom response patterns of the cells were accounted for by the tuning curve model that was used to generate the predicted responses.

RESULTS

Overview

This study describes responses of 32 neurons recorded from the DNLL of Mexican free-tailed bats in response to tone bursts and species-specific calls. All cells were binaural and were excited by stimulation of the contralateral ear and inhibited by ipsilateral stimulation. BFs ranged from 11.8 to 37.5 kHz, with most cells between 20 and 30 kHz. The hearing range of this bat extends from 5 to 80 kHz, although frequencies from 15 to 30 kHz are overrepresented in its auditory system (Vater and Siefer 1995). Responses to tones were simple and homogeneous in all DNLL cells. The tuning curves were V shaped. The Q10 dB (bandwidth at 10 dB divided by the BF) values ranged from 2.7 to 25.1, although the Q10 dB of 81% (26/32) of the neurons were between 7 and 12. The average was 9.1. All cells responded monotonically to increases in sound intensity for both BF tones and tones that were off BF (Fig. 2).

Responses evoked by species-specific calls

Here we describe the responses evoked by 10–20 species-specific calls emitted by Mexican free-tailed bats. Most were...
SCs and two were ECs. The most prominent feature was that almost all neurons were unselective for one or another call in that each neuron responded to any call so long as the call had energy that encroached on its ERR. Additionally, the strengths of the responses were, in most cases, similar to the degree of overlap between the spectral energy of the sound and the cell’s ERR. Thus the simple response properties evoked by tones were reflected in the responses evoked by the more complex species-specific calls we presented.

These features are illustrated by the neuron in Fig. 3, which shows responses to 20 species-specific calls when the peak level of each signal was 40 dB above BF threshold. Overlaid on the spectra of each call is a gray band indicating the cell’s ERR at 40 dB above BF threshold and a line indicating the cell’s BF threshold. Notice that the neuron fired to 16 calls and failed to respond to 4 calls. Fifteen of the calls that evoked responses had suprathreshold energy within the cell’s ERR (indicated by the circle regions in each panel). One call, SC14, had energy at ERR frequencies that were either at or just below threshold and evoked a weak response. The four calls to which the neuron did not respond, calls SC2, SC3, SC16, and SC18, had spectral components that corresponded to the cell’s ERR but the energy in those components was below threshold in each of the four calls.

These features were reflected in the population. Figure 4 shows the percentage of calls that evoked responses in the 32 neurons when presented at 10–20 dB above BF threshold and at 30–40 dB above BF threshold. At the lower intensity, the cells, on average, responded to 51% of the calls, and only six cells (20%) responded to 80% of the calls. When the intensity was increased, they responded, on average, to 78% of the calls, and the majority (69%) responded to ≥80% of the calls. The explanation is simply that the spectral energy of some calls was below threshold at 10–20 dB, and thus those calls failed to
excite some of the neurons. Increasing the overall level of the signals raised the appropriate spectral components of each call, and thus many more calls evoked a response at the higher intensity.

The primacy of the excitatory tuning for evoking responses to the complex signals is also illustrated by the responses evoked when the calls were played backward. The idea is that if a cell’s responses were determined by something in addition to the signal’s encroachment on its ERR, such as some temporal structure of the forward call (e.g., a downward sweep through many frequencies), then when the sound was reversed, the cell should respond less vigorously or not at all to that signal because the temporal structure would be reversed. Thus comparing responses to the forward and reverse calls should reveal the degree to which the cell’s responses can be explained simply by the signal’s entry into the cell’s ERR.

Figure 5 is a representative example showing that responses were evoked by the same set of calls when they were presented normally or reversed. Moreover, the neuron also responded with similar spike counts to the forward and reverse versions of each call. This trend was seen for all cells. Figure 6 plots the number of spikes that each of the 32 cells fired to the forward and reverse versions. All signals were presented at 40 dB above threshold.

Response profiles to communication calls are accurately predicted by responses to tones

In the preceding examples, we showed that responses evoked by tones could accurately predict whether or not each neuron would respond to each of the species-specific calls we presented. Here we show that activity in the neuron’s ERR not only predicts whether or not a response would be evoked by each call, but it also predicts the neuron’s response profile, defined as the latency and temporal discharge pattern evoked by each call.

We show this by convolving the spectrogram of each signal with the neuron’s ERR. A convolution, as explained in METHOD...
ODS, is a mathematical process by which two matrices are slid past each other and the amount of overlap between the two is quantified. One matrix was constructed from the spectrograph of each signal, and the other matrix was constructed from the responses in the ERR. The convolution returns a prediction of the response profile that should be evoked by that call, given only the excitatory frequency tuning as a determining factor. The envelopes of the predicted profiles and the profiles actually evoked by each call were plotted, and the similarity between the envelopes was measured by correlation.

The predictions of whether or not each call would evoke a response, termed the binary predictions, were highly accurate. The median predictability for the population was 0.90, showing that the ERR was able to predict whether or not a cell would respond to a call 90% of the time. The binary predictions were just as good for the reverse versions of the calls (median of 0.90) and were not statistically different from those of the forward versions ($P = 0.33$, paired $t$-test).

The convolutions also accurately predicted the response profiles evoked by each call. Figure 7 shows the predicted response profiles and the profiles actually evoked by 20 species-specific calls in one neuron. The only “error” was for call SC11, for which no response was predicted although a robust response was evoked. For the other 19 calls, however, the predicted and obtained profiles were very similar to each other. The correlations of the predicted and evoked responses for the 19 calls were high and ranged from 0.54 to 0.99. The average correlation of the 20 calls for this cell was 0.79, which was representative of the population whose median was 0.72. In addition, predictions were equally good for the reverse versions of the calls. When combined with the results for forward calls, the median correlation of response profiles for all species-specific calls and their reverse versions was 0.72 (Fig. 8A).

While the preceding correlations were high, differences in the predicted and evoked responses were due, in part, to the random variations in responses to signals that occur from trial to trial. We partially correct for random variability by presenting 10 of the species-specific calls 20 times, twice. The responses to the 10 sounds from each of the data collection sets were then correlated with each other following the same procedures used for the correlation of predicted and evoked responses described in the preceding text. This correlation was termed the “ceiling,” and all correlations between evoked and predicted response profiles were normalized to each cell’s ceiling. After making the ceiling correction, the average correlations between predicted and evoked responses increased. The median for the correlations of the response profiles of the species-specific calls was 0.85 (Fig. 8B), a value substantially higher than the previous correlations.
higher than the median correlation of 0.72 that was computed without the ceiling correction.

Responses to species-specific calls are homogeneous among DNLL cells

One of the most striking features of DNLL cells is that they are remarkably homogeneous in that neurons having similar BFs responded to each of the species-specific calls with similar response profiles. This homogeneity is illustrated in Fig. 9, which shows responses evoked by five different calls in five DNLL cells with similar BFs and in one neuron with a different BF. The BFs of the five closely tuned cells ranged from 29.2 to 28.6 kHz. It is apparent from just a cursory inspection that the responses of each cell to a particular call were similar. The homogeneous responses of similarly tuned cells is in marked contrast to the responses evoked by the same calls for the DNLL cell tuned to 20.3 kHz, shown in Fig. 9, bottom.

The homogeneity among similarly tuned DNLL cells is further underscored by the ability to predict the response profiles of a particular DNLL cell from the ERR of another similarly tuned DNLL cell. Stated differently, because similarly tuned cells are homogeneous, it should not matter which cell’s ERR generated the prediction and which generated the evoked response. This was in fact what was found. Figure 10 shows the predicted and evoked responses of six DNLL neurons for three calls, SC4, SC5, and SC8. Five neurons, 1–5, had similar BFs that ranged from 29.2 to 28.6 kHz. The BF of neuron 6 was 20.3 kHz and was different from the other neurons. The predicted responses for all six neurons are the same and are the predicted responses derived from convolving the ERR of neuron 4 (enclosed by the rectangle) with calls, SC4, SC5, and SC8. We now asked how well the predicted responses for neuron 4 correlated with the obtained responses of neurons 1–5. For neuron 4, the correlations among the evoked responses and those predicted from its own ERR were 0.53 for call SC4, 0.73 for call SC5, and 0.86 for SC8. What is noteworthy is that the correlations of the predicted responses for neuron 4 and the evoked responses for the other neurons were as good, and often better, than they were for neuron 4. For example, the correlations of responses predicted by the ERR of neuron 4 and the evoked responses of neuron 2 were 0.76 for call SC4, 0.93 for call SC5, and 0.89 for call SC8. The average correlation between the predicted and evoked responses of the five closely tuned cells was 0.74 without the ceiling correction. By way of comparison, the average correlation between the responses predicted by the ERR of neuron

![Figure 8](image_url)  
**Fig. 8.**  
A: graph showing the average correlation between evoked and predicted responses to all species-specific calls, including their temporal reverses, that were presented to each neuron.  
B: corrected correlations between evoked and predicted responses to forward and reversed versions of each call. The correlation coefficients were corrected for random variations in responses, i.e., each were corrected for a ceiling factor.

![Figure 9](image_url)  
**Fig. 9.** Neurons with similar BFs respond to species-specific calls with similar responses. Responses of 6 neurons evoked by 5 species-specific calls are shown. The BF of each neuron is shown on the far left of each response set. Five neurons had similar BFs ranging from 29.2 to 28.6 kHz, and 1 neuron had a BF of 20.3 kHz (bottom). All signals were presented between 30 and 40 dB above each cell’s threshold.
and those evoked in the one differently tuned cell, cell 6, was only 0.03. This provides further evidence that DNLL cells within an isofrequency contour are a homogenous population in which the parameters that determine the responses of one cell to a complex sound are the same parameters that determine the responses to the same sound for all cells in that contour.

**Prediction of responses to other complex sounds**

In the previous sections, we showed that responses to tone bursts in the neuron’s ERR accurately predicted responses to species-specific calls. One interpretation of these results is that the neural circuitry of the free-tailed bats is designed to process their own calls. We believe this is not the case but rather that the simplicity of the DNLL endows its neurons to respond to any complex signal in a manner that is predictable by its responses to tone bursts. To firmly establish this feature, we also evaluated the ability of tonal responses to predict the responses evoked by a variety of other complex sounds. These sounds are not normally heard by this species and thus eliminated the factors of familiarity and behavioral relevance.

Three sets of sounds that were not species-specific were used. One set was sounds from a human environment (environmental sounds) ranging from the human voice to the ringing of a telephone. Another set consisted of vocalizations of other animals (animal vocalizations). Both the environmental sounds and animal vocalizations were increased in frequency and decreased in time by a factor of 8 to move their component frequencies into the ultrasonic hearing range of the bats. The final set of sounds consisted of completely synthetic sounds (synthesized sounds) designed to mimic some of the harmonic and temporal structures of the species-specific sounds.

By convolving these sounds with each neuron’s ERR, the
binary predictions for all sound groups were as good as those of the species-specific sounds, with average binary predictability for all sound groups between 0.77 and 0.89. It appeared that the ERR could account for which sounds a cell would respond to regardless of the nature of the sound.

The ceiling-corrected correlations among the predicted and evoked response profiles of the complex sounds that were not species specific were also high, although not always as high as those for the species-specific calls. The correlations of the animal vocalizations were as high as those for the species-specific calls. The median for the animal vocalizations was 0.86 and was not statistically different from the 0.85 median value for the species-specific calls ($P = 0.52$). The median correlation for the synthetic sounds was 0.70 and for environmental sounds was 0.55. Although these were still fairly high correlations, they were significantly lower than the correlations for the species-specific calls and calls of other animals.

These numbers indicated that by simply taking into consideration responses evoked by tonal frequencies in the neuron’s excitatory response region in a linear model, one can account for 73% of the nonrandom features of the temporal response profiles of the cell to species-specific sounds and between 74 and 30% of the response profiles of any complex sound. It appeared that the nature of the sound determined, to some extent, the accuracy with which the frequency response area could predict the response profile of the sound, though most sounds still had a median correlation of over 0.50 indicating that the tuning curve was still a moderately good predictor of response profiles.

**DISCUSSION**

The main finding of this study is that DNLL responses evoked by complex sounds can be largely explained by a linear summation of the excitation in each neuron’s response region. Thus responses evoked by almost all the complex sounds we presented were accurately predicted by convolving the spectrotemporal features of each call with the responses evoked by tonal frequencies. Moreover, DNLL neurons tuned to the same or similar frequencies responded to the same complement of calls, and they responded to those calls with latencies and temporal discharge patterns that were so similar as to be virtually interchangeable.

What this suggests is that the DNLL population response to any given call is simply a neural representation of the signal’s spectral-temporal structure. To illustrate this representation, we show in Fig. 11 the responses of all the DNLL neurons in our sample that were driven by calls SC7 and SC8. The responses are plotted on the ordinate according to their BF and are plotted on the abscissa according to their relative latencies and discharge durations evoked by each call. Superimposing the spectrograms of each call on the responses evoked by that call reveals that the DNLL population response re-creates both the spectral and the temporal features of each signal.

**Comparing responses to species-specific calls in the DNLL and IC**

The simplicity and homogeneity of the DNLL responses to species-specific calls is in marked contrast to the selective and diverse responses evoked by species-specific calls in the IC. The disparity is surprising because the DNLL is only one synapse below the IC, and both nuclei receive a large convergence of inputs from a similar complement of lower centers (Grothe et al. 1994; Huffman and Covey 1995; Shneiderman et al. 1988; Yang et al. 1996). The differences are apparent, however, when responses to species-specific calls of DNLL neurons are compared with those of IC neurons as reported in the companion paper by Klug et al. The comparison is especially relevant because both studies were conducted on Mexican free-tailed bats and many of the calls used to evaluate responses in the IC were the same calls presented to DNLL neurons in this study.

We point out four principal differences. First, in contrast to the DNLL, where a particular call evokes similar response
profiles among isofrequency DNLL cells, IC neurons with similar BFs respond to a particular call with a wide variety of response profiles. Second, whereas DNLL cells are nonselective and respond to any call that encroaches on their excitatory response regions, IC cells are highly selective in that they respond only to some calls but not to others even though the signals they fail to respond to have energy that encroaches on their ERRs. In many cases, IC cells fail to respond to any of the calls presented, although other IC cells with the same BF respond to a few and yet others to many of the calls. Third, the predictions of response profiles to complex calls, based on responses to tonal stimuli (convolutions), are much less accurate in the IC than in the DNLL, suggesting that IC processing is more complex. Fourth, while DNLL responses are accounted for almost exclusively by the excitation in their response regions, the responses of IC cells are influenced significantly by inhibition. In the majority of IC neurons, selectivity is greatly reduced when inhibition is blocked, and tonal stimuli become more accurate predictors of response profiles to complex stimuli. Thus IC neurons are diverse and highly selective when inhibitory innervation is active, but when inhibition is blocked, they become more like DNLL cells. The transformation, however, is not complete; unlike DNLL cells, isofrequency IC cells with blocked inhibitory inputs still display a variety of response profiles to a particular call and never display the homogeneity of isofrequency DNLL cells.

Response homogeneity is a feature DNLL cells may share with lower auditory nuclei

As mentioned in the preceding text, one feature that distinguishes the DNLL from the IC is response homogeneity, the similar responses among isofrequency DNLL cells to a given call. Response homogeneity, at least for tones and other relatively simple stimuli, is also a characteristic feature of most auditory nuclei below the IC. Typically, the neurons in each lower nucleus respond to tone bursts with one, or with a few principal discharge patterns, with only minor variations among the population. Some examples are the lateral superior olive (LSO) (Joris and Yin 1995; Park et al. 1997), the medial nucleus of the trapezoid body (MNTB) (Banks and Smith 1992), the medial superior olive (MSO) (Yin and Chan 1990), and the subdivisions of the ventral nuclei of the lateral lemniscus (VNLL) (Covey and Casseday 1991; Haplea et al. 1994; Huffman et al. 1998a,b; Vater et al. 1997), where each nucleus is distinguished by its own set of response properties that are different from the properties expressed by the other nuclei. Indeed, it is the homogeneity of features in each nucleus that provides the justification for cataloging the responses and types of currents in MNTB, LSO, or MSO, or for studying the binaural processing in LSO and contrasting it with the binaural processing in MSO. If we eliminate the variable of frequency tuning and consider only cells tuned to the same frequency, any signal, whether spectrally simple or complex, should evoke a similar discharge pattern in most, if not all, isofrequency neurons in the particular nucleus. This does not mean to say that neurons in every lower nucleus are as simple as DNLL neurons; some lower nuclei, such as the dorsal cochlear nucleus (DCN) (Davis and Young 2000; Nelken et al. 1997) and the columnar division of the VNLL (VNLLc) (Huffman et al. 1998a,b; Zhao and Wu 2001) clearly process information more nonlinearly and thus respond to some signals but not to others. The significant point, however, is that an isofrequency population, whether in the DCN, the LSO, or the VNLLc, should respond in the same way to a given signal or the isofrequency population should fail to respond to a different signal. Contrast this with the IC, where a spectrally complex or a spectrally simple signal evokes a diversity of responses among isofrequency neurons.

The features discussed in the preceding text suggest that one of the principal transformations that occurs in the IC is a change from processing that emphasizes similarity in lower nuclei to one that emphasizes diversity and selectivity in the IC. Two key features could account for these response differences. The first is the difference in the ways that inputs are distributed along isofrequency contours in lower nuclei compared with the IC. The second is the complement of voltage-gated channels in the population. The DNLL, as an example of a lower nucleus, receives innervation from a number of lower nuclei, but most isofrequency DNLL neurons, or neurons in the frequency contours of any lower nucleus, presumably receive the same complement of projections and have a similar complement of voltage-gated channels (Bajo et al. 1999; Fu et al. 1997; Huffman and Covey 1995; Wu and Kelly 1995; Yang et al. 1996). This can explain why a given signal evokes a similar response profile in isofrequency DNLL neurons. The innervation pattern in the IC is markedly different, where inputs from most lower nuclei form bands that innervate restricted regions of each frequency contour (Gabriele et al. 2000; Oliver et al. 1995, 1997; Shneiderman and Henkel 1987; Whitley and Henkel 1984). The band from a particular lower nucleus only partially overlaps the bands from other nuclei. Thus any segment of a frequency contour receives a unique complement of inputs from a subset of lower nuclei that differs from the subset that innervates adjacent regions of the contour. In this way, the excitatory inputs to an IC cell would arise only from a subset of lower nuclei, while the inputs to cells in the same contour a short distance away would be somewhat different and so on across the contour. Such differences in excitatory inputs could account for the IC diversity among isofrequency neurons seen even when inhibition was blocked (Klug et al. 1995; Park and Pollak 1994; Pollak and Park 1993). The response of each IC neuron also undergoes substantial modification by inhibition. The inhibitory innervation, which modifies response profiles and creates selectivity, also derives from a subset of lower nuclei, where the different subsets distribute differently along a frequency contour. By mixing and matching excitatory and inhibitory inputs, a large number of potential innervation combinations can be created that can be enlarged even more by combining the inputs on cells with different complements of voltage- and/or ligand-gated channels (Moore et al. 1998; Nelson and Erulkar 1963; Peruzzi et al. 2000; Sivaramakrishnan and Oliver 2001; Wagner 1994). In this way, neurons in each frequency contour could respond to complex signals with far more diversity than occurs in frequency contours of the lower nuclei that innervate the IC.

The strategy of the ascending auditory pathway is that each lower nucleus transforms the incoming spike trains in a unique way, and spike trains that emerge from each lower nucleus then provide either excitatory or inhibitory innervation to the IC. The various outputs that emerge from each nucleus are then distributed regionally along frequency contours, where the
information from the lower nuclei is integrated and expressed as diverse and complex responses. Exactly what information each lower nucleus presents to the IC in response to species-specific calls is largely unknown, although obtaining such information is critical if a deeper understanding of how and why the IC responds to such signals is to be achieved. This study represents a first step in providing that information. Here we showed that one lower nucleus, the DPLL, presents GABAergic innervation to the IC that is an accurate neural representation of the spectrotemporal features of the sound received at the ears.

We thank C. Resler for technical support. We also thank F. Theunissen for critical advice and assistance.

This work was supported by National Institute on Deafness and Other Communication Disorders Grant RO1 DC-00268-16. Present addresses: E. E. Bauer, Virginia Merrill Bloedel Hearing Research Center, University of Washington, Seattle, WA 98195-7923; and A. Klug, Oregon Hearing Research Center, Oregon Health Sciences University, Portland, Oregon 97201.

REFERENCES


