Title: Two distinct representations of social vocalizations in the basolateral amygdala

Running Head: Two representations of social vocalizations in the amygdala

Authors: Marie A. Gadziola¹,², Sharad J. Shanbhag¹, and Jeffrey J. Wenstrup¹,²

Authors’ addresses: ¹Department of Anatomy and Neurobiology, Northeast Ohio Medical University, Rootstown, Ohio 44272. ²School of Biomedical Sciences, Kent State University, Kent, Ohio 44240.

Author Contributions:
Conceived/designed study: MAG, SJS, JJW
Performed experiments: MAG
Analyzed data: MAG, SJS, JJW
Wrote manuscript: MAG, SJS, JJW

Corresponding author:
Jeffrey Wenstrup, Ph.D.
Department of Anatomy and Neurobiology
Northeast Ohio Medical University
4209 State Route 44, PO Box 95
Rootstown, Ohio 44272
Telephone: (330) 325-6630
Fax: (330) 325-5916
Email: jjw@neomed.edu

Copyright © 2015 by the American Physiological Society.
ABSTRACT

Acoustic communication signals carry information related to the types of social interactions by means of their “acoustic context”, the sequencing and temporal emission pattern of vocalizations. Here we describe responses to natural vocal sequences in adult big brown bats (*Eptesicus fuscus*). We first assessed how vocal sequences modify the internal affective state of a listener (via heart rate). The heart rate of listening bats was differentially modulated by vocal sequences, showing significantly greater elevation in response to moderately aggressive sequences than appeasement or neutral sequences. Next, we characterized single-neuron responses in the basolateral amygdala of awake, restrained bats to isolated syllables and vocal sequences. Two populations of neurons distinguished by background firing rates also differed in acoustic stimulus selectivity. Low background neurons (<1 spike/s) were highly selective, responding on average to one tested stimulus. These may participate in a sparse code of vocal stimuli, in which each neuron responds to one or a few stimuli, and the population responds to the range of vocalizations across behavioral contexts. Neurons with higher background rates (≥1 spike/s) responded broadly to tested stimuli and better represented the timing of syllables within sequences. We found that spike timing information improved the ability of these neurons to discriminate among vocal sequences and among the behavioral contexts associated with sequences, compared to a rate code alone. These findings demonstrate that the BLA contains multiple robust representations of vocal stimuli that can provide the basis for emotional/physiological responses to these stimuli.

Keywords: Acoustic Communication, Bat, Electrocardiogram, Eptesicus fuscus, Vocalizations
INTRODUCTION

Social communication by sound depends on acoustic signals composed of basic spectro-temporal elements—syllables—that are combined in temporal patterns. Both the types of vocal syllables emitted and the temporal emission patterns, or sequencing, carry significant information related to the behavioral state of the sender and can influence the behavior of conspecifics: bats (Gould and Weinstein, 1971; Russ et al., 2005; Bastian and Schmidt, 2008; Janssen and Schmidt, 2009; Gadziola et al., 2012a), primates (Fichtel and Hammerschmidt, 2002; Fischer et al., 2002; Rendall, 2003), tree shrews (Schehka et al., 2007), ground squirrels (Leger and Owings, 1977; Mateo, 2010), dogs (Yin and McCowan, 2004; Pongracz et al., 2006), African elephants (Soltis et al., 2009), and guinea pigs (Arvola, 1974). An accurate interpretation of these social signals thus depends on appropriate analysis of the “acoustic context”, defined here as the sequencing and temporal emission pattern of vocalizations within which they are presented. This study examines how a brain region critical for the response to social signals, the amygdala, responds to the natural sequencing of syllables used in vocal communication.

Based on its central roles in establishing associations between sensory cues and biologically important events and in orchestrating emotional responses (LeDoux, 2000, 2007), the amygdala likely serves an important function in acoustic communication (Morris et al., 1999; Sander and Scheich, 2001; Johnstone et al., 2006; Naumann and Kanwal, 2011; Gadziola et al., 2012b; Parsana et al., 2012; Peterson and Wenstrup, 2012; Grimsley et al., 2013). Auditory information from the medial geniculate body and auditory association cortex enters the amygdala via the lateral nucleus (LeDoux et al., 1991; Mascagni et al., 1993). With widespread projections to hypothalamic, midbrain, and medullary nuclei, the central nucleus of the amygdala is well-positioned to coordinate behavioral, autonomic and neuroendocrine responses to biologically
significant stimuli (Cardinal et al., 2002; Price, 2003; Sah et al., 2003). Amygdalar responses also modulate subsequent auditory processing through direct projections to auditory cortex (Amaral and Price, 1984), nucleus basalis (Sah et al., 2003), and, at least in bats, the inferior colliculus (Marsh et al., 2002). This circuitry establishes the basis for emotional responses to social vocalizations, but how amygdalar neurons respond to these signals and encode the emotional content within them remain open questions.

Several studies utilizing different species suggest different answers to these questions. Although a common finding is that neurons generally display stronger excitatory discharge in response to vocalizations associated with negative affect (Naumann and Kanwal, 2011; Gadziola et al., 2012b; Parsana et al., 2012; Peterson and Wenstrup, 2012), some studies report excitatory responses to stimuli associated with both positive and negative affect (Gadziola et al., 2012b; Peterson and Wenstrup, 2012). These and another study (Grimsley et al., 2013) describe a key role for the temporal patterning of spike discharge to convey information about acoustic stimuli and the associated behavioral context. A limitation of most of these studies, however, is that they have relied on the presentation of social syllables in isolation (Naumann and Kanwal, 2011; Gadziola et al., 2012b; Peterson and Wenstrup, 2012), and show considerable variation in choice of stimulus repetition rates and randomization procedures. Because syllable ordering and temporal emission patterns provide the necessary acoustic cues for distinguishing the affective meaning of vocalizations (Bastian and Schmidt, 2008; Bohn et al., 2008; Gadziola et al., 2012a), such methodological differences in stimulus presentation may alter the meaning of a stimulus to a listener. Amygdalar neurons, which appear to encode the meaning of vocal stimuli (Wang et al., 2008a; Parsana et al., 2012; Grimsley et al., 2013) should be affected by these methodological features.
The current study examines responses to natural, species-specific vocal sequences in adult big brown bats. This species (*Eptesicus fuscus*) is a highly social and long-lived (Paradiso and Greenhall, 1967; Kurta and Baker, 1990; Brunet-Rossini and Austad, 2004) mammal with conspecific relationships that last over multiple years (Willis et al., 2003; Willis and Brigham, 2004). In its social interactions, it relies heavily on acoustic communication and possesses a complex vocal repertoire that mediates mother-infant interactions (Gould and Weinstein, 1971; Gould et al., 1973; Monroy et al., 2011), social foraging behavior (Wright et al., 2014), a range of affiliative and aggressive behaviors during roosting (Gadziola et al., 2012a), and, presumably, mating (no reports). Across these studies, over 20 distinct syllables have been identified. Bats are also one of the few mammalian taxa to display vocal learning (Esser, 1994; Boughman, 1998; Knörnschild et al., 2009), and this may extend to the family that includes the big brown bats (Knörnschild, 2014). For all of these reasons, we believe this species forms a valuable mammalian model of social vocal interactions that may be mediated by the amygdala in humans.

We first assess how vocal sequences modify the internal affective state of a listener (via heart rate), then characterize single-neuron responses in the basolateral amygdala to both isolated syllables and vocal sequences. The neurophysiological studies addressed several questions: 1) Do neuronal responses distinguish among individual vocal sequences, 2) Do neuronal responses distinguish among vocalizations associated with different behavioral contexts, 3) By what response features do neurons distinguish among vocalizations or contexts? Since previous studies suggest that both stimulus timing and spike timing are relevant to the BLA encoding of vocal stimuli, we compared responses to syllables and sequences, examined responses as a function of syllable repetition rate, and assessed the role of spike timing in coding. The results show that vocal sequences differentially modify the internal affective state of a listening bat,
with moderately aggressive vocalizations evoking a greater change in heart rate compared to appeasement or neutral sequences. Further, amygdalar neurons form distinct populations that participate in very different representations of social vocalizations. Finally, the results indicate that the timing of action potentials plays an important role in one of these forms of representation.

MATERIALS AND METHODS

Data were obtained from big brown bats (*Eptesicus fuscus*) maintained in a captive research colony at Northeast Ohio Medical University. Electrocardiogram (ECG) recordings were obtained from four adults (3 males, 1 female), and neurophysiological recordings were obtained from the basolateral division of the amygdala (BLA) in five additional adults (1 male, 4 females). All animal husbandry and experimental procedures were approved by the Institutional Animal Care and Use Committee at Northeast Ohio Medical University (protocol number 12-033).

Surgical Procedures

Animals received an intraperitoneal injection of Torbugesic (5 mg/kg, IP, Fort Dodge Animal Health, Fort Dodge, IA) before surgery, then were anesthetized to effect with Isoflurane (2-4%, Abbott Laboratories, North Chicago, IL). Hair overlying the dorsal surface of the skull was removed using depilatory lotion and the skin was disinfected with iodine solution. A midline incision was made and the underlying muscle was reflected laterally to expose the skull. A metal pin was cemented onto the skull to secure the head to a stereotaxic apparatus. For bats used in neurophysiological studies, a sharpened tungsten ground electrode was inserted through a small hole in the skull and cemented in place. Based on stereotaxic coordinates, a craniotomy (<1 mm
diameter) was made to access the amygdala. The underlying dura was removed and the
craniotomy covered with sterile bone wax. For bats used in ECG monitoring studies, hair
overlying the dorsal flanks on both sides of the back was removed to allow electrodes to contact
the skin. Immediately following surgery, local anesthetic (Lidocaine) and antibiotic cream were
applied to the surgical areas, an injection of Carprofen (3-5 mg/kg, s.c., Butler Schein Animal
Health Inc., Pittsburgh, PA) was administered, and the animal returned to the holding cage for a
two-day recovery period before physiological experiments.

Acoustic Stimuli

Digitized acoustic stimuli, including both synthetic stimuli and pre-recorded
vocalizations, were converted to analog signals at 250-500 kHz and 16-bit depth using SciWorks
(DataWave Technologies). Signals were anti-alias filtered (cutoff frequency of 125 kHz) and
attenuated (Tucker-Davis Technologies (TDT)), then amplified and sent to an electrostatic
speaker (TDT, model ES1). The system response had a gradual roll-off of approximately 3 dB
per 10 kHz. Harmonic distortion components were not detectable ~45 dB below the signal level.
Speaker placement differed in the two experiments and will be described separately.

Big brown bat social vocalizations were previously recorded from captive animals in
defined social interactions and characterized in detail (Gadziola et al., 2012a). The present study
used as stimuli both isolated syllables as well as vocal sequences. A ‘syllable’ is the smallest
acoustic unit of a vocalization, defined as one continuous emission surrounded by background
noise. A ‘vocal sequence’ is a longer emission pattern that can be defined by unique syllable
sequencing and inter- and intra-call boundaries. Syllable types were named according to their
acoustic structure.
These vocalizations were assigned to a behavioral context based on non-vocal, overt behavioral displays (Gadziola et al., 2012a). Lower aggression vocalizations were emitted by roosting bats in conjunction with aggressive behavioral displays (baring teeth, biting) after they had been jostled or disturbed by a conspecific. Bats transitioned to high aggression vocalizations after prolonged or repeated disturbances, producing distinct behavioral displays (snapping, darting, and defensive posturing before physical contact). Appeasement vocalizations were never associated with aggressive displays, but were instead related to increased social contact among conspecifics. Further, the distinction between lower and high aggression calls was supported by differences in the change in heart rate (HR) in response to tactile irritation: high aggression calls occurred after longer periods of tactile irritation and were associated with larger HR changes (duration and magnitude). Both males and females emitted social vocalizations in each behavioral context.

Although our previous study identified 18 distinct social syllable types, we reduced our stimulus set to include 12 syllable types (Fig. 1), choosing the most probable syllables and those with diverse spectro-temporal characteristics. An echolocation syllable type was also included to represent a neutral context. Because there are many more syllables used in aggressive interactions than in appeasement interactions, our syllable set includes more aggression than appeasement syllable types (n = 9 aggression, 3 appeasement, 1 neutral). In choosing natural vocal sequences, we balanced our stimulus design by choosing two or three characteristic exemplars to represent each of the four behavioral contexts: lower aggression, high aggression, appeasement and neutral (Fig. 2). Vocalizations recorded under these behavioral contexts were unique in terms of the syllable type and sequencing of syllables (Fig. 2) (Gadziola et al., 2012a). For example, lower aggression vocal sequences were typically composed of the same, rapidly
repeating DFMs syllable (Fig. 2C), whereas high aggression vocal sequences included a more
diverse range of syllable types emitted at slower repetition rates (Fig. 2D). Appeasement vocal
sequences also contained unique syllable types and sequencing structure (Fig. 2B). All vocal
stimuli selected were recorded under a natural behavioral context and not provoked by the
experimenter. Appeasement vocalizations are typically emitted at a level ~30 dB less intense
than aggression or biosonar vocalizations.

**Electrocardiogram Monitoring**

ECG monitoring procedures were identical to those described in Gadziola et al. (2012a). Briefly, animals were lightly anesthetized by Isoflurane inhalation prior to attaching the HR monitor. The ECG was recorded differentially with two surface electrodes temporarily fixed to the dorsal flanks, allowing full range of movement. Physiological signals were transmitted via a wireless headstage (4.7 g) to a receiver (Triangle Biosystems, Inc.), permitting bats to behave freely within the experimental chamber. On an initial training day, we first tested that the bats tolerated the weight of the headstage and placement of the surface electrodes by allowing them to move freely within the test chamber for one hour. Although the headstage was 15-25% of body weight, the animals moved freely in their cage and would also interact with other bats if given the opportunity. The loudspeaker was placed in the center of the chamber floor to maintain the most consistent sound field for bats hanging from the ceiling or walls of the experimental chamber.

Analog output from the receiver was amplified and bandpass filtered (100-500 Hz; Model 3600, A-M Systems, Inc., Sequim, WA), digitized (10 kHz/channel, 12-bit depth; DataWave Technologies Inc., Loveland, CO), and stored on a computer. DataWave software controlled
physiological and video data acquisition. A threshold was used to detect the positive deflection of each R-wave within the QRS complex, and the corresponding time stamps saved.

Bats were allowed 30 minutes to recover from anesthesia, acclimate to the experimental chamber, and achieve a stable baseline HR. The ECG was recorded over 5-min trials, along with video and vocalizations. The first 60 s of a trial were used to monitor the baseline HR. After 60 s, sounds were presented from a loudspeaker for 30 s, then instantaneous HR was measured for the remainder of the trial. Within a recording session, bats were presented with acoustic sequences assigned to one of five behavioral contexts: lower aggression, high aggression, appeasement, and a neutral control sequence of pure tones at two different intensity levels. The order of sequence presentation was randomized within each session. Vocalizations were presented at sound levels consistent with naturally occurring intensities; appeasement vocalizations were attenuated by 20 dB relative to aggression vocalizations. To directly test level-dependent changes, pure tone sequences were presented at 0 dB and 20 dB attenuation. As a reference, a 25 kHz tone (the average peak frequency across vocal sequences) at 0 dB attenuation corresponds to 97 dB SPL. For calls associated with a behavioral context, stimuli were composed of three exemplars that were randomized with equal weighting to minimize habituation to repeated sounds. Stimuli were presented at a rate of 1/s for 30 trials. Bats were not tested with isolated syllables.

For further analysis of ECG data, DataWave files were imported into MATLAB (MathWorks, Natick, MA). Instantaneous HR was computed as the reciprocal of R-R intervals in beats per minute (bpm). Movement artifact was reduced by removing outliers and applying a sliding window with a shape-preserving piecewise cubic interpolation to fill missing values. Baseline HR was calculated as the instantaneous HR averaged over the first 60 s of each trial.
For each trial, change in HR was calculated by subtracting the baseline from the maximum HR value.

**Neurophysiological Recordings**

General methods used in auditory neurophysiological experiments were described previously (Gadziola et al., 2012b). These experiments were conducted in a single-walled acoustic chamber. On experiment days, animals were placed in a custom-built stereotaxic device used to guide electrode penetrations and aid histological reconstruction of recording sites. Animals were secured with a non-adhesive wrap and placed inside a small tube where they rested comfortably while head-fixed within the stereotaxic apparatus. Bats were not sedated during experiments. If an animal showed signs of discomfort or distress, the experiment was terminated for the day. Recording sessions did not exceed 6 hours and were limited to one per day. Physiological recordings were obtained from both sides of the amygdala. On average, animals were used in 7 recording sessions (range: 4–9 sessions) on different days.

Single-neuron activity was recorded using micropipette electrodes filled with 1 M NaCl (tip diameters of 2–5 µm) and advanced by hydraulic micropositioner (Kopf 560). Extracellular potentials were amplified, bandpass filtered (300–5000 Hz, A-M Systems 3600), and sampled at 40 kHz with 16-bit resolution (DataWave Technologies). DataWave software recorded spike occurrence and displayed peri-stimulus time histograms (PSTHs) and raster plots in real time. Spikes of constant waveform shape and amplitude determined single-neuron activity. Peak voltage of spikes typically exceeded background noise by a factor of 5. Neuronal spiking was identified either through background discharge or in response to acoustic search stimuli (noise bursts and social vocalizations). An explicit feature of the experimental design was an attempt to characterize auditory responses at any site with clear spiking activity (i.e., at least two
overlapping spike waveforms per 5 min recording period). As a result, our sample includes many more auditory unresponsive neurons, highly selective neurons, and neurons with low background discharge (< 1 spike/s) than our previous study of the big brown bat amygdala (Gadziola et al., 2012b).

The loudspeaker was placed 10 cm from the animal and 25° into the sound field contralateral to the amygdala under study. Synthetic acoustic stimuli (noise and tone bursts, 2 ms rise-fall time, 30 ms duration; tonal sequences) and previously recorded social vocalizations from unfamiliar bats (13 syllables, 7 vocal sequences) were presented in blocks of 30 trials at a rate of 1/s. Duration of syllables ranged from 4–59 ms, and vocal sequences ranged from 148–766 ms. Peak amplitude was normalized across all stimuli. Once a neuron was encountered and isolated, initial data collection involved a standard series of rate/level functions in response to social vocalizations (syllables and sequences), broadband noise, and tones. Stimuli were typically presented at three attenuation levels: 0, 20, and 40 dB. Simple stimuli were used to assess onset latency and responsiveness to pure tones or broadband noise.

Since the repetition rate of individual syllables within a vocal sequence may distinguish different vocalizations or behavioral contexts, we examined the responsiveness of neurons to repetition rate. Artificial sequences consisting of repeated presentations of 5 different syllables were constructed. The syllables used were two aggression syllables (DFMs, rBNBs), one appeasement syllable (DFMI-QCFI-UFM) and two neutral stimuli (echolocation pulse, 25-kHz tone burst). Repetition rates were varied (1, 2, 4, 10, 30, 50, 80, 110 Hz) while holding the number of syllable repeats the same (4 repetitions). The same syllable was used for all four repetitions within the sequence. Syllables with short durations (DFMs, echolocation pulse) were presented at the maximum repetition rates while longer duration syllables (e.g., DFMI-QCFI-
UFM, rBNBs) were limited to repetition rates with periods greater than the syllable duration. Each artificial sequence stimulus was presented 30 times.

Verification of recording location

Stereotaxic coordinates developed in our laboratory for this species guided electrode penetrations through the amygdala (Gadziola et al., 2012b). Responses were recorded from an amygdala in several successive penetrations, the last of which was marked by a deposit of a neural tracer (FluoroGold, FluoroChrome, Inc., Englewood, CO or FluoroRuby, Molecular Probes, Eugene, OR). Thus, nearly all electrode penetrations were placed close to electrode tracks marked by deposits. Tracers were iontophoretically deposited using a constant current source; the electrode was kept in position for 10 additional minutes, then removed from the brain. Details of tracer-specific techniques have been described previously (Marsh et al., 2002; Yavuzoglu et al., 2010). Tracer deposit sites were photographed with a SPOT RT3 camera and SPOT Advanced Plus imaging software (version 4.7) mounted on a Zeiss Axio Imager M2 fluorescence microscope.

Spike data analysis

Only neurons that were presented with a complete syllable and sequence stimulus set were included in the analysis. Post-hoc spike sorting using spike features and principal component analysis confirmed well-isolated single-neuron recordings and allowed removal of artifact when necessary. In some cases (n=5 sites), two well-isolated neurons were present at the same recording site. Spike times and stimulus information were then exported to MATLAB for further analysis. A baseline background or “spontaneous” discharge rate was calculated from an initial 30 s non-stimulus control period. As background discharge can vary with stimulus history,
a pre-stimulus background rate was also calculated from the 200 ms prior to stimulus onset for each block of trials featuring the identical stimulus type and level. Mean firing rates across trials were calculated in 10 ms bins, along with the 95% confidence interval. The criterion for a significant stimulus-evoked response was at least one 10 ms bin between stimulus onset and 100 ms post-stimulus offset that did not have overlapping confidence intervals with the pre-stimulus background. For neurons with little or no pre-stimulus background discharge, the firing rate in a 10 ms bin was required to exceed 10 spikes/s. For neurons with high background rates, the response of each neuron was captured in a population response plot by determining if the activity within each 10 ms bin was excitatory (red), inhibitory (blue) or not significantly different (black) by comparing each bin’s confidence interval with the average confidence interval computed across the 200 ms pre-stimulus background.

The ability of amygdalar neurons to synchronize to the individual syllables of repeated stimuli was assessed by computing the vector strength (VS) for each neuron’s response to each stimulus repetition. The Rayleigh statistic was used to assess significance of the VS measure, with p = 0.001 indicating significant synchronization (Mardia and Jupp, 2000).

We examined the correspondence between each neuron’s significant responses to the 13 individual syllables in our stimulus set and to the natural vocal sequences in which those syllables were present. For each neuron, we determined the number of sequences for which a significant response would be expected based upon the syllables that elicited a significant response. This expectation assumed that if a neuron responded to a syllable, it would respond to every sequence containing that syllable. The number of sequences that elicited significant responses was then divided by the expected number and expressed as a percentage. A complementary analysis used the sequences eliciting a significant response to determine a count
of syllables for which a significant response was expected. Again, this expectation assumed that if a neuron responded to a sequence, it would respond to every syllable, presented in isolation, within that sequence. While this is less likely the case, our sequences were composed of relatively few syllable types. The composition of each test sequence is indicated in the legend for Figure 2.

The spike distance metric (SDM), developed by Victor and Purpura (1996, 1997) and implemented in MATLAB by the Spike Train Analysis Toolkit (http://neuroanalysis.org/toolkit/), was used to assess the relative importance of spike timing to the underlying neural code and to determine whether spike discharge patterns can be used to discriminate among individual stimuli or the behavioral contexts associated with these (Victor and Purpura, 1996, 1997). According to the SDM, $D_{spike}[q]$, the distance (or dissimilarity) between two spike trains is the minimum total “cost” to transform one spike train into the other by a series of insertions, deletions, and time shifts of spikes. Both the insertion and deletion of a spike is set to a cost of 1, whereas shifting spikes by an amount of time, $t$, is set to a cost of $qt$. The cost parameter $q$ (in units of 1/s) can be varied to characterize the relative importance of spike times and spike counts. When $q=0$, the cost to move a spike in time is “free”, and distances are computed solely on the number of spikes (i.e. a rate code). As $q$ increases, timing of individual spikes begins to exert a stronger influence on the distance metric, since a change in spike time of $1/q$ will have the same cost as simply deleting the spike.

A confusion matrix was constructed to assess how well responses to the same stimulus cluster together compared to pairs of responses from different stimuli. Each row represents instances of stimulus conditions while each column represents instances of predicted stimulus conditions; correct assignments fall along the diagonal. For each trial of a given stimulus, $D_{spike}$
was calculated between a given spike train and spike trains from other stimuli. The confusion matrix bin that obtained the minimal value was incremented by 1 to indicate the predicted stimulus, or in the case of a tie, was divided equally among the stimuli that shared the minimum $D_{\text{spike}}$ value. The information, $H$ (in units of bits) was then computed to evaluate the ability of the neuron to classify stimuli based on temporal patterns of auditory responses. $H$ will be zero when classification is random, whereas perfect discrimination of sound stimuli corresponds to $H = \log_2(13) = 3.7$ bits for syllables, or $H = \log_2(8) = 3$ bits for vocal sequences.

The above process was repeated for a range of cost values ($q=0, 4, 8, 10, 12, 13, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 1000, 2000$) to characterize the associated temporal integration window at which maximal information is found ($H_{\text{max}}$). The contribution of spike timing to the stimulus classification was assessed by subtracting the information when the cost, $q$, was 0 ($H_{\text{rate}}$) from $H_{\text{max}}$.

To further control for any bias associated with using a finite data set, the chance level of information, $H_{\text{shuffle}}$, was estimated by randomly re-assigning spike trains to stimulus conditions, and computing the information held in the confusion matrix of shuffled values (Victor and Purpura, 1996). This process was repeated 100 times for each neuron to obtain a robust estimate of $H_{\text{shuffle}}$. All reported values of $H$ calculated from the observed responses were corrected by subtracting $H_{\text{shuffle}}$ from the original value. Any negative $H$ values after correction were set to zero.

To determine whether spike discharge patterns transmit information about the behavioral context of vocal stimuli, a second SDM analysis was conducted as described above, but combined all spike trains belonging to the same associated behavioral context instead of by stimulus identity. Pure tone sequences were excluded from this analysis, leaving four behavioral
context categories: biosonar, appeasement, lower aggression, and high aggression. Thus, perfect discrimination of behavioral context in this analysis corresponds to $H = \log_2(4) = 2$ bits.

RESULTS

HR response of listeners to acoustic stimuli

To understand how social vocalizations modify the behavioral and affective state of a listener, we first monitored ECG signals from animals during presentation of acoustic signals in five different categories: conspecific vocalizations associated with lower aggression, high aggression, and appeasement, and neutral controls of pure tone sequences that were presented at each of two intensity levels.

Bats showed an average baseline HR of 460±58 bpm during the 60 s prior to sound stimulus presentation. The HR of listeners showed significant elevations in response to sound on the majority of trials (81%, or 57 trials). However, conspecific vocalizations evoked a greater change in HR compared to the neutral pure tones (142±35 vs. 76±26 bpm, two-sample t-test, $p<0.01$).

There were two major findings from analysis of changes in HR. First, aggression vocalizations evoked the largest change in HR compared to appeasement vocalizations (171±123 vs. 85±72 bpm; one-way ANOVA, $p<0.05$) or neutral tones (171±123 vs. 77±70; one-way ANOVA, $p<0.01$). Second, and more surprising, lower aggression vocalizations evoked a greater change in HR compared to all other behavioral contexts, including high aggression (220±121 vs. 122±107 bpm; one-way ANOVA, $p<0.05$; Fig. 3A, B). An additional finding was that HR responses to neutral tones showed sound-level dependent changes, with larger elevation of HR
for higher-intensity tones than for the lower-intensity tones (106±34 vs. 47±36 bpm; independent t-test, p<0.05; Fig. 3A).

Amygdalar responsiveness is related to background firing rate

Neurophysiological approaches were used to examine how neurons in the basolateral amygdala (BLA) respond to and discriminate among vocal sequences, and to compare differences in responses to isolated syllables and vocal sequences. Sixty-four neurons were histologically localized to the BLA. From these, 43 neurons (67%) were considered auditory-responsive based on our criteria (see Methods).

Although significant periods of inhibition were observed in 19% of neurons, inhibition was generally difficult to detect unless there was a sufficiently high background firing rate. Testing responsiveness with 50 ms time bins, rather than the 10 ms bins normally used, resulted in an additional eight neurons being classified as auditory responsive, and was more likely to detect significant suppression. However, the longer time bins were not useful for the temporal analysis of response patterns and were therefore not employed in subsequent analyses.

Background Firing Rates

We first examined the distribution of mean firing rates of the background activity recorded from an initial 30-s period before any test stimulus presentation. This ensured that background firing rates were not influenced by stimulus history. The average background firing rate was 1.9 spikes/s for all neurons, and 2.8 spikes/s for auditory-responsive neurons. Among the auditory-responsive neurons, background rate departed from a normal distribution (Kolmogorov–Smirnov test, p<0.001) and was bimodal (coefficient of bimodality, b = 0.73; Fig. 4). For subsequent analyses, the population of auditory responsive neurons was split into two groups based on a break in the distribution at a firing rate of 1 spike/s. Roughly half of the
auditory responsive neurons (22 of 43 neurons) showed background firing rates less than 1
spike/s, and are referred to as the “low background” group; the remaining neurons showed
background rates that were equal to or greater than 1 spike/s, and are referred to as the “high
background” group. Among auditory-unresponsive neurons (n=21), the mean background rate
was 0.1 spikes/s, with all such neurons showing background rates less than 1 spike/s. No
differences were found in action potential half-width or recording depth between the low and
high background groups.

To test whether the background activity of BLA neurons was modulated during the
stimulus testing period, mean firing rates calculated during the initial 30 s silent period were
compared to the 200 ms window prior to each stimulus onset. The majority of neurons (72%, or
46 of 64 neurons) showed significant modulation of the background activity to at least one
stimulus. The background rates of most auditory responsive neurons were modulated by acoustic
stimulation (81%, or 35 of 43 neurons): roughly half of the responses showed an increase in
background firing and the other half showed a decrease. A smaller proportion of auditory
unresponsive neurons had significant modulation of background firing during the testing period
(52%, or 11 of 21 neurons). For each of these “auditory unresponsive” neurons, background
firing increased in response to at least one stimulus. To control for this stimulus-induced
modulation of background activity, all measures of stimulus-evoked spiking were computed
relative to the background activity recorded during the 200 ms prior to stimulus onset.

Neuronal Responsiveness to Acoustic Stimuli

We next examined the relationship between background firing and the number of
syllables and sequences to which a neuron responded (Fig. 5A). Neurons with low background
rates were highly selective, responding on average to less than one tested syllable and one tested
sequence (Fig. 5A, black circles). Across the entire suite of stimuli, most of these neurons responded to only one tested stimulus. In contrast, neurons with high background rates responded to most tested stimuli (Fig. 5A, yellow-filled squares). On average, high background neurons responded to 10 of 13 tested syllables and 7 of 8 tested sequences. The low and high background groups had minimal overlap in the number of stimuli to which the neurons responded (Fig. 5A). The distributions of responsiveness to syllables and sequences were both non-normal (Kolmogorov-Smirnov test: syllables, p <= 0.001; sequences p < 0.01) and bimodal (coefficients of bimodality: syllables, b = 0.74; sequences, b = 0.77).

Neurons in the BLA responded to a broad range of acoustic stimuli, including isolated syllables, vocal sequences, and synthetic acoustic stimuli. In general, tested stimuli were similarly effective at evoking a response among the sample of neurons. However, there was a dramatic difference in the proportion of neurons responding to stimuli between the low and high background groups (Fig. 5B). On average, 83±11% of high background neurons responded to each stimulus (Fig. 5B, left), while only 7±5% of low background neurons responded to each stimulus (Fig. 5B, right). Among low background neurons, vocal sequences were more likely than syllables to evoke responses (2.6±1.1 neurons per sequence vs. 1.2±1.4 neurons per syllable; independent t-test, p<0.05). Note that several low background neurons were not responsive to either individual syllables or the natural sequences; such neurons responded only to BBN, tones, or to syllables presented at certain repetition rates.

These differences in responsiveness between high and low background neurons are illustrated in Figures 6 and 7. In Figure 6, a neuron with low background firing (0.1 spikes/s) displayed a significant response to only two of the tested syllables and one of the tested vocal sequences. For this neuron, mean firing rates across a 200 ms window ranged from 0 to 11
spikes/s, but there was a robust response (peak firing rate of 77 spikes/s) only to the appeasement syllable, DFMl-QCFl. The high selectivity of this neuron is illustrated by the fact that closely related syllables (DFMl or DFMl-QCFl-UFM) evoked very little or no significant spiking. To sequences, there was little or no response; in particular, the neuron responded poorly to repetitions of the DFMl-QCFl syllable within the Appeasement 2 sequence. This result suggests that neuron’s firing depended substantially on the acoustic context within which the excitatory syllable was presented.

In stark contrast, the high background neuron (4.5 spikes/s) in Figure 7 was responsive to all tested syllables and vocal sequences presented at the 20 dB attenuation level. Average firing rates calculated over a 200 ms window from stimulus onset ranged from 3 to 41 spikes/s across all stimuli, with peak firing rates within the same window ranging from 17 to 113 spikes/s. To syllables (Fig. 7A), the neuron’s response magnitude and duration varied but were unrelated to syllable duration or behavioral context. The only consistent pattern was that most of the strongest responses were to syllables with an initial downward frequency modulated component (e.g., E, DFMls, DFMl, DFMl-QCFl, DFMl-QCFl-UFM). The pattern of response to vocal sequences was more varied than the response to syllables. For several sequences (e.g., Biosonar, Appeasement 1, and Appeasement 2), the response was well locked to the temporal patterning of the constituent syllables. For other sequence responses (e.g., Lower Aggression 1 and Lower Aggression 2), the relationship to the syllable response was less clear, in part due to the high repetition rate of syllables in those sequences. It is noteworthy that both these sequences (Lower Aggression 1 and Lower Aggression 2) are exclusively composed of the DFMls syllable, yet the temporal response patterns are quite different. This suggests that the acoustic context within which syllables are presented affected the response of this neuron.
Since high background neurons responded to most stimuli yet show considerable variability in their response properties, we explored how social vocalizations were represented across the sampled population of high background neurons within the BLA. The proportion of neurons showing significant excitation was calculated for each 10-ms time bin in response to syllables and sequences in order to estimate the synchronized activity across the sample population (Figs. 8, 9).

In response to all tested syllables, BLA neurons showed similar patterning of excitation regardless of syllable structure or behavioral context (Fig. 8). Isolated syllables typically evoked a single response peak across the sampled BLA neurons. Specifically, Figure 8 shows that nearly all tested syllables, including both aggression and appeasement syllables, evoked a response in nearly 50% of high background neurons during a 10 ms post-stimulus bin. The noteworthy exception was the DFMs syllable, evoking a response among many fewer neurons. There was no clear preference for syllable type or calls associated with a particular behavioral context at any of the three sound levels presented (data not shown).

In response to vocal sequences, the population displayed temporal activation patterns that usually featured multiple response peaks (Fig. 9). Across the population, aggressive vocal sequences tended to evoke a more robust and potentially synchronized population response than did appeasement sequences, and did so across attenuation levels. Further, the population response patterns to the aggressive vocalizations differed substantially. These results suggest that the temporal patterning of discharge among the population may discriminate among vocal sequences. As with single neurons (previous section), the temporal patterns of the population did not always correspond to the response patterns to individual syllables. For example, the population response to single appeasement syllables (Fig. 8) was much stronger than to the
appeasement sequences (Fig. 9). Conversely, the population response to the Lower Aggression sequence was much stronger than the response to isolated presentation of its constituent DFMs syllable.

The single neuron and population responses described in this section suggest that low and high background neurons participate in different amygdalar representations of vocal stimuli. Responses of low background neurons may form a sparse code of vocal stimuli, based on their highly selective, relatively high firing rate response to a particular stimulus. In contrast, the high background neurons appear to participate in a neural code based on broad responsiveness, in which discrimination among vocal stimuli may be expressed by differences in response timing and magnitude. For both groups of neurons, the data suggest that acoustic context influences the manner in which the neuron responds to constituent syllables. We explore this issue in the next section.

**Responses to syllables and sequences suggest a role for acoustic context**

Since most studies of amygdalar coding of social vocalizations have examined the response to syllables, it is important to ask how the response to syllables relates to the response to sequences containing that syllable. If a neuron responds to a syllable in isolation, does it also respond to the sequences that contain that syllable? Conversely, if a neuron responds to a particular sequence, does it respond to the isolated syllable or syllables contained in that sequence? How does the response to syllables change as the repetition rate increase? The responses of the neurons illustrated in Figures 6-9 suggest that the context within which syllables occurs may be important.

We assessed the correspondence between a high background neuron’s response to syllables and its response to the sequences that contain those syllables by dividing the number of
observed sequence responses by the number expected on the basis of syllable responses. This value, expressed as a percentage, quantifies the correspondence between syllable responses and sequence responses, with 100% indicating maximum correspondence. Figure 10A displays this correspondence across all tested neurons that showed at least one syllable response. Across neurons, the median value of the correspondence was 73.3%, indicating an observed response to 73.3% of the sequences to which a neuron was expected to respond. A neuron’s response to a syllable was not a highly reliable predictor of its response to a sequence containing that syllable.

Do sequence responses predict the syllables to which a neuron will respond? We used each high background neuron’s response to sequences to develop an expectation of the number of syllables to which it would respond and then divided the number of observed syllable responses by the number expected on the basis of sequence responses. Figure 10B displays this correspondence across all tested neurons that showed at least one sequence response. Across neurons, the median value of the correspondence was 22.5%, indicating an observed response to 22.5% of the syllables expected on the basis of sequence responses. A neuron’s response to sequences was a poor predictor of its response to syllables contained within the sequence. These observations suggest that the acoustic context within which syllables are presented may alter the neuron’s response.

The rate at which syllables are repeated is a specific feature of acoustic context that carries behavioral contextual information in big brown bats (Gadziola et al., 2012a) and other species (Fischer et al., 2002; Schehka et al., 2007; Bastian and Schmidt, 2008). Figures 6-9 suggest that low and high background amygdalar neurons may differ in their ability to follow the rate of syllable repetition in natural sequences. To explore this further, we examined responses of 25 neurons to tones and syllables repeated at rates from 1 to 110 per second, rates that occur
naturally among vocal sequences used by big brown bats (Gadziola et al., 2012a). Figure 11 displays the results of this analysis. For low background neurons \((n = 13)\), very few responded to any of these stimuli, an expected result of the high selectivity of these neurons for acoustic stimuli. Furthermore, there was virtually no evidence that these neurons could follow natural rates of syllable repetition: in only one neuron’s response was there significant locking to a repeated sequence at rates above 2 per second. For high background neurons \((n = 12)\), most neurons responded to most presented rates of syllable presentation. However, synchronization to repeated syllables occurred in a minority of neurons and rarely occurred at tested rates higher than 10 per second. These results indicate differences between the two groups of neurons in the ability to respond to and follow syllable repetitions, but in general showed that BLA neurons do not follow the high rates of stimulus repetition present in many vocal sequences, such as the low aggression sequences shown in Figure 9.

**Neural coding strategies**

To investigate coding strategies by which BLA neurons discriminate among social vocalizations, we analyzed spike discharge using Victor and Purpura’s (1996, 1997) spike distance metric (SDM). The SDM characterizes the dissimilarity between spike trains. For our analysis, the SDM was used examine the differences in spike trains elicited by vocal stimuli and then quantify the relative stimulus discrimination contributions of spike rate and spike timing across a range of temporal precision values. Analyses were performed using an 800-ms post-stimulus onset spike train time window to encompass all stimulus durations. Preliminary analyses showed that this longer analysis window was more informative than shorter analysis windows for both sequences and isolated syllables (data not shown).
The SDM analysis revealed that the amount of information carried in a pure rate code was similar for vocal sequences and isolated syllables (Fig. 12A, top). However, additional information was gained in encoding vocal sequences when spike timing was considered, with best performance in this neuronal sample achieved using an integration window of 40 ms (Fig. 12A, top). In contrast, discrimination performance was not improved by considering spike timing in response to isolated syllables. To evaluate the information added by temporal discharge patterns, the difference between the maximum amount of information for each neuron across the different temporal integration windows, $H_{\text{max}}$, was compared to that using a rate code alone, $H_{\text{rate}}$ (Fig. 12A, bottom). The median information difference across all neurons was greater than zero for both syllables and sequences, indicating a role for spike timing in coding strategies in the BLA. However, spike timing information had a greater contribution to discrimination among sequences than among syllables (Wilcoxon rank sum test: $p<0.001$). These analyses suggest greater variation in the temporal patterns of responses to vocal sequences, compared to isolated syllables. This analysis is consistent with the spiking patterns shown by high background neurons in Figures 7-9.

Can neurons in the BLA discriminate among vocal stimuli based on their associated behavioral context? To address this question we ran a second SDM analysis that collapsed vocal stimuli into 4 categories: biosonar, appeasement, lower aggression, and high aggression (Fig. 12B). Responses to the pure tone sequence were not included. This analysis revealed that information was available to discriminate the behavioral context associated with vocal stimuli. Information was higher in response to vocal sequences when spike timing was considered, with best performance achieved with an integration window of 40 ms (Fig. 12B, top). The median information difference between $H_{\text{max}}$ and $H_{\text{rate}}$ across all neurons was greater than zero for both
syllables and sequences, and spike timing information had a greater contribution to
discrimination among behavioral contexts in response to sequences than syllables (Wilcoxon
rank sum test: p<0.001) (Fig. 12B, bottom). Consistent with the spiking patterns observed in high
background neurons in Figure 7-9, these analyses suggest that the temporal patterning of
responses to vocal sequences can be used to discriminate the behavioral context of vocal stimuli.

DISCUSSION

Since a fundamental role of the amygdala is to orchestrate emotional responses to sensory
stimuli, we (Gadziola et al., 2012b; Peterson and Wenstrup, 2012; Grimsley et al., 2013) and
others (Naumann and Kanwal, 2011; Parsana et al., 2012) have hypothesized that amygdalar
neurons should respond to and encode vocal stimuli that are associated with different behavioral
states or social contexts. To test this hypothesis, we examined responses of BLA neurons to
natural vocal sequences and their constituent syllables. The vocal sequences used here evoke
increases in HR that are scaled to the associated behavioral context, with greater increases in HR
associated with aggressive calls than to appeasement calls or neutral sequences. Among BLA
neurons, we report a new finding that auditory responsiveness was bimodally distributed and that
the degree of selectivity to vocal stimuli is closely associated with background discharge rate.
These data indicate that distinct populations of BLA neurons participate in different
representations of vocal stimuli. Low background neurons form a sparse code in which each
neuron responds to one or very few vocal stimuli, and different neurons respond to different
vocalizations. High background neurons respond to most stimuli. A new finding among BLA
neurons is that the temporal patterning of action potentials carries more information than a rate
code alone, but this applies only to encoding of vocal sequences, not syllables. The temporal
patterning also conveys information about the behavioral context of vocal sequences. These findings demonstrate that the BLA contains multiple robust representations of vocal stimuli that can provide the basis for emotional/physiological responses to these stimuli by an animal.

Several analyses of syllable and sequence responses support the conclusion that the acoustic context within which syllables are presented, i.e., their sequencing and rate, has a major impact on the response of BLA neurons. From a methodological view, this raises questions about the utility of studies, ours and others, that examined BLA responses to isolated syllables. From a functional view of the amygdala, this reinforces the concept that the sequencing and rate of acoustic stimuli can have a major impact on the amygdala’s response to a sound, and by extension, on the animal’s emotional response to these stimuli.

**Emotional content in vocalizations**

Our previous work shows that the HR of vocalizing bats provides a measure of internal state that is well matched to investigator classifications of social context based on behavior and vocalizations (Gadziola et al., 2012a). In the current study, we find that the HR of a listener is differentially modulated by vocalizations from different behavioral contexts, with aggressive vocalizations evoking greater elevations than appeasement or neutral sequences. Each of these findings is consistent with the general understanding that vocalizations represent the internal state of a vocalizing animal and that they may alter the internal state of the listening animal. While not novel, our current result is important in demonstrating a link between vocalizations and emotional state in our animal model.

The novel feature of these results is that we observed indication of a mismatch between the states of the sender and the receiver. Thus, we found that lower aggression sequences evoke the greatest increase in HR, even when compared to high aggression sequences. Multiple lines of
evidence (Gadziola et al., 2012a) support our original classification of the relationship between vocalizations and the sender’s internal state. What explains the mismatch?

One explanation is partly methodological. We employed three randomly presented sequences for each social context, and it is possible that the chosen high aggression sequences do not all have the same impact on the emotional state of a listener. In fact, our neurophysiological results reveal differences in amygdalar responses between the two tested high aggression sequences. Future experiments will test whether these sequences elicit distinct heart rate/behavioral responses. This may reveal subtler categories of high aggressive vocal interactions than our initial studies revealed.

A second explanation may relate to a real mismatch between a sender’s “intent” and a listener’s interpretation. In the present case, a potential explanation for the mismatch relates to the proximity of the listener. When a vocalizing bat produces high aggression calls, the listening bat is not typically in close proximity, resulting in a decreased probability of biting behavior relative to lower aggression contexts (Gadziola et al., 2012a). Thus, one might predict that high aggression vocalizations may actually be less emotionally arousing to a listening bat that is out of harm’s way.

**Subpopulations of amygdalar neurons**

Our experimental approach included an attempt to record from neurons of all background rates and the use of an extended pre-stimulus recording period to assess this rate. As a result, we recorded many more neurons with low background discharge compared to our previous study (Gadziola et al., 2012b). With few exceptions, these low background neurons were either unresponsive to our subset of auditory stimuli or highly selective for acoustic stimuli. These response types are much more common in the current study than in our previous work. This
The association between background discharge and stimulus selectivity has been described in studies of other brain regions (Kerlin et al., 2010; Ison et al., 2011; Mruczek and Sheinberg, 2012). In these studies, there was a clear linkage between these features and cell type: projection neurons have low background discharge and high stimulus selectivity, while interneurons have high background discharge and low stimulus selectivity. Does this hold for auditory responsive neurons in the BLA?

The BLA contains two main cells types. Glutamatergic, pyramidal-like projection neurons make up the large majority (>90%) of BLA neurons, with the remaining identified as local-circuit inhibitory neurons (Washburn and Moises, 1992; Rainnie et al., 1993; Sah et al., 2003). Background discharge rate is a useful but not definitive feature in characterizing these two cell types. Neurons with rates below 1 spike/s are generally considered to be projection neurons (Paré and Gaudreau, 1996; Rosenkranz and Grace, 1999, 2001; Likhtik et al., 2006). However, other projection neurons have background rates as high as 7 spikes/s, with a substantial number in the 2-5 spikes/s range (Likhtik et al., 2006). Likhtik and colleagues proposed that neurons above 7 spikes/s are almost certainly interneurons, but neurons with rates in the 2-7 spikes/s could not be authoritatively classified. They further showed that spike width is not a reliable identifier of the two cell types in the BLA, unlike other brain regions (Domich et al., 1986). The differences in background discharge may relate to underlying differences in resting membrane potential, synaptic inhibition, and voltage-gated conductances that have been characterized in BLA neurons (Washburn and Moises, 1992; Rainnie et al., 1993; Pare et al., 1995; Sah et al., 2003).

Except for the inclusion in our study of four neurons with rates above 7 spikes/s, our distribution of background rates from awake big brown bats corresponds remarkably well to the
rates of projection neurons identified in isoflurane-anesthetized cats (Likhtik et al., 2006). This correspondence, and the established literature described above, strongly suggests that the neurons we categorized as “low background” are projection neurons. Also based on this correspondence, we conclude that the “high background” group in our study likely includes a substantial number of projection neurons, as well as interneurons. This conclusion is supported by the predominance of projection neurons in the BLA (Washburn and Moises, 1992; Rainnie et al., 1993; Sah et al., 2003). In our study as in that of Likhtik and colleagues, there was no clear relationship between background discharge, spike width, and, by inference, cell type.

Our results clearly show a close relationship between background discharge rate and auditory stimulus responsiveness/selectivity. Neurons with background discharge in the 0-1 spike/s range generally responded to 0-1 syllables and 0-1 vocal sequences. Neurons with background discharge only slightly higher, in the 2-5 spikes/s range, generally responded to at least 75% of syllables and vocal sequences. The new finding for the BLA is that there are two forms of representation of vocal stimuli that likely characterize BLA outputs. Some previous studies of BLA responses to vocal stimuli appear to record from subsets of these neural populations. Our previous study in big brown bats (Gadziola et al., 2012b) recorded mostly from high background neurons with average background firing of ~5 spikes/s and found that nearly all of these neurons were highly responsive to vocal stimuli. Parsana et al (2012) reported similar background firing rates of ~5 spikes/s in their freely moving rats. Such neurons may be a high background subset of all rat BLA neurons as recorded by Bordi and coworkers under similar conditions (Bordi et al., 1993). It is not clear how selective these neurons were to a full range of social vocalizations. In the mustached bat, two studies report similar background discharge rates (~3.7 spikes/s) that appear to include both low and high background neurons as defined in our
study (Naumann and Kanwal, 2011; Peterson and Wenstrup, 2012). Both mustached bat studies show population selectivity for syllables that is overall higher than that demonstrated in the big brown bat. While species may vary in the use of representation strategies involving high selectivity versus high responsiveness types, our data suggest that a consideration of the background discharge rates among the sample of neurons may be crucial for understanding the nature of the vocal representation.

**Neural Coding Strategies**

There are likely multiple neural circuits within the basolateral amygdala, responsible for coding different aspects of behaviorally relevant stimuli. Other groups have found strong evidence for valence-coding neurons (Paton et al., 2006; Belova et al., 2007, 2008), intensity-coding neurons (Anderson and Sobel, 2003; Small et al., 2003), and even fear extinction-resistant neurons (An et al., 2012). The present study suggests that neurons within the basolateral amygdala employ a variety of neural strategies for encoding social vocalizations.

*Representation of Vocal Stimuli by Low Background Neurons.* In the big brown bat BLA, presumptive projection neurons with low background discharge (≤ 1 spike/s) provide a distinct representation of vocal stimuli. These neurons display a robust response to one or a very few stimuli. The response is usually phasic and does not show synchronous firing to the individual syllables within a vocal sequence or to a train of acoustic stimuli. The sampled population responded to each of the syllables and sequences in roughly equal proportion. These observations suggest a sparse code of vocalizations, in which the firing of an individual neuron indicates the occurrence of a particular vocalization and the population as a whole may efficiently encode a range of vocalizations (Olshausen and Field, 2004). There is no clear
indication among this sample of a preference for vocalizations with negative or positive affect, but the sample size may be too small to determine.

Similar to the low background neurons in this study, the majority of BLA neurons in the mustached bat are highly selective, preferentially responding to only a few of the tested syllables (Naumann and Kanwal, 2011; Peterson and Wenstrup, 2012). However, the mustached bat sample included both high and low background neurons (data from Peterson and Wenstrup 2012). Furthermore, although both antagonistic and affiliative calls evoked excitation, there was a predominance of excitatory responses to antagonistic signals, especially to a broadband noisy signal, similar to the rBNBI, used in aggressive interactions. Thus, the predominant form of representation, at least to syllables, was a sparse code in which vocal stimuli are represented by a population of selective neurons that use a rate code to discriminate among vocal stimuli.

Representation of Vocal Stimuli by High Background Neurons. The present study finds that both isolated syllables and vocal sequences evoke stimulus-dependent responses across a population of high background neurons. In big brown bats, this population forms a second, distinct representation of vocal stimuli that is based on broadly responsive neurons that discriminate among stimuli on the basis of response magnitude and timing. Gadziola et al. (2012b) demonstrated the importance of response duration in discriminating among syllables in the vocal repertoire, and the present study further supports the importance of spike timing to the neural code. Spike timing becomes particularly informative when neurons respond to natural vocal sequences.

We find that neurons may employ different coding strategies depending on the complexity of acoustic signals. We relied on a spike distance metric to compare the relative performance of rate and temporal codes within the population of high background neurons when
responding to syllables or vocal sequences. We focused our analysis on the high background neurons, as these broadly responsive neurons had the potential to discriminate among the stimulus set. These metrics are not suitable for neurons with a high degree of selectivity for stimuli, or for neurons that discharge no, or very few, spikes in response to many of the tested stimuli. Among the high background population, we find that neural responses are able to transmit more information about vocal sequences than syllables, and that vocal sequences show significant gains in information when temporal information is considered. From work in other sensory systems, it has been suggested that pure rate codes are sufficient at signaling gross sensory differences, but that spike timing is important for signaling finer differences (Samonds and Bonds, 2004; Victor, 2005). Thus, spike timing may be particularly important for coding complex acoustic sequences (Narayan et al., 2006; Schnupp et al., 2006; Huetz et al., 2009) and for signaling more subtle differences between the intensity of a behavioral state (i.e. lower and high levels of aggression). Further, as natural vocal sequences often contain several rapidly repeated elements, different neural strategies may be adopted depending on the repetition rate (Wang et al., 2008b). The observed enhancement in discriminating social context provided by vocal sequences is carried, at least in part, by the temporal patterning of spikes. These findings contribute significantly to our understanding of valence coding in the amygdala. By highlighting the importance of spike timing to the neural code, our results provide a potential explanation for why some human imaging studies have failed to detect valence coding within the amygdala (Anderson and Sobel, 2003).

Although our spike timing analyses in the present and previous study (Gadziola et al., 2012b) demonstrate the presence of valence-coding by BLA neurons, the recorded population of high background BLA neurons reveals a striking inconsistency in responses to the two examples
of vocal sequences that are frequently observed during high aggression interactions (Gadziola et al., 2012a). These two sequences differ acoustically from one another in that the “High Aggression 1” sequence included many low-frequency, broadband noise syllable types (e.g. rBNBs, rBNBl) and evoked robust responses in the amygdala, whereas the “High Aggression 2” sequence contained frequency-modulated syllables (e.g. sAFM, sHFM, UFM) and was relatively ineffective at eliciting a strong population response. Although these frequency-modulated syllable types evoke strong population responses when presented alone, responses are suppressed when combined within natural sequences. What underlies this apparent discrepancy?

Although both example sequences are commonly observed during behavioral interactions classified as high aggression, there may have been a more subtle distinction that we were unable to differentiate in an earlier behavioral study (Gadziola et al., 2012a) based on overt behavioral displays. While classification labels of “aggression” help specify the type of negative affect associated with the vocalization, there are several distinct forms of aggression (e.g. conspecific, defensive, predatory). Thus, there is not likely to be a unitary state with a single neural representation, but rather multiple forms depending on the underlying survival circuit activated (Moyer, 1976; LeDoux, 2012). Moreover, within an aggressive social interaction, it is possible that the same vocalizing bat may oscillate between different forms of aggression, or even between different emotional states. This differential responsiveness may explain why high aggression vocalizations evoked less change in HR compared to lower aggression signals: if increased HR reflects activity within the amygdala, then the “High Aggression 2” sequence would not be expected to evoke as large a change in HR as the High Aggression 1 sequence.
Challenges in the study of auditory responses of BLA neurons

In our evaluation of the role of the amygdala in responses to natural acoustic stimuli, it has become clear how much of the conclusion depends on a broad set of experimental variables. Thus, our observation of two forms of representation of acoustic stimuli depended on the recording of neurons with a broad range of background firing rates, including those neurons that are difficult to detect with standard approaches.

A second consideration includes the role of species differences. There are differences in the representation of vocal stimuli in different species of bats, and further differences in representation in rats and mice. For instance, rat BLA neurons are primarily excited by antagonistic vocalizations, while many big brown bat neurons respond across the affective range. Species effects may include different forms of representation of vocal stimuli, but may also include species-typical responses to restraint, a contextual feature that alters BLA neuronal function (Shekhar et al., 2005; Kavushansky and Richter-Levin, 2006; Roozendaal et al., 2006), including responsiveness to sounds (Grimsley et al., 2015).

While there has been clear evidence that other sensory context and internal state affects amygdalar responses to sensory stimuli (Yasoshima et al., 1995; Yan and Scott, 1996; LeDoux, 2000, 2007; Kuraoka and Nakamura, 2007; Grimsley et al., 2013), previous work and the current study indicate additional effects related to acoustic stimuli. As syllable ordering and temporal emission patterns provide the necessary acoustic cues for distinguishing the affective context of vocalizations (Banse and Scherer, 1996; Fichtel and Hammerschmidt, 2002; Fischer et al., 2002; Rendall, 2003; Lucas et al., 2004; Arnold and Zuberbuhler, 2006; Bastian and Schmidt, 2008; Bohn et al., 2008; Gadziola et al., 2012a; Coutinho and Dibben, 2013), we hypothesized that amygdalar neurons would show better discrimination of vocalizations, and of the social context,
when presented with natural vocal sequences rather than isolated syllables. The results support
this hypothesis, showing that the important role of spike timing in neuronal discrimination of
vocal stimuli and of valence is only revealed in testing based on vocal sequences. Our study
further shows that BLA neurons do not generally display synchrony to the elements of acoustic
trains presented at high rates. Vocal sequences that include high syllable rates elicit either onset
responses or temporal summation effects, e.g., for lower aggression sequences. Finally, the
response to syllables is not highly informative of the response to sequences that contain those
syllables, suggesting that combinatorial interactions may play important roles. Even here,
however, there may be species effects. The study by Parsana et al. (2012) suggests that there is
relatively little difference in responsiveness or valence coding by BLA neurons when responding
to negative-valence 22 kHz vocal sequences vs. 22 kHz continual tones. We have no doubt that
the high background BLA neurons in the big brown bat would distinguish between these stimuli.

Other stimulus presentation parameters, such as stimulus repetition rate or randomization,
will likely alter the meaning of the stimulus to a listener and modulate amygdalar
responsiveness. Thus, comparisons across studies may prove difficult if stimulus presentation
parameters are not the same. For example, in the mustached bat amygdala, Peterson and
Wenstrup (2012) find that amygdalar neurons are less responsive to artificial stimuli, more
selective for vocalizations, and display a greater degree of persistent activity compared to the
results of Naumann and colleagues (2011). These differences are likely attributable to
methodological differences in stimulus repetition rates, as well as stimulus ordering (e.g.
repeated presentation vs. pseudo-randomized). Future studies should focus on the specific effects
of repetition rate and stimulus ordering procedures to better understand the influence of the
acoustic context on amygdalar responses.
ACKNOWLEDGEMENTS

We thank Kelsey Gerbig for assisting with electrocardiogram recordings, Carol Grose for technical assistance, and Jasmine Grimsley and the Auditory Neuroscience Group at Northeast Ohio Medical University (NEOMED) for discussion of data. We are grateful to the Ohio Department of Natural Resources for permission to collect bats.

Present address of Marie Gadziola: Department of Neurosciences, School of Medicine, 2109 Adelbert Avenue, Case Western Reserve University. Cleveland, OH, 44106-4975

GRANTS

This study was supported by research grant R01 DC00937-23 (JJW) from the National Institute on Deafness and Other Communication Disorders of the U.S. Public Health Service.
REFERENCES


**FIGURE LEGENDS**

**Figure 1.** All syllable types used as stimuli. Waveforms (*upper trace*) and spectrograms (*lower trace*) for 13 representative syllable types previously identified as neutral (E), aggressive (DFMs, UFM, rBNBs, rBNBl, QCF, torQCF, sHFM, sAFM, and sinFM) or appeasing (DFMl, DFMI-QCFl, and DFMI-QCFI-UFM) based on the behavioral context in which they were predominately observed (Gadziola et al., 2012a). Peak amplitude of all syllables is normalized. The time interval between syllables is for illustrative purposes and does not reflect natural emission rates. Syllable types were named according to acoustic structure (Gadziola et al., 2012a).

**Abbreviations:** E, echolocation; DFM, downward frequency modulation, short; UFM, upward frequency modulation; rBN, rectangular broadband noise burst, short; rBNBl, rectangular broadband noise burst, long; QCF, quasi-constant frequency; torQCF, torus quasi-constant frequency; sHFM, single-humped frequency modulation; sAFM, single-arched frequency modulation; sinFM, sinusoidal frequency modulation; DFMl, downward frequency modulation, long; DFMI-QCFl, downward frequency modulation long to quasi-constant frequency long; DFMI-QCFI-UFM, downward frequency modulation long to quasi-constant frequency long to upward frequency modulation.

**Figure 2.** Acoustic sequences used as neurophysiological stimuli. Waveforms (*upper trace*) and spectrograms (*lower trace*) for 8 acoustic sequences. For each context (*A-D*), two examplars were used. Except for the tone sequence, natural vocal sequences were selected that are representative of such sequences (Gadziola et al., 2012a). Sequence composition: Neutral 1 (Biosonar), natural sequence of echolocation calls; Neutral 2 (Pure Tones), sequence of five 25 kHz tone bursts; Appeasement 1, repeated syllable pairs of DFMI-QCFI-UFM and DFMI; Appeasement 2, repeated quartet of DFMI-QCFI-UFM, DFMI-QCFI-UFM, DFMI-QCFI-UFM,
Figure 3. Differential modulation of heart rate (HR) in response to sound. (A) Mean change in HR calculated by subtracting the baseline rate from the maximum heart rate value. Values were averaged across all trials for each of the five behavioral contexts. Error bars are 95% confidence intervals. Post-hoc tests with Bonferroni correction revealed that lower aggression calls evoked a greater change in HR when compared to all other behavioral contexts. Higher-intensity tones evoked a greater change in HR compared to lower-intensity tones.

Asterisk, $p<0.05$, double asterisk, $p<0.01$, triple asterisk, $p<0.001$. (B) Mean instantaneous heart rate and 95% confidence intervals averaged across lower aggression (red) and high aggression (black) trials. HR was first normalized to the average baseline value across all trials. Black bar represents timing of sound presentation.

Figure 4. Distribution of background firing rates. A histogram of background firing rates from all recorded neurons. Among auditory responsive neurons, there is a non-normal (Kolmogorov–Smirnov test, $p<0.001$) and bimodal distribution (coefficient of bimodality, $b>0.55$) with a break at 1 spike/s. Histogram bins are 0.5 spikes/s.

Figure 5. Background discharge is related to auditory responsiveness. (A) Background firing rate for high (yellow-shaded squares) and low (black circles) background neurons plotted against the number of syllables (left) and sequences (right) that evoked a response from each neuron. Horizontal dashed line indicates 1 spike/s cutoff for low background neurons. Although some high background neurons were highly selective, most responded to multiple stimuli (mean = 11/13 syllables, and 7/8 sequences). Neurons with low background rates respond to very few
stimuli. \textbf{(B)} Left: High background neurons (n=21) responded to most acoustic stimuli. Social vocalizations were named according to acoustic structure (Gadziola et al., 2012a). Right: Low background neurons (n=22) responded selectively, but represented the range of acoustic stimuli tested.

\textbf{Figure 6.} A low background neuron (0.1 spikes/s) was highly selective, with a clear preference for the DFMI-QCFl syllable. For the thirteen syllables \textbf{(A)} and eight sequences \textbf{(B)} tested, mean firing rate \textit{(black line)} and 95\% confidence intervals \textit{(light grey outline)} are shown, binned at 10 ms. The \textit{Appeasement 2} sequence includes the DFMI-QCFl syllable, but note the clear suppression in response relative to when the syllable is presented alone. \textit{Black bar} represents stimulus timing and duration for syllables, whereas the waveforms of vocal sequences are plotted above. \textit{Inset:} overlaid spike waveforms demonstrate a well-isolated neuron.

\textbf{Figure 7.} A high background neuron (4.5 spikes/s) responded to all of the tested stimuli, but showed clear differences in discharge pattern. For syllables \textbf{(A)} and sequences \textbf{(B)}, mean firing rate \textit{(black line)} and 95\% confidence intervals \textit{(light grey outline)} are shown, binned at 10 ms. \textit{Black bar} represents stimulus timing and duration for syllables, whereas the waveforms of vocal sequences are plotted above. Note that responses to individual syllables do not always predict responses to sequences. \textit{Inset:} overlaid spike waveforms demonstrate a well-isolated neuron.

\textbf{Figure 8.} Population responses to isolated syllables. Eight representative syllables are illustrated, organized by the associated behavioral context: neutral syllable \textit{(black label)}, appeasement syllables \textit{(blue labels)}, and aggression syllables \textit{(red labels)}. Colored plots of population activity represent each high background neuron in a single row. Each 10 ms response bin is colored based on excitation \textit{(red)}, inhibition \textit{(blue)}, or no significant difference \textit{(black)}
compared to the 200 ms pre-stimulus background window. Order of neurons is identical across all plots. Population responses are summarized below by calculating the proportion of neurons showing excitatory discharge at each 10 ms bin. Peak responsiveness is similar across the majority of syllables, regardless of associated behavioral context. Black bars represent stimulus timing and duration.

**Figure 9.** Population responses to eight vocal sequences at two different sound levels (0 dB and 40 dB attenuation). Responses show unique temporal activation patterns that could not be predicted from responses to isolated syllables. Aggressive vocalizations (red labels) show a greater proportion of neurons with excitatory discharge than appeasement vocalizations (blue labels). At the lowest sound level (40 dB attenuation), population responses to three aggression sequences remained strong, while responses to appeasement and neutral sequences (black labels) dramatically decreased.

**Figure 10.** Correspondence between syllable and sequence responses. (A) Distribution of the percent correspondence values based on syllable responses. The number of significant sequence responses that are expected based upon observed significant responses to syllables was counted for each neuron and expressed as a percentage. Median percentage was 73.3%, indicated by gray line on histogram. (B) Distribution of the percent correspondence values based on sequence responses. The percentage of significant syllable responses that are expected based upon significant observed responses to sequences was computed for each neuron. Median percentage was 22.5 (gray line).

**Figure 11.** Responses of amygdalar neurons as a function of syllable or tone repetition rate. *Left:* low background neurons (n=13). *Right:* high background neurons (n=12). Each graph displays the percentage of sample neurons responding to (white bars) or synchronized to (black bars).
bars) stimulus repetitions. n.s. indicates no significant auditory response (white bars) or no significant synchronization (black bars) to any rate of repetition of the test syllable or tone. Most low background neurons did not synchronize to or respond to repetitive syllables or tones. In contrast, most high background neurons responded across syllable repetition rates. However, very few of these neurons synchronized to repetition rates above 10/s.

**Figure 12.** Comparison of neural coding strategies based on spike distance metric. (A) Discrimination among vocal stimuli. Top: Information for discriminating among syllables (filled circles) and sequences (open squares) across a range of temporal integration values. Error bars show 95% confidence intervals. Bottom: Difference between information at best temporal integration value versus information in rate code, for both syllables and vocal sequences. Vocal sequences carried more information than syllables in the rate code, and gained a significant amount of information when the timing of spikes was considered. The amount of information gained from timing information was significantly different for syllables vs. sequences (Wilcoxon rank sum test, p<0.001). Error bars show 95% confidence intervals of the median. (B) Discrimination among behavioral contexts associated with vocal stimuli. Top: Information for discriminating among the contexts associated with vocal stimuli (biosonar, appeasement, lower aggression, and high aggression) for syllables (filled circles) and sequences (open squares) across different temporal integration values. Bottom: Difference between context discrimination information at best temporal integration window for syllables and sequences. Context discrimination information was higher in response to vocal sequences when spike timing was considered, and spike timing information contributed more to discrimination among behavioral contexts in response to sequences than syllables (Wilcoxon rank sum test: p<0.001).
A  Neutral 1 - Biosonar

B  Appeasement 1

C  Lower Aggression 1

D  High Aggression 1

Neutral 2 - Pure Tones

Appeasement 2

Lower Aggression 2

High Aggression 2
A

Low background
High background

Background Firing Rate (spikes/s)

Number of Syllables Eliciting Response

Number of Sequences Eliciting Response

B

High Background Rate

Low Background Rate

Percent Responsive Neurons

Artificial
Appeasement
Biosonar
Aggression

BBN
Tone
Neutral Sequence
Biosonar Sequence
Appeasement 1
Appeasement 2
Lower Aggression 1
Lower Aggression 2
High Aggression 1
High Aggression 2
E
DFMI
DFMI-QCFI
DFMI-QCFI-UFM
DFMs
UFM
rBNB
rBNI
QCF
torQCF
sHFM
sAFM
sinFM

Percent Responsive Neurons
No Change

Suppression

Excitation

DFM1

sAFM

rBNBl

DFM1-QCFI

DFM1-QCFI-UFM

DFMs

sinFM

Percent Neurons

Neuron ID

Percent Neurons

Neuron ID

Time (ms)
Synchronized Response

Auditory Response

% Neurons (n = 13)

Low Background

High Background

25 kHz

DFMl-QCFI-UFM

DFMs

rBNNs

Repetition Rate (Hz)