Representations of Cat Meows and Human Vowels in the Primary Auditory Cortex of Awake Cats

Ling Qin, Jing Yu Wang, and Yu Sato

1Department of Physiology, Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Chuo, Japan; and 2First Affiliated Hospital, China Medical University, Shenyang, People’s Republic of China

Submitted 11 October 2007; accepted in final form 27 February 2008

Qin L, Wang JY, Sato Y. Representations of cat meows and human vowels in the primary auditory cortex of awake cats. J Neurophysiol 99: 2305–2319, 2008. First published February 27, 2008; doi:10.1152/jn.01125.2007. Previous investigation of neural responses to cat meows in the primary auditory cortex (A1) of the anesthetized cat revealed a preponderance of phasic responses aligned to stimulus onset, offset, or envelope peaks. Sustained responses during stationary components of the stimulus were rarely seen. This observation motivates further investigation into how stationary components of naturalistic auditory stimuli are encoded by A1 neurons. We therefore explored neuronal response patterns in A1 of the awake cat using natural meows, time-reversed meows, and human vowels as stimuli. We found heterogeneous response types: ~2/3 of units classified as “phasic cells” responding only to amplitude envelope variations and the remaining 1/3 were “phasic-tonic cells” with continuous responses during the stationary components. The classification was upheld across all stimuli tested for a given cell. The differences of phasic responses were correlated with amplitude-envelope differences in the early stimulus portion (<100 ms), whereas the differences between tonic responses were correlated with ongoing spectral differences in the later stimulus portion. Phasic-tonic cells usually had a characteristic frequency (CF) <5 kHz, which corresponded to the dominant spectral range of vocalizations, suggesting that the cells encode spectral information. Phasic cells had CFs across the tested frequency range (<16 kHz). Instantaneous firing rates for natural and time-reversed meows were different, but mean rates for different categories of stimuli were similar. Evidence for cat’s A1 preferring conspecific meows was not found. These functionally heterogeneous responses may serve to encode ongoing changes in sound spectra or amplitude envelope occurring throughout the entirety of the sound stimulus.

INTRODUCTION

Species-specific vocalizations play an important role in animal communication. These vocalizations are characterized by multidimensional acoustic parameters such as intensity, FM, AM, component frequencies, and pitch. They also contain a variety of behavioral referents such as the presence of different kinds of predators, social relationships, and food (Seyfarth and Cheney 2003). It has been shown that a marmoset’s vocalizations may be represented by the discharges of scattered neuronal populations in A1, which are synchronized to the stimulus events during the course of a vocalization (Wang et al. 1995). The amplitude envelope was shown to be a preferred acoustic feature of the marmoset’s A1 neurons to encode the monkey’s twitter vocalizations (Nagarajan et al. 2002). Because the twitter calls, which are composed of a series of transient phrases, are rich in changes in the amplitude envelope, these kinds of calls could be well represented by neural responses during their entire course.

With regard to the neuronal representations of vocalizations in the auditory cortex, an earlier hypothesis was that species-specific vocalizations were selectively represented by the activity of call-specific neurons, i.e., call detectors. However, electrophysiological studies showed that the percentage of call-specific cells was very small, at least in the initial stages of the auditory cortical pathway (Funkenstein and Winter 1973; Manley and Muller-Preuss 1978; Newman and Wollberg 1973; Winter and Funkenstein 1973; Wollberg and Newman 1972). More recent results suggest instead that vocalizations are encoded by the activity of neuronal populations. In other words, a particular cluster of neurons may respond to a particular acoustic feature or a combination of acoustic features. Whole vocalizations are represented by the discharge patterns of spatially distributed neuronal populations that encode the decomposed acoustic features (Creutzfeldt et al. 1980; Gourevitch and Eggermont 2007; Nagarajan et al. 2002; Rauschecker et al. 1995; Recanzone 2000; Rotman et al. 2001; Wallace et al. 2005; Wang et al. 1995).

The auditory cortex is hierarchically organized. The neurons in primary auditory cortex (A1) tend to encode simple acoustic features such as the temporal envelope for a call and have less selectivity for vocalizations (Glass and Wollberg 1983a; Nagarajan et al. 2002; Pellegr-Toiba and Wollberg 1991; Schnupp et al. 2006; Wallace et al. 2005; Wang et al. 1995). The neurons in the specify areas are more sensitive to a specific combination of acoustic features and thus have higher selectivity for vocalizations (Cohen et al. 2007; Esser et al. 1997; Gourevitch and Eggermont 2007; Rauschecker et al. 1995; Romanski et al. 2005; Theunissen and Doupe 1998; Tian et al. 2001). Combining inputs from different acoustic features may constitute neural representations of the behavioral referents of vocalizations.

It has been shown that a marmoset’s vocalizations may be represented by the discharges of scattered neuronal populations in A1, which are synchronized to the stimulus events during the course of a vocalization (Wang et al. 1995). The amplitude envelope was shown to be a preferred acoustic feature of the marmoset’s A1 neurons to encode the monkey’s twitter vocalizations (Nagarajan et al. 2002). Because the twitter calls, which are composed of a series of transient phrases, are rich in changes in the amplitude envelope, these kinds of calls could be well represented by neural responses during their entire course.

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.
However, for vocalizations with few abrupt changes in the amplitude envelope such as a cat’s meows, neurons in cat A1 were generally reported to mainly respond to the onset and/or offset of the meows (Gehr et al. 2000; Gourevitch and Eggermont 2007; Rotman et al. 2001; Sovijärvi 1975). To increase their coding efficiency, A1 neurons were proposed to use multiple parameters, including the firing rate, spike timing, and spike synchronization, to encode the acoustic information of meow calls (Gourevitch and Eggermont 2007; Schnupp et al. 2006). It has been demonstrated that low spike rates, if temporally aligned, can convey stimulus related information (de-Charms and Merzenich 1996).

However, one limitation of previous studies on neural responses to meow calls is that they were conducted on anesthetized cats (Gehr et al. 2000; Gourevitch and Eggermont 2007; Rotman et al. 2001). In the auditory cortex of the guinea pig, anesthesia has been found to suppress the temporal patterns of responses to vocalization, thereby facilitating onset responses (Syka et al. 2005). Thus there is a possibility that the A1 neurons in awake cats have additional discharge patterns that are suppressed by anesthesia. This study therefore examined the discharge patterns in awake cats to investigate their functional significance in vocalization discriminations. We also examined the ability of A1 neurons to process the behavioral referents of cat meows.

We investigated the neural responses to cat’s natural meows, time-reversed meows, and human vowels (a, o, u, e, i) in the A1 of awake cats. Time-reversed meows have all the acoustic features of natural meows with the exception that the temporal order is reversed, thereby resulting in the loss of potential behavioral referents. Human vowels may have no behavioral meanings for naïve cats, but they share many similar acoustic features with meows such as long duration, low spectral range, and lacking abrupt temporal change. We used these vocalizations with similar acoustic features as stimuli to examine whether cat’s A1 could differentiate conspecific meows from other ethological irrelevant foils, based on potential behavioral referents.

We found that some A1 neurons in the awake cat showed additional tonic discharge patterns, responding to spectral features of the stimuli. Additionally, we found no differential responses among the natural meows, time-reversed meows, and human vowels.

METHODS
Animal preparation, recording, and histology

Experiments were performed in accordance with the Guidelines for Animal Experiments, University of Yamanashi, and the Guiding Principles for the Care and Use of Animals approved by the Council of the Physiological Society of Japan. Animal preparation, recording, and histological procedures were the same as described previously (Chimoto et al. 2002; Qin et al. 2005, 2007). Four cats were chronically prepared for single-unit recordings from the auditory cortex. Under pentobarbital sodium anesthesia (initial dose, 40 mg/kg) and aseptic conditions, an aluminum cylinder (inner diameter, 12 mm) was implanted bilaterally in the temporal bone for microelectrode access at an angle of 10–20° from the sagittal plane. A metal block was embedded in a dental acrylic cap to immobilize the head. After >1 wk of postoperative recovery, the cat’s body was gently wrapped in a cloth bag. The head was restrained with holding bars for a short period. In successive daily sessions, the restraining period was lengthened, and the cats were familiarized to sitting in an electrically shielded, sound-attenuated chamber. The animals were given food and drink during the sessions, and after each session, they were returned to their home cages. The conditioning process lasted for ≥2 wk. When the recording experiments began, the cats sat with no sign of discomfort or restlessness. A day prior to the beginning of the recording session, the bone (diameter, 1–2 mm) at the bottom of the cylinder was removed, leaving the dura intact; the procedure was performed under ketamine anesthesia (initial dose, 15 mg/kg).

The recording session began the following day in a soundproof chamber. The dura was pierced with a sharpened probe, and a single epoxylite-insulated tungsten microelectrode (FHC, impedance: 2–5 MΩ at 1 kHz) was advanced into the A1 with a remote-controlled micromanipulator (Narisige, MO-951). The extracellular single-unit activity was discriminated using a template-matching discriminator (ASD, Alpha-Omega Engineering). The digital ASD outputs of the spike occurrence time (time resolution: 50 μs) were stored on a hard disk. The characteristic frequency (CF) of the recording site was determined by measuring the tonal responses of the surface neurons in each electrode track.

The cat’s face, particularly the eyes, was continuously observed on a monitor connected to a charge-coupled-device camera. When drowsiness was suspected, the cat was alerted by gently tapping the body with a remote-controlled device or by briefly opening the door. For each animal, the daily recording sessions lasted for 3–5 h over 2–6 mo. At the end of each daily recording session, the recording chamber was rinsed with sterile saline and antibiotic fluid and sealed with Exafine (GC Corporation) and an aluminum cap. The animal was returned to its cage. All cats remained healthy throughout the experimental period. At the end of the experiment, some recording sites were marked by electrolytic lesions (25 μA, 10 s). The animal was deeply anesthetized with sodium pentobarbital and perfused with 10% formalin before the brain was removed. The brain surface was photographed. The cerebral cortex was cut in coronal sections and stained with neutral red. Based on the lesion locations and electrode tracks, the recording sites were reconstructed.

Acoustic stimulus presentation

Acoustic stimuli were presented from a speaker placed 2 cm away from the auricle contralateral to the recording site. The sound delivering system was calibrated to produce a flat spectrum (128–16,000 Hz, ±5 dB) at the entrance of the cat’s meatus. Cat vocalizations were collected from spontaneously vocalizing cats recorded individually in a sound attenuated room. The sounds were recorded by a microphone (ONO SOKKI LA-5110). The recorded signals were passed through a low-pass filter (cutoff frequency = 20 kHz), then connected to a computer and digitized using Spike2 software (Cambridge Electronic Design) with a sampling rate of 100 kHz. The vocalization stimuli were recorded from the cats of our breeding colony. We selected five well-defined vocalizations (meows 1–5) recorded in different situations from five different cats. Meows 1 and 2 were meows recorded prior to feeding. Meow 3 was emitted when a cat was slightly aggravated by a plastic bar touching its face. Meow 4 was a typical kitten call recorded when a kitten was isolated from its mother. Meow 5 was a short meow call elicited when a cat was first placed in the recording chamber. Brown et al. (1978) showed that the cat’s calls, emitted in these different situations, presented temporal and spectral structures containing different behavioral meaning and identity information. The amplitudes and spectrographic representations of the five calls are illustrated in Fig. 1, A–E. To generate the spectrographic representations, a 5,000-point (50 ms) Fourier transform was applied to consecutive time segments (sliding in 5-ms step) of the call. Therefore for each call, one pixel in the spectrogram covered 5 ms of the time domain and 20 Hz of the frequency domain. The power of each frequency bin ranged from 0 to 20 kHz and was illustrated in grayscale (Fig. 1, A–E, middle). The pixels with maximum power
were represented in black, those with less power (≤0 dB), in white. As shown by the spectrographs, the main energy of our stimuli concentrated on the low-frequency range. For the convenience of observation, the spectrograms within low-frequency range (<5 kHz) were amplified and displayed in Fig. 1, A–E, bottom.

The frequency contours of the harmonics in each call were extracted by using the method similar to DiMattina and Wang (2006). The peak pixel having the maximum amplitude in the whole spectrogram was found first. Then at its neighboring time window, we searched the next pixel having the maximum amplitude to extend the frequency contour. This frequency search procedure was consecutively conducted along the two directions of time axis, while restricting our frequency search range within ±100 Hz relatively to the preceding pixel. Using this method, we got the time-frequency contour $f_1(t)$ of the strongest harmonic, which has 5-ms time resolution and 20-Hz frequency resolution. Next the frequency contour of the second strongest harmonic, $f_2(t)$, was abstracted by repeating preceding search procedure on the spectrogram excluded $f_1(t) ± 100$ Hz. The third strongest harmonic, $f_3(t)$, was also defined after abstracted the first two harmonics. The small numbers inside the spectrograms in Fig. 1 marks the three strongest harmonics of each stimulus.

Each call had distinct acoustic features in both the spectral and temporal domains. For example, the spectral width of the harmonic components was wide in meows 1 and 4 (~4 kHz), medium in meows 2 (3 kHz) and 5 (2 kHz), and narrow (1 kHz) in meow 3. The fundamental frequency was relatively low (~300 Hz) in meows 1 and 2, medium in meows 5 (600 Hz) and 3 (500 Hz), and high (1,000 Hz) in meow 4. Many of the vocalizations exhibited a strong FM component. For example, meow 1 had upward FM at an early short period of the call and downward FM at a late period. The amplitude rise at the onset and the amplitude fall at the offset were relatively rapid in meow 2 but slow in meow 1. Meow 4 had a sharp peak in the middle of the

FIG. 1. Vocalization stimuli. A–E: amplitude envelopes (top) and spectrograms (middle, frequency range: 0–20 kHz; bottom, 0–5 kHz) of the 5 natural meows (meows 1–5) used in this study. The intensity is coded in grayscale with black as the maximum in the spectrogram. F–J: amplitude envelopes and spectrograms of the vowel stimuli. The numbers inside each bottom panel mark the 3 strongest harmonics of each spectrogram.
wave envelope. The five natural meows were reversed in the time domain to construct the time-reversed meow stimuli (rMeows 1–5).

The five vowels (/a, o, u, e, i/) in Japanese were also used as stimuli. The vowels were recorded from a male Japanese speaker and had been earlier used as stimuli in our human psychological experiments (Sakayori et al. 2002). The amplitude and spectrographic representations of the vowels are illustrated in Fig. 1, F–J. The vowels were characterized by a static harmonic spectrum without obvious frequency and AM. The fundamental frequency of this speaker’s vowels was 180 Hz. Different vowels could be easily identified by the first and second formant frequencies (F1 and F2) (Potter and Steinberg 1950). The respective values of F1 and F2 were 760 and 1,120 Hz in /a/; 420 and 850 Hz in /o/; 390 and 1,400 Hz in /u/; 540 and 2,000 Hz in /e/; and 310 and 2,400 Hz in /i/. These values are within the normal range of Japanese vowels (Tokizane 1951).

The natural meows, time-reversed meows, and vowels were presented in a random order at a peak level of 50 dB sound pressure level (SPL, dB re 20 μPa). Each vocalization was repeated 8–16 times (interstimulus interval >1.5 s).

Data analysis

For each neuron and each vocalization, the peristimulus time histogram (PSTH) of the firing rate was computed in 5-ms bin width and smoothed by Gaussian function with 5 ms SD. The firing rates during the 0.5 s before the onset of each stimulus were considered as background. The average of background firing rates was calculated and subtracted from the height of each time bin in the PSTH to obtain the driven rate histogram. A driven rate was considered significant when it was >2 SD of the background firing rates (response threshold). In the PSTH of each stimulus, we counted the number of 5-ms bins with significant driven rate during stimulus period plus 100 ms after stimulus offset. This number was multiplied by 5 ms to represent the “response duration” of the vocalization stimuli. Because offset responses in our data usually lasted no longer than 100 ms, the response duration in this study represents the overall length of both the response during stimulus and the offset response.

To compare the temporal response patterns across a cell population, the driven rate PSTHs of individual cells were presented together in three-dimensional plots (Fig. 2). First the response duration for meow 1 stimuli was measured for each cell, and all the recorded cells were sorted in ascending order of response duration for meow 1. Second, the individual PSTHs were normalized with the maximum height as 1 and the area less than the response threshold as 0. The normalized PSTHs were aligned at stimulus onset and arranged in the order of meow 1 response duration, constructing the population response patterns to different stimuli. The PSTH height was represented by a color scale (red represents 1, and blue represents 0). We used the same cell order, the rank of response duration for meow 1, to construct the 15 plots of different stimuli in Fig. 2 for convenience to compare the temporal response patterns across stimuli. We chose the response duration for meow 1 as the index to rank the cells because meow 1 had the longest duration (1.44 s) in our 15 stimuli.

Some cells showed a strong brief response during a restricted period (<50 ms) after the onset, offset, or sharp mid-portion envelope peak of the vocalizations. We termed such responses “phasic responses.” Some cells also showed a period of tonic excitatory activity during the portion of the vocalizations. Such responses were termed “tonic responses.”

RESULTS

We collected a total of 218 single units from both hemispheres of the A1 (CF between 128 and 16,000 Hz) of four awake cats. Of these, 157 (72%) cells responded to ≥1 of the 15 vocalization stimuli (Fig. 1). This report is based on the 157 A1 cells. The single-unit activities were recorded at a depth of 400–2,000 μm from the first encountered unit of each track. Histological reconstruction showed that the recording sites were in the caudal part of the middle ectosylvian gyrus, the banks of the dorsal tip of the posterior ectosylvian sulcus, and a small portion of the adjacent posterior ectosylvian gyrus, i.e., the caudal part of A1 (Reale and Imig 1980). The border between A1 and the adjacent posterior auditory field (PAF) was distinguished by the frequency reversal of the tonotopic order.

Neural response duration

Using 15 vocalization stimuli, including five natural meows, five time-reversed meows, and five vowels (Fig. 1), we investigated the temporal response patterns of A1 neurons in awake cats. Neural responses to the 15 vocalization stimuli are illustrated in Fig. 2 where the PSTHs (driven rate of each 5-ms bin in color scale) of all 157 cells are arranged in order of the response duration of meow 1 (see METHODS for details).

We noted that the cells showed various temporal response patterns with different response durations. Some cells with a shorter response duration (plots, bottom) showed tonic responses throughout the stimulus period. The driven rate (color of the dots) changed differently with the stimulus time across different cells, resulting in a complex map of responses.

We investigated the distribution of the response duration in 157 cells. The meow 1 response durations of individual cells were plotted in ascending order in Fig. 3A. The curve has a smooth and continuous rise from 0 to 1.44 s (duration of meow 1 of 1.44 s plus poststimulus analysis time of 0.1 s is 1.54 s), indicating that A1 has a cell group with diffusely distributed response durations for meow 1. Thus there was a continuum of response patterns among A1 from cells that only responded to stimulus onset/offset to those that responded throughout the stimulus period.

The PSTHs for the other 14 vocalizations were arranged in the same cell order as meow 1 (Fig. 2). Generally, cells with short responses to meow 1 (Fig. 2, top of plots) also responded to other calls longer. This trend is quantitatively shown in Fig. 3B, which compares the response durations of meows 1 and 2 in individual cells. Because the response duration across different vocalizations was constant in some cells with phasic responses locked to sound onset, it was inappropriate to normalize the response duration by the stimulus duration. Thus we directly compared the response durations of different vocalizations. There was a significant positive correlation (r = 0.71, P < 0.01, t statistic having 155 df) between the response durations of meows 1 and 2. Similarly, the response duration of meow 1 also correlated with that of rMeow 1 and /a/ (Fig. 3, C and D; r = 0.91 and 0.68, respectively). We further compared the response durations between any pair of the 15 vocalizations. The results are illustrated in Fig. 3B by the correlation matrix, which presents the 15 acoustic stimuli along a single column and a single row. Each cell in the matrix corresponding to the intersection of any unique pair of stimuli is assigned a gray scale proportional to the Pearson r value. All the correlations of the response durations were significant (P < 0.01, r
ranged from 0.43 to 0.93), indicating that a cell responds with a relatively consistent duration, regardless of stimulus type.

Additionally, a pair of natural and time-reversed meows showed relatively higher correlation, displayed by the dark cells paralleling to the diagonal line in Fig. 3E. This may be a consequence of the similarity of the acoustic features between the natural and time-reversed meows. As shown by the dark bottom-right corner in Fig. 3E, the correlations between the response durations of five vowels as a whole were higher than the correlations between vowels and meows or reversed meows. This may be also attributable to the similarity in acoustic features among vowels (Fig. 1). Thus the neural response duration also showed some variations depending on the acoustic features of stimuli. The more the stimuli resembled each other, the more similar the response durations were.

**Efficacy of different vocalizations**

Next, we examined whether the stimuli from three different categories (natural meows, time-reversed meows, and human
vowels) activated A1 neurons with similar efficacy. The result would help us to know whether the A1 neurons could process the behavioral referents of meows. If the natural meows evoked different neural responses than the time-reversed meows and human vowels, this would suggest that A1 could capture the behavioral referents of meows.

Because the response duration was diverse among the neurons, it was inappropriate to evaluate a neuron’s responsiveness based on the discharge rate averaged over the entire stimulus duration, which will undervalue the response strength for cells with only phasic responses. Therefore we divided the neurons’ responses into a series of 50-ms segments starting from the onset of stimuli and then consecutively evaluated the mean driven rate over each 50-ms segment. If the driven rate during a specific 50 ms of the vocalization stimulus was >2 SD of the background firing rate, we considered the segment of vocalization to be effective in evoking the spike responses of a neuron. The efficacy of the first 50-ms segment of different vocalizations was expressed as the percentage of evoked neurons (Fig. 4A, □). For all 15 stimuli, the percentage of evoked neurons was always ~60% of the total number of recorded neurons. There was no significant difference between the percentages for the vocalization stimuli from the three different categories (ANOVA, P > 0.05). Thus all the vocalizations were equally effective in evoking neural responses during the first 50 ms.
For the 300- to 350-ms poststimulus time segment, the percentage of evoked neurons decreased to ~20% in all stimuli (Fig. 4A). This suggests that the onset segment was more effective in evoking a population response than the late segment. Further, in the 300- to 350-ms segment, there was no significant difference between the percentages of evoked neurons for the different stimulus categories (ANOVA, $P > 0.05$). This was also the case when we consecutively compared the percentages of evoked neurons during each 50-ms segment from 0 to 350 ms after stimulus onset (Fig. 4B).

We further examined the magnitude of the population response evoked by different vocalizations. The magnitude of the population response was evaluated by averaging the driven rate across all evoked neurons. Figure 4C shows the mean magnitude of the population response for each stimulus category as a function of the poststimulus time (50-ms segment). The magnitudes of responses to all stimulus categories (meows, reversed meows, and vowels) decreased similarly with shifting a 50-ms time window. No significant difference among the response magnitudes for different stimulus categories was found at any 50-ms segments (ANOVA, $P > 0.05$).

The analysis of both the percentage of evoked neurons and response magnitude revealed two facts. One was that the onset segment of vocalizations was more effective in evoking neural responses than the late segments. The other is that the natural meows, time-reversed meows, and vowels were equally effective in evoking neural responses in any segment.

Multiple temporal response patterns

Although the continuous distribution of response durations in A1 cells (Fig. 3A) suggested a continuum between the cells with and without the tonic responses, we assumed that different processing may underlie the phasic and tonic responses. Thus we attempted to classify the cells into two groups on the basis of the response duration of meow 1.

We examined whether the cells with different response durations have different response patterns. Figure 5 shows the means ± SE PSTHs during meow 1 stimulus for the cells in the first, second, and third 1/3 part of response duration distribution. The first 2/3 cells showed dominant phasic responses at stimulus on- and offset, while the last 1/3 cells showed tonic response throughout the stimulus with a weak phasic peak at the on- and offset. One interesting observation is that the magnitude of phasic response increased with increasing response duration (compare the PSTHs of the 1st and 2nd 1/3 cells). However, when the tonic response became dominant in the last 1/3 cells, the magnitude of the phasic response decreased. The other observation is that the first third of cells demonstrated significant suppression throughout the entire stimulus interval. Because this suppression was below spontaneous spiking rates, it is very likely to be created by inhibitory processes. All these observations support the thesis that “once phasic, always phasic,” meaning that the stimulus does not influence individual neuronal response phasicness/tonicness.

We therefore classified the cells into two groups according to

FIG. 4. Efficacy of different vocalizations in evoking neural responses. A: the percentage of neurons that responded to the individual stimuli during the 0- to 50-ms ($c$) and 300- to 350-ms ($d$) poststimulus periods. B: the percentage of responsive neurons plotted against the poststimulus time. The percentages of responsive neurons for vocalizations of the same category (meow, reversed meow, or vowel) were pooled together, and the averaged values of each category were represented by different symbols. C: the mean driven rate of excitatory responses plotted against the poststimulus time. The driven rate was averaged over the results from the same vocalization category.

FIG. 5. PSTHs of different parts of cells responding to meow 1 stimulus. The 1st, 2nd, and 3rd 1/3 cells in the response duration distribution were pooled separately. The curve shows the mean driven rate; shade region indicates ± SE.
the distribution of response duration of *meow 1*. The first 2/3 cells \((n = 104, \text{ response duration } < 0.485 \text{ s})\) with only phasic responses were termed the “phasic cells,” and the last 1/3 cells \((n = 53)\) with both phasic and tonic responses, “phasic-tonic cells.”

The average PSTHs of 104 phasic cells and 53 phasic-tonic cells to the 15 vocalizations are shown in Fig. 6 by red and blue lines, respectively. All the vocalizations evoked distinct phasic and tonic responses. As demonstrated in the preceding text, the phasic cells with only phasic responses to *meow 1* responded to all other stimuli in a phasic manner, whereas the phasic-tonic cells with both phasic and tonic responses to *meow 1* responded to other stimuli in a similar manner. The consistency of the responses across the different vocalization stimuli suggests that our classification was cell-specific rather than vocalization-specific.

**Correlation between the neural response and stimulus amplitude envelope**

By comparing the response time courses with the amplitude envelopes of vocalizations (Fig. 6, top in each panel), it was clear that the phasic responses of phasic cells were elicited only at the on- and offset of the vocalizations and the sharp peaks in the mid-portion of the amplitude envelope (for example, see the response to *meow 4* in Fig. 6). The phasic-tonic cells also showed some phasic responses at similar time points of stimuli (blue lines in Fig. 6). Because the onset, offset, and envelope peak of the stimuli were characterized by an abrupt change in the amplitude envelope, such a transient signal may elicit phasic responses from A1 neurons. The magnitude and speed of an amplitude change may be different at the onset or offset between any pair of different stimuli. Figure 7A shows the sound waves during the first 25 ms of the /*u*/ and /*i*/ stimuli in which the ordinate is the voltage of the output from the voice recording microphone. The rising speed of the amplitude envelope is obviously faster in /*i*/ (light line in Fig. 7A) than in /*u*/ (dark line). Therefore the question is whether the different amplitude envelope of sound onset could evoke different neural responses. In other words, could the A1 neurons detect the difference in amplitude envelope of stimulus within a short time interval?

![Fig. 6](http://jn.physiology.org/Download?http://jn.physiology.org/content/jn/99/5/2312/F6.large.jpg)
To answer this question, we examined the difference in the neural responses to the /u/ and /i/ stimuli. We first evaluated the onset response magnitude of each individual cell (the driven rate during the time window between 0 and 25 ms after stimulus onset) and then calculated the difference (absolute value) of the response magnitudes to the /u/ and /i/ stimuli in each cell. Figure 7B shows the distributions of the response difference for the 104 phasic and 53 phasic-tonic cells. Most of the phasic and phasic-tonic cells showed a difference in the response magnitudes to the two stimuli, suggesting that the neural responses were sensitive to the onset signals of vocalizations.

The stimulus “amplitude-envelope difference” during the 1-ms onset period of paired stimuli was quantified by the difference (absolute value) between their maxima of amplitude (in units of millivolt). The 105 amplitude-envelope differences (1 for each unique pair of stimuli) were plotted against the corresponding mean response difference of 104 phasic and 53 phasic-tonic cells, respectively (Fig. 7D and E). The phasic cells showed a significant positive correlation between the difference in response and that in the amplitude envelope ($r = 0.53, P < 0.01$), whereas the phasic-tonic cells did not ($r = 0.18, P = 0.07$). This suggested that the phasic cells were sensitive to the change of amplitude envelope during the first 1 ms of stimulus onset.

The amplitude-envelope difference between the paired stimulus onsets was also quantified within other longer time windows, including 2.5, 5, 10, 15, 20, and 25 ms after stimulus onset. Figure 7F illustrates the correlation coefficients between the mean response difference and the amplitude-envelope differences within various time windows. The time window for evaluating neural responses was fixed at 0–25 ms after stimulus onset, whereas the time window for evaluating amplitude envelope was extended gradually from 0–1 to 0–2.5, 0–5 ms, and so on. During the first 1 ms of onset, the correlation in the phasic cells was higher than in the phasic-tonic cells ($r = 0.53$).
With the elongation of stimulus duration, the correlation gradually decayed in the phasic cells, but incremented in the phasic-tonic cells (● vs. ○ in Fig. 7F). The firing rate of phasic-tonic cells became more correlated with the amplitude envelope than those of phasic cells when the longer period of the amplitude envelope was taken into consideration (5–20 ms poststimulus).

In summary, both the phasic and phasic-tonic cells were sensitive to the amplitude envelope of stimulus onset. The detectable temporal range of the stimulus amplitude envelope was different between the cell groups: the phasic cells were more sensitive to the earlier (<5 ms) signal, whereas the phasic-tonic cells responded to the later (>5 ms) signal (Fig. 7F).

Similar procedures were also performed to analyze the correlation between the offset neural responses (0–25 ms after stimulus offset) and the amplitude envelope of stimulus offset. The results showed that the offset response difference was significantly correlated with the amplitude envelope difference during the last 10 ms before stimulus offset (data were not illustrated). No clear difference was found between the offset responses of the phasic and phasic-tonic cells. For both the phasic and phasic-tonic cells, the maximum correlation coefficient appeared (r = 0.31 and 0.41, respectively) when the amplitude envelope was evaluated during the last 1 ms before stimulus offset. Thus the phasic responses, including both the on- and offset responses, were sensitive to the amplitude envelope within a short time interval of stimulus. However, the onset responses were more tightly correlated with the amplitude signals than the offset responses.

**Correlation between the neural response and stimulus spectrum**

In the preceding section, we showed that the phasic responses (onset/offset response) of the phasic and phasic-tonic cells correlated with the amplitude envelope in stimulus on- and offset. It should also be examined whether the phasic responses also correlate with the stimulus spectral contents. Therefore we measured the peak frequency of each stimulus spectrum (the frequency with the largest amplitude in the spectrum) during 0–25 ms after stimulus onset and then calculated the difference (absolute value) in the peak frequencies between the paired stimuli. The peak frequency represents the spectral location of the main energy of the stimulus; the peak frequency difference largely reflected the extent of spectral difference between the paired stimuli.

Correspondingly, the difference in the onset responses (0–25 ms after stimulus onset) between each stimulus pair was also measured. The mean response differences of phasic and phasic-tonic cell groups were plotted against the peak frequency difference in Fig. 8, A and B, respectively. No clear correlation between the mean response difference and peak frequency difference was found in either group of cells (r = 0.04 and 0.01, respectively). Similar procedures were also

---

**Figure 8.** Correlation between the neural response and stimulus spectrum. **A:** the peak frequency difference (during 25 ms after stimulus onset) between each stimulus pair vs. the mean response difference of the phasic cells. **B** and **C:** the peak frequency difference (0–25 ms and 150–175 ms) vs. the mean response difference of phasic-tonic cells. **D:** the time functions of correlation coefficient between the mean response difference and the differences of $f_1$, $f_2$, and $f_3$ (red, blue, and black lines, respectively). The abscissa is the start time of the shifting 25-ms window to calculate driven rate. **E:** the coefficient of correlation between the mean response difference and the difference of $f_1$ plotted against time. Lines with different colors indicate the results obtained by using various widths of time windows to calculate driven rate. **F:** the coefficient of correlation between the mean response difference and the amplitude-envelope difference plotted against time.
conducted to compare the offset responses (0–25 ms after stimulus offset) with the stimulus spectra during the 25-ms period before stimulus offset. No significant correlation between the offset response and peak frequency was found in either the phasic or phasic-tonic cells ($r = 0.18$ and 0.16, respectively). These results indicate that the phasic responses (both on and offset responses) were less sensitive to the difference in the stimulus spectrum.

On the contrary, when we conducted a similar analysis on the neural response and peak frequency within a 25-ms period in the middle of the stimuli (150–175 ms after stimulus onset), the tonic responses of the phasic-tonic cells showed an obvious correlation between the differences in the peak frequencies and the mean responses ($r = 0.58$, $P < 0.01$; Fig. 8C). This result suggests that the late neural responses to the mid-portion of the stimuli are sensitive to the spectral contents.

To further investigate the time course of neuronal sensitivity to the spectral contents, we systematically examined the correlation between stimulus spectral difference and mean response difference in our data. To evaluate the stimulus spectral difference, we conducted consecutive spectral analysis on our vocalization stimuli and extracted the frequency contours of three largest-amplitude harmonics from the spectrograms ($f_1$, $f_2$, and $f_3$, METHODS for details). The absolute difference between $f_1$ of any two stimuli was calculated and termed as “$f_1$ difference.” The “$f_2$ difference” and “$f_3$ difference” were also calculated similarly.

To measure the mean response difference, the difference in neural responses (driven rate) between any pair of stimuli was calculated by using a 25-ms time window shifted in 5-ms steps from 0 to 350 ms after stimulus onset. Because the phasic cells usually showed no spiking response to the mid-portion of the stimulus, this calculation was only conducted on the 53 phasic-tonic cells and the results were averaged across the 53 samples to get the mean response difference.

The correlation coefficient between $f_1$ difference (the strongest harmonic) and the mean response difference was calculated consecutively from the stimulus onset in 5-ms steps. The results were plotted against the start time of the analysis time window. The neural responses (driven rate) were re-analyzed by using various bin sizes from 10 to 100 ms. The result is shown in Fig. 8E. The $f_1$ difference was always used in this procedure. The time functions of the correlation coefficients had no obvious change when the bin size varied from 10 to 50 ms except a left shift. However, when the neural response was calculated by 100-ms bins, the mean response difference significantly correlated with $f_1$ difference throughout the stimulus duration, including the first 100-ms responses. This is explainable by assuming that the frequency-correlated responses started between 50 and 100 ms after stimulus onset and were included by the first 100-ms analysis window.

For a control, we investigated the relationship between the stimulus amplitude-envelope difference and the mean response difference within each analysis time window in the 53 phasictonic cells. The correlation coefficients were also plotted against the start time of the analysis time window, thereby constructing the time function of amplitude-envelope sensitivity (color lines in Fig. 8F). In contrast to the frequency sensitivity time function (Fig. 8E), the amplitude-envelope sensitivity time function was high during the early period soon after stimulus onset but quickly decreased and crossed the significant level at 100 ms after stimulus onset. This result further indicated that the phasic responses, occurring within 100 ms after stimulus onset, were sensitive to the amplitude-envelope, whereas the tonic responses, >100 ms after stimulus onset, were not.

The results in Fig. 8 reveal that A1 cells may represent vocalizations by using multiple encoding strategies. At the start of stimulus, the phasic responses encode the amplitude envelope of stimulus onset, whereas during the mid-portion of the stimulus, the tonic responses encoded the frequency contents. The amplitude change was prominent at stimulus onset and mild during the mid-portion of a cat’s meows and human vowels. Correspondingly, A1 shifts the encoding strategies from amplitude coding to frequency coding to discriminate the stimuli. Thus the possession of multiple encoding strategies may increase the efficiency of A1 cells to represent vocalizations.

Relation between response duration and CF

Of the 157 cells that responded to vocalizations, characteristic frequency (CF, frequency of the lowest amplitude pure tone that elicits a response) was available in 137 cells. We plotted the response duration of meow 1 against CF for each individual cell in Fig. 9A to investigate the relation between the response duration and CF. It was apparent that the cells having longer response duration were mostly concentrated in the range of low CF, whereas the cells having shorter response duration dispersed across all the tested CF range. The horizontal dotted line in Fig. 9A indicates the criterion level which we used to categorize the phasic and phasic-tonic cells in this report. Figure 9B is the scatter plot of the response durations for all the 15 stimuli (normalized by the duration of sound stimulus) versus cell’s CF, in which the distribution of long-duration responses were also biased to low CF. Hence, the vocalization stimuli could evoke the phasic response in a wide range of A1 area, whereas the tonic response only in low CF area.

We further examined the spectral range of our vocalization stimuli. The power spectrum of each stimulus was obtained by applying Fourier analysis on the whole sound wave. The spectra were normalized by their individual maximum SPL as 0. Figure 9C displays the energy distribution (mean and SD) of our stimuli at each frequency component (frequency resolu-
Patterns were also not changed by the temporal reversal of the 1st and 2nd rows of Fig. 6). The phasic-tonic response responded to reversed meows also in phasic patterns (red lines in neurons responding to natural meows in phasic patterns respond to reversed vocalizations, we could know the effect of the temporal order of vocalizations because the two stimuli have the same spectral content but reversed temporal contents. The appearance of tonic responses was not dependent on the type of vocalization but on the cell’s own properties. A1 neurons responded to vocalization exemplars from different categories with similar response durations and strength (percent of evoked neurons and mean driven rate). We hypothesize that A1 neurons might use multiple coding strategies to encode different acoustic features of vocalizations. The amplitude envelope at stimulus onset or offset was detected by the phasic response, whereas the difference in the frequency content might be encoded by the tonic response during the steady mid-portion of the stimuli. Tonic responses were mainly found in the portion of A1 where the cells’ CF matched the prevalent spectra of vocalizations, whereas the phasic responses were found throughout the tested A1 area (CF <16 kHz). And the PSTHs of spike responses to vocalizations were altered when the vocalizations were temporally reversed. These findings contribute toward our understanding of how A1 analyzes the acoustic features of complex sounds.

**DISCUSSION**

This study is the first to investigate neural responses to cat meows and human vowels in A1 of an awake cat. Five major findings were presented. 1) A1 neurons showed phasic responses to stimulus onset, stimulus offset, and sharp amplitude envelope peaks. Some also showed tonic responses to the steady mid-portion of stimuli. The appearance of tonic responses was not dependent on the type of vocalization but on the cell’s own properties. 2) A1 neurons responded to vocalization exemplars from different categories with similar response durations and strength (percent of evoked neurons and mean driven rate). 3) We hypothesize that A1 neurons might use multiple coding strategies to encode different acoustic features of vocalizations. The amplitude envelope at stimulus onset or offset was detected by the phasic response, whereas the difference in the frequency content might be encoded by the tonic response during the steady mid-portion of the stimuli. 4) Tonic responses were mainly found in the portion of A1 where the cells’ CF matched the prevalent spectra of vocalizations, whereas the phasic responses were found throughout the tested A1 area (CF <16 kHz). And 5) the PSTHs of spike responses to vocalizations were altered when the vocalizations were temporally reversed. These findings contribute toward our understanding of how A1 analyzes the acoustic features of complex sounds.

**Multiple temporal response patterns to vocalizations**

Previous studies on anesthetized cats have generally reported that A1 neurons mainly responded to some transient segments of a cat’s calls such as the onset, offset, or other envelope peaks but did not respond to other parts of the calls (Gehr et al. 2000; Rotman et al. 2001). This is in agreement with the general findings that the auditory neurons in anesthetized animals responded to sound stimuli in a phasic manner (deCharms and Merzenich 1996; DeWeese et al. 2003; Eggermont 1997; Heil 1997; Phillips 1985; Schnupp et al. 2001). In this study on awake cats, we found that the A1 neurons not only responded to these specific transient segments (phasic responses) but also to the remaining portion of the vocalizations (tonic responses). Our results are consistent with other recent studies of awake animals, which revealed both phasic and sustained responses to continuous acoustic stimuli (Chimowitz et al. 2002; Lu et al. 2001; Malone et al. 2002, 2007; Petkov et al. 2007; Recanzone 2000; Wang et al. 2005).

A1 neurons formed a continuum of temporal response patterns, from responding only to stimulus onset/offset to responding throughout the stimulus period (Fig. 2), suggesting that the phasic and tonic response groups are not completely

**Effects of the temporal order of vocalizations**

By comparing the neural responses to natural and time-reversed vocalizations, we could know the effect of the temporal order of vocalizations because the two stimuli have the same spectral content but reversed temporal contents. The neurons responding to natural meows in phasic patterns responded to reversed meows also in phasic patterns (red lines in the 1st and 2nd rows of Fig. 6). The phasic-tonic response patterns were also not changed by the temporal reversal of stimulus (blue lines in Fig. 6). However, the PSTHs of natural meows were not completely symmetrical to those of the reversed meows, especially in the case of meows 4 and 5. The instantaneous firing rate responding to some segments of a stimulus was different between the natural and time-reversed meows. This is the case for both phasic and tonic responsive cells. Thus although A1 may work as a processor to analyze the acoustic features (amplitude envelope or frequency difference) within vocalizations, the analysis is affected by the sequential context of vocalizations.

![Graph](http://jn.physiology.org/)

**FIG. 9.** Relation between response duration and characteristic frequency (CF). A: duration of neural response to meow 1 vs. the cell’s CF. Horizontal dotted line indicates the criterion level for dividing phasic and phasic-tonic cells. B: response durations of all the 15 vocalizations (normalized by the duration of vocalization) vs. CF. C: energy distribution of our vocalization exemplars. Solid line indicates the mean SPL of the 15 vocalizations at each frequency component. Dotted line indicates the mean driven rate.)
separate in A1. However, most A1 neurons responded to the different vocalizations in a cell’s specific temporal response pattern (Figs. 3 and 6), suggesting that different neuronal mechanisms underlie the different response patterns.

We found a functional difference between the phasic and tonic responses. Phasic responses were sensitive to the amplitude envelope at sound on- and offset rather than to the spectral content in spectral contents (Figs. 7 and 8). On the other hand, the tonic response, which occurred at 100 ms after stimulus onset, was sensitive to the spectral difference of stimuli rather than to the amplitude envelope (Fig. 8, E and F). This is consistent with our previous findings that the tonic responses are sensitive to spectral information such as the spectral envelope (Qin et al. 2004a), spectral edge (Qin et al. 2004b), and the fundamental frequency of harmonics (Qin et al. 2005). Thus different acoustic features (amplitude envelopes versus spectral contents) within a sound signal are processed by A1 neurons in different discharge manners (phasic vs. tonic discharges).

We also found a difference between the CF distributions of the phasic and tonic response types (Fig. 9). The cells with tonic responses to vocalizations usually had CFs within the spectral range of vocalizations. It is therefore reasonable to hypothesize that the tonic responders are sensitive to the spectral information of vocalizations. On the other hand, the phasic responders were distributed throughout the A1 area tested (CF <16 kHz). That is, the cells with CFs that were largely outside the vocalization spectra could also respond to the stimuli, suggesting that the cells have a broad spectral receptive field to cover the stimulus spectra. This is consistent with our previous finding that the phasic responses have a broader spectral receptive field than the tonic responses (Qin et al. 2003).

Does A1 encode the behavioral referents?

Previous studies used the vocalizations of non-human primates, in which the behavioral and social contexts were well defined by the behavioral observations, to investigate neural responses involved with behavioral referents of species-specific vocalizations (Cohen et al. 2007; Gifford et al. 2003; Romanski et al. 2005). However, in the case of cat meows, the behavioral and social contexts have not yet been clearly defined. To overcome this shortcoming, we used three different categories of vocalizations, i.e., natural meows, time-reversed meows, and human vowels. If the conspecific vocalizations (meows) evoked more vigorous neural responses than ethologically irrelevant foils (time-reversed meows and human vowels), it would be strongly suggested that A1 neurons are sensitive to the behavioral referents of cat meows. However, the neural responses to the different categories of stimuli were similar in many aspects, including the percentage of responsive neurons (Fig. 4B), discharge rate (Fig. 4C), and response duration (Fig. 2). The similarities among responses to different vocalization types could not be attributable to poor sound stimulus quality because our stimulus vocalizations were recorded at 100 kHz/s with a cutoff frequency of 20 kHz, which ensures that most acoustic features were recorded. Furthermore, correlation analyses have shown that the stimuli were characterized by different acoustic features (amplitude envelope and frequency content) and evoked different responses in A1 neurons (Figs. 7F and 8D). The similarity of overall responses suggested that the natural meows could not evoke more vigorous neural responses than time-reversed meows and human vowels. This negative result supports the conclusion that the A1 neurons do not encode the behavioral referents of meows as also shown in other studies on A1 (Glass and Wollberg 1983b; Nagarajan et al. 2002; Schnupp et al. 2006).

Before an overly general conclusion is drawn, however, several limits of our current experiments should be mentioned. One is that we only used a small set of vocalizations to stimulate A1. It could not be excluded that A1 neurons may prefer some specific meows not been tested in this study. Furthermore, the cats were passively listening to the vocalizations during the experiments. We did not require the cats to make any behavior reflecting the behavioral relevance of vocalizations. It was also unclear whether our exemplars have significant behavioral relevance for all the tested cats.

Both the phasic and tonic responses contribute to vocalization discrimination

Because the neural responses under anesthesia were dominated by the phasic responses locked to the onset, offset, or envelope peaks of vocalizations (Gehr et al. 2000; Gourevitch and Eggermont 2007; Rotman et al. 2001; Schnupp et al. 2006; Sovijärvi 1975; Wang et al. 1995), the vocalization discrimination of A1 was previously considered to depend on the transient responses evoked by abrupt changes in the sound signal. Relative to the long duration of vocalizations such as a cat’s meow, the response duration of an individual neuron was quite short, which limited the information carried by a single neuron’s activity. Hence A1 neurons were suggested to discriminate vocalizations by using multiple parameters of spike activities such as the firing rate, spike timing, and spike synchronization (deCharms and Merzenich 1996; Furukawa and Middlebrooks 2002; Gehr et al. 2000; Gourevitch and Eggermont 2007; Rotman et al. 2001; Schnupp et al. 2006; Sovijärvi 1975; Wang et al. 1995). Nevertheless, the sparse spikes observed in the A1 neurons of anesthetized animals limited the amount of information conveyed.

In this study of awake cats, we found that A1 neurons could use multiple response patterns to discriminate vocalizations based on their acoustic features. A1 neurons showed phasic responses soon after the on- and offset of stimuli, which could be used to detect rapid changes in the amplitude envelope within a short time interval. The amplitude changes occurring at different times after stimulus onset were detected by different cells. The phasic responses of phasic cells were sensitive to the amplitude envelope during the first 5 ms after stimulus onset, and the phasic responses of phasic- tonic cells were sensitive to the amplitude during the next 5–20 ms (Fig. 7F). Thus this population could discriminate both the early and late amplitude signals at sound onset.

A cat’s meows and human vowels are usually characterized by a long mid-portion with few rapid amplitude changes. During this period, some A1 neurons exhibited tonic responses, which are sensitive to the spectral contents rather than to the amplitude envelope (Fig. 8D). Although the percentage of neurons responding to the mid-portion of vocalizations was smaller than that of neurons responding to the onset (Fig. 4B), the presence of tonic responses filled in the gaps between the discrete phasic responses. The tonic response connecting the
on- and offset of a meow may play a role in recognizing whether a meow is continuous or not.

A recent study on the human auditory cortex (Hewson-Stoate et al. 2006) found that the main difference between the vowel- and noise-evoked electroencephalography (EEG) signals was a large, vertex-negative, sustained response to vowels. The vowel-specific sustained response began with the time window of P1 deflection in the transient response at sound onset (30–50 ms after sound onset) and lasted throughout stimulus presentation. In agreement with this, our results in cat also showed that spectral sensitive neural responses began around 50 ms and lasted throughout stimulus presentation (Fig. 8D). Therefore the processing of the spectral components of long communication sounds such as cat meows and human vowels may require sustained brain activity in both cats and humans, which may correspond to the long timescale (200–300 ms) cortical processing suggested by fMRI studies on human auditory cortices (Boemio et al. 2005).

Based on the results from present and previous studies, we propose that vocalizations are represented in cat A1 in two ways. One, as suggested by many previous studies, is through synchronized responses that encode the occurrence of a sharp change in the amplitude envelope. The other is through tonic responses that encode static acoustic features such as the frequency content. The combination of the two contributes to the representation of vocalizations throughout their durations, especially those lacking dynamic changes in the amplitude envelope.

Effect of the temporal order

It should be noted that although the natural and reversed meows had temporally symmetrical spectrograms, their neural responses were not completely symmetrical (Fig. 6). Previous studies on anesthetized cats also reported differences in the 50-ms bin responses to natural and reversed meows (Gehr et al. 2000; Gourevitch and Eggermont 2007). These results suggest that A1 neurons have a memory for past acoustic signals, something that affects the response to current signals. The sequential context of the stimulus has been shown to modulate the activity of A1 neurons in various ways (Bartlett and Wang 2005). The responses to current signals could be increased or decreased according to the spectral, temporal, and intensity properties of the proceeding signals. The complicated responses to the stimulus context may result in failure to predict complex sound responses by the use of pure tone tuning profiles (Bar-Yosef et al. 2002; Rotman et al. 2001). Thus although A1 neurons abstract the acoustic features within the vocalizations, the abstraction is not a simple time-invariant process.

Acknowledgments

The helpful comments of the anonymous reviewers are greatly acknowledged. We thank N. Yaguchi for technical assistance.

Grants

This study was supported by Grant-in-Aid for Young Scientists Grant 18890076 from the Japan Society for the Promotion of Sciences, National Nature Science Foundation of China Grant 30700938 to L. Qin, and a Grant-in-Aid for Scientific Research from University of Yamanashi to L. Qin and Y. Sato.

References


