Behavioral/Systems/Cognitive

Neural and Behavioral Discrimination of Sound Duration by Cats

Ling Qin,[1,2] **Yongchun Liu,**[1] **JingYu Wang,**[1] **ShuNan Li,**[1] **and Yu Sato**[1]

[1]Department of Physiology, Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Chuo, Yamanashi 409-3898, Japan, and [2]Department of Physiology, China Medical University, Shenyang, 110001, China

Sound duration is important for distinguishing auditory object. Previous studies on the neural representation of duration have usually lacked psychophysical data obtained from the same species; hence, the correspondence between neural and behavioral discrimination of duration remains obscure. We addressed this issue in cats by using the signal detection theory to investigate both neural activities in the primary auditory cortex (A1) and the cat's behavioral performance. We found that 320 ms duration can be well discriminated from 10 ms duration by some A1 neurons with specific response patterns: sustained response extended proportionally with the increase of stimulus duration and the On–Off response synchronizing stimulus onset and offset. Neurons with only On response cannot discriminate duration. The discrimination performance of both sustained and On–Off responses deteriorated as the target duration decreased from 320 to 20 ms and the percentage of discriminative neurons (correct rate >0.75) decreased from 40 to 2%. Compared with the psychophysical results, we found that the psychometric functions of cats well matched the neurometric functions of most sustained-response neurons and a small number of On–Off-response neurons. Pooling the spikes of multiple units improved neural discrimination, which may be attributable to the salience (noise reduction) of the responses in pooled data. Our results suggest that the sustained and Off responses of A1 neurons underlie the duration discrimination behavior of cats.

Introduction

Sound duration is an important temporal feature of acoustic information. For example, the duration of fricative noise is the cue for fricative–affricate distinction (Repp et al., 1978; Mitani et al., 2006). For this reason, the neural substrate for processing sound duration has been intensively investigated by previous electrophysiological studies. It is well established that the neurons in the low level of auditory pathway can potentially encode sound duration through their temporal response patterns, such as responses at stimulus onset and offset, or through tonic responses over the duration of the sound (Kitzes et al., 1978; Calford and Webster, 1981; Rhode and Smith, 1986; Grothe, 1994). Neurons that selectively respond to a narrow range of durations were only found at or above the level of the inferior colliculus (IC) or its homolog. Duration-selective neurons have been described in the central auditory system of many vertebrates, including frogs (Potter, 1965; Narins and Capranica, 1980; Hall and Feng, 1986; Gooler and Feng, 1992), bats (Pollak and Schuller, 1981; Pinheiro et al., 1991; Casseday et al., 1994; Fuzessery, 1994; Ehrlich et al., 1997; Galazyuk and Feng, 1997; Faure et al., 2003; Mora and Kössl, 2004), cats (He et al., 1997), chinchillas (Chen, 1998), mice (Brand et al., 2000), and rats (Pérez-González et al., 2006). Al-

though a large volume of literature has been accumulated, the cortical substrate for duration representation has seldom been reported, and most previous studies were conducted on anesthetized animals. Hence, it is worth investigating duration representations in the auditory cortex of awake animals.

More importantly, the previous neuronal data of duration representation frequently lacked corresponding psychological data about the behavioral discrimination of animals. There is still no evidence demonstrating whether the observed neuronal activities really account for different duration perceptions. To address this issue, we recorded neural responses to tones of various durations in the primary auditory cortex (A1) of awake cats and also measured the behavior performance of cats discriminating these tones. Although the physiological and psychological data were obtained from separate cats, our results show for the first time excellent correspondence between the neural representation and discrimination behavior of sound durations in the same species.

Materials and Methods

Animal preparation, recording, and histology. Experiments were performed in accordance with the University of Yamanashi *Guidelines for Animal Experiments* 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 as described previously in detail (Qin et al., 2005, 2007, 2008a,b). For microelectrode access in the chronic recording, an aluminum cylinder (inner diameter, 12 mm) was implanted into the temporal bone above the auditory cortex, which was identified using a stereotaxic atlas of the cat brain. Extracellular recording was conducted using an epoxylite-insulated tungsten microelectrode (impedance, 2–5 MΩ at 1 kHz; FHC), which penetrated the brain through a micromanipulator (MO-95; Nar-
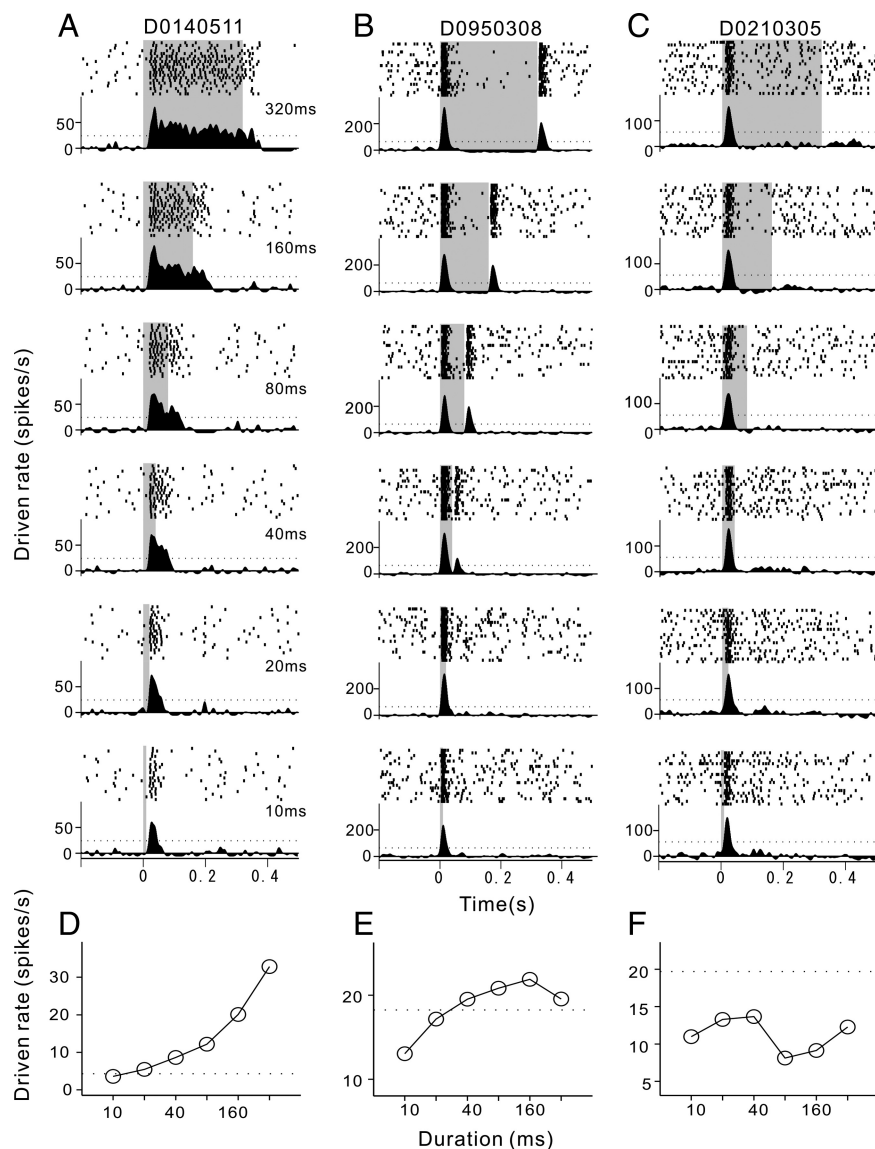
Figure 1. Discharge patterns and duration tuning functions of the three representative units. *A–C*, Raster plots and mean PSTHs of 20 trials in response to tone bursts of different durations (10, 20, 40, 80, 160, and 320 ms). Shaded areas indicate stimulus duration. *D–F*, Mean driven rate as a function of stimulus duration. Dotted line shows the level of 2 SDs of the background firing rate.

ishige) mounted on the implanted cylinder. The coordinate of each track was recorded to construct a map. At the end of the experiment, some tracks were reapproached and marked by electrolytic lesions. The animal was then deeply anesthetized with sodium pentobarbital and perfused with 10% Formalin. The brain was removed, cut into coronal sections, and stained with neutral red. Based on the relative coordinates between the electrolytic lesions and other tracks, the locations of recording sites in the brain were reconstructed. This report was based on units collected from A1, which was determined by the anatomic location and tonotopic organization.

Sound generation and delivery. The sound delivery system was controlled by a custom-made program written in MATLAB (MathWorks). Digitally generated waveforms of sound stimuli were fed into a 16-bit digital-to-analog converter (PCI-6052E; National Instruments) at a sampling frequency of 100 kHz and to an eight-pole Chebyshev filter (P-86; NF Electric Instruments) with a high cutoff frequency of 20 kHz. The outputs were played through a pair of speakers (K1000; AKG) placed 2 cm from each auricle of the cat. We calibrated the sound delivery system between 128 and 16,000 Hz at the frequency step of 8 Hz, and the output varied by ±5 dB. Harmonic distortion was less than −60 dB.

Electrophysiological recording and data analysis. Single-unit activities were discriminated using a template-matching discriminator (ASD; Alpha Omega Engineering). Once a single unit was isolated, we estimated its frequency tuning by randomly presenting a sequence of pure tones (500 ms in duration including 5 ms rise/fall time) in the frequency range of 128–16,000 Hz (in 128 Hz steps) and at sound pressure level (SPL) of 50 dB. The best frequency (BF) was defined as the tone frequency that elicited the largest number of spikes during the 500 ms stimulus period. We then evaluated the effect of sound duration on neuronal responses by presenting BF tones with six different durations (10, 20, 40, 80, 160, and 320 ms) including 5 ms rise/fall time. Each duration was repeated 20 times; in total, 120 tone bursts were presented in pseudorandom order with an interstimulus interval of at least 1 s. If single-unit activity was still well isolated after completing the recording at 50 dB SPL, the same stimulus procedure was repeated at 70 and/or 30 dB.

The peristimulus time histogram (PSTH) in 1 ms bin width was constructed by counting the number of spikes in each trial. Neural discharge was sometimes suppressed by a tone stimulus to lower than the spontaneous firing rate (Fig. 1*B*). To represent both excitatory and suppressive activities, the height of PSTH was transformed into the "driven rate" by subtracting the background firing rate (firing rate averaged over 120 trials during the 0.5 s period before each stimulus onset). PSTH was then smoothed by a Gaussian function with 5 ms SD. To assess the temporal discharge pattern of the neuron, mean PSTH averaged over 20 trials of the same duration stimuli was also constructed (Fig. 1*A–C*, bottom). Based on the mean PSTH of 320 ms tones, we classified the neuronal responses into three patterns: (1) "sustained response," in which the duration of excitatory responses (bins of PSTH higher than 2 SDs of the background firing rate) during the 320 ms stimulus period was longer than 160 ms (Fig. 1*A*); (2) "On–Off response," in which the duration of excitatory responses during the stimulus period was shorter than 160 ms, and excitatory responses were also found within the 100 ms period after stimulus offset (Fig. 1*B*); and (3) "On response," in which excitatory responses were only found during the stimulus period, and the response duration was shorter than 160 ms (Fig. 1*C*).

For each neuron, we evaluated the change in neural response as a function of the duration of the tone burst. For this, the duration tuning function of the unit was constructed by plotting its mean driven rate, averaged over the 400 ms period after stimulus onset as a function of stimulus duration (Fig. 1*D–F*). The 2 SDs level of the background firing rate averaged over the same length of time window was also plotted to evaluate the reliability of the tuning function (Fig. 1*D–F*, dotted line).

Neurometric function analysis. The receiver-operating characteristic (ROC) based on signal detection theory (Green and Swets, 1966) has been used successfully to examine whether neural responses can provide sufficient information to discriminate different stimuli (Britten et al., 1992, 1996; Parker and Newsome, 1998). Here, we applied PSTH-based ROC analysis following Walker's method (Walker et al., 2008). The difference between two individual PSTHs (1 ms bin, 400 ms analysis window, normalized by the maximum response height of the

cell) was quantified by the Euclidian distance (ED), calculated as ED $= \sqrt{\sum_{i=1}^{n_{bin}}(x_i - y_i)^2}$, where n_{bin} is the total number of bins, and x and y are bin heights.

For each unit, ED between two single-trial PSTHs of the same 10 ms tone was referred to as a within-category distance, whereas that between PSTHs of 10 ms and one of the other durations (20, 40, 80, 160, or 320 ms) was referred to as an across-category distance. An ROC curve (Fig. 2A) for a given duration (20– 320 ms) was constructed by comparing the distributions of within- and across-category distances in a series of 40 criterion values spanning the full range of both distributions; that is, for each criterion, the proportion of across-category distances that exceeded the criterion (hit rate) was plotted against the proportion of within-category distance that exceeded the criterion (false alarm rate). If the two distributions were identical, this curve would trace the line of identity on unit axes (Fig. 2A, diagonal), and the two stimuli would be indistinguishable. If the distribution of across-category distance has a median value greater than that of the within-category distance, these points will trace a curve above the diagonal. Integrating the area under the ROC curve by the trapezoid rule yields a measure of discrimination performance (expected proportion of correct responses) that is independent of a potential criterion bias. The area under the ROC corresponds to the correct rate of cat's behavioral responses in the psychophysical experiments, which is calculated as p(correct) $= [p$(hits) $+ (1 - p$(false alarms))]/2. The area under the ROC was plotted against the duration of target tone to construct the neurometric function (Fig. 2D).

Psychophysical testing procedure. Behavioral performance was measured in three male cats different from those used for electrophysiological recording. The cats were trained using a go/no-go procedure to discriminate whether a target tone was different from a standard tone. Initially, the cats were deprived of food to 80% of their free-feeding weight. Subjects responded in a custom-built behavioral box in a soundproofed room. The behavioral apparatus and sound generation were controlled and monitored by software developed in MATLAB, running on a computer interfaced with digital input–output hardware (PCI-6052E; National Instruments). The auditory stimuli were delivered via a pair of calibrated speakers (K1000; AKG) placed 2 cm outside the grid walls of the behavior box. A video camera and photoelectric sensors were used to monitor the cat's position and movements. The cats were first trained to lick a metal pipe on sound presentation to obtain a drop of liquid food. Each trial was initialized only when the cats stood ready and kept their head in the observing position in front of the metal pipe, where the speaker calibration was performed. Cats learned this procedure within 1 week and were then trained to discriminate the duration difference between two successively presented tones. The first tone was always a standard tone (10 ms in duration), whereas the second tone was presented 250 ms later and was either a standard or target tone (320 ms). Subjects were required to lick the metal pipe when the target signal was presented ("hit") and not to lick when the standard signal was heard ("correct rejection"). There were also two kinds of error responses: licking on the presentation of the standard signal ("false alarm") and not licking on the presentation of the target signal ("miss"). Subjects were positively reinforced only for the hit response. False alarm response resulted in a timeout, during which the lights were extinguished and the training program paused for a period of 10–30 s.
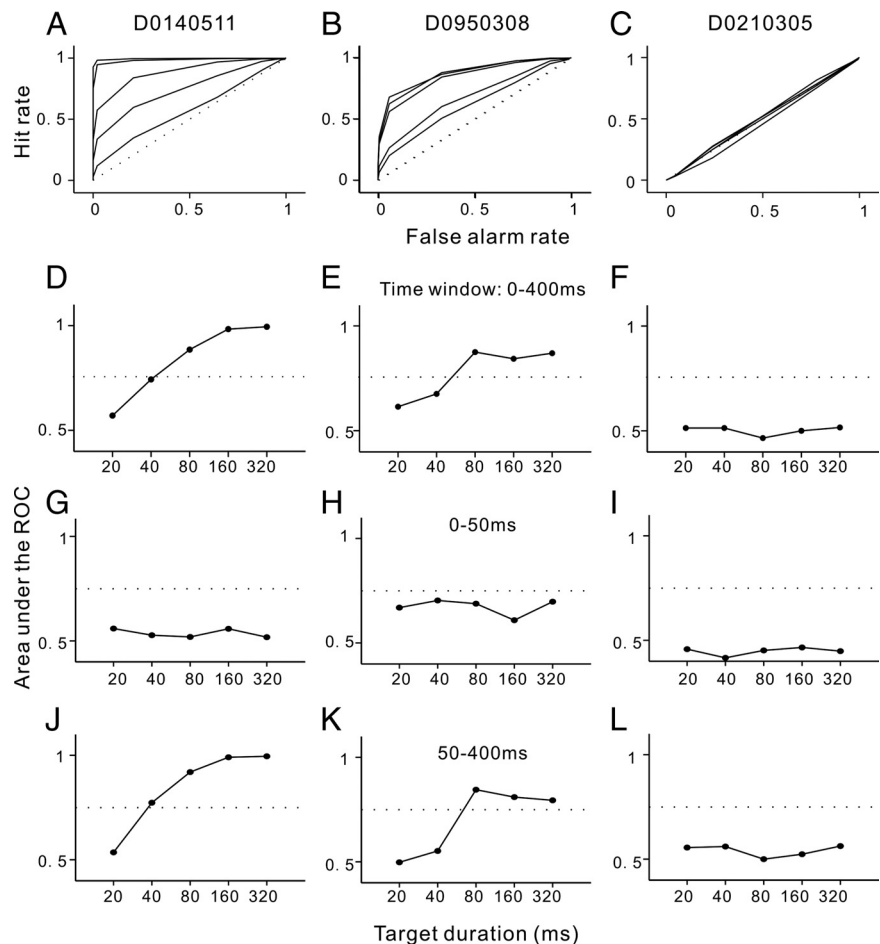

Figure 2. Results of ROC analysis in three representative units. *A–C*, ROC curves constructed from the two distributions of ED: the ED between two single-trial PSTHs, both from the standard duration (10 ms) group and the ED between two PSTHs of the standard and target duration (20, 40, 80, 160, or 320 ms), respectively. *D–F*, Neurometric functions constructed by plotting the area under the ROC curves shown in *A–C* against the target duration. Dotted line shows the level of 0.75. Time window used in ROC analysis is 0 – 400 ms after stimulus onset. *G–I*, Neurometric functions obtained from the time window of 0 –50 ms. *J–L*, Neurometric functions obtained from the time window of 50 – 400 ms.

The target and standard signals were randomly presented in an equal ratio. One session consisted of 100 trials. Cats could usually finish two to four training sessions every day.

The subjects' daily performance in this task was quantified via the measure d' from signal detection theory (Green and Swets, 1966; Macmillan and Creelman, 1991). d' is calculated as the z-score of the probability of a false alarm response subtracted from the z-score of the probability of a hit response. The cats required 1–2 months of training (5000–6000 trials) to achieve performance levels of $d' \geq 2.0$ (~84% correct). The threshold of d' is usually defined as 1.0 in visual studies. In this auditory task, the performance of cats could not be maintained in successive sessions if the training procedure stopped at $d' > 1.0$. We therefore adopted a high criterion of $d' = 2.0$, as used by Otto et al. (2005). Once cats reached the performance of a $d' \geq 2.0$ for five successive sessions, they advanced to additional steps, in which the duration of the target tone was decreased in a sequence of 160, 80, 40, and 20 ms. This procedure ceased at the step when the subject's performance could not reach the criterion after 2000 training trials. The frequency and/or SPL of tone bursts were then changed, and the training procedure was restarted from 320 ms duration.

After completing the training paradigm, we evaluated the cat's discrimination performance by randomly presenting various pairs of durations in one session, which consisted of 50 trials of standard comparison (10 –10 ms) and 10 trials of each target comparison (10 –320, 10 –160, 10 –80, 10 –40, and 10 –20 ms), respectively. Each cat performed 12 test-
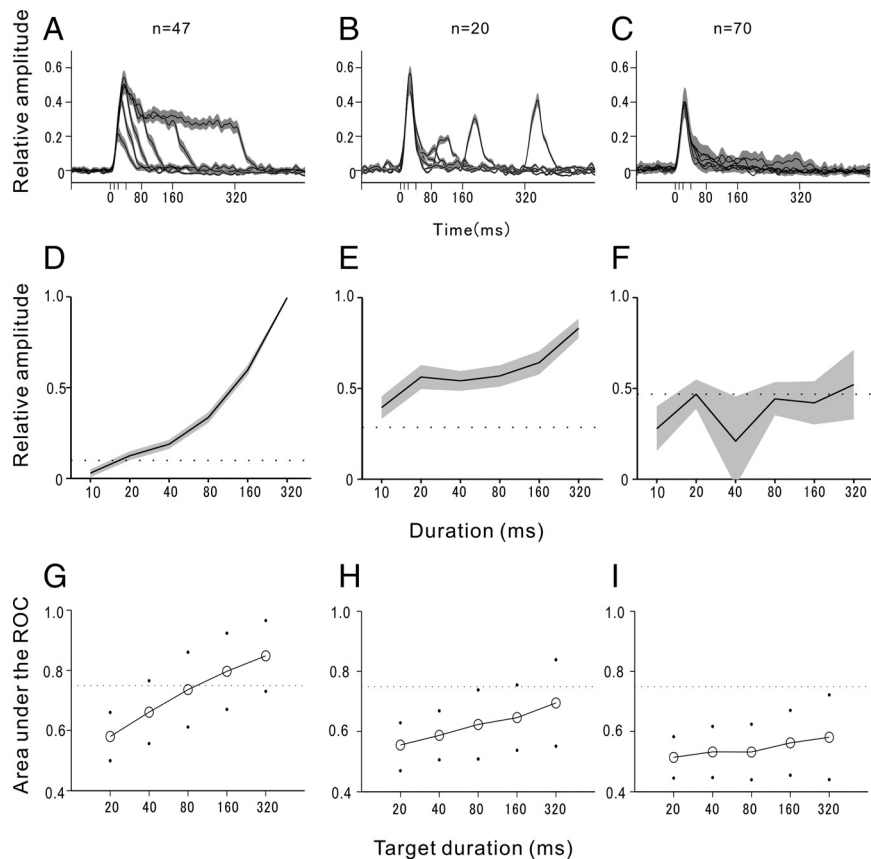
Figure 3. Population properties of the three groups of A1 neurons. *A–C*, Mean PSTHs of sustained-, On–Off-, and On-response units, respectively. PSTHs of 10, 20, 40, 80, 160, and 320 ms are superimposed to show the effect of stimulus duration on each unit group. Shaded area shows SE. *D–F*, Mean duration tuning function averaged over each group. Shaded area shows SE. Dotted line is the mean level of 2 SDs of background firing. *G–I*, Mean neurometric function averaged over each group. Dots show SD. Horizontal line is the level of 0.75.

ing sessions. Each session tested one combination of four frequencies (2, 4, 8, or 16 kHz) and three SPLs (30, 50, or 70 dB). The psychophysical function of duration discrimination was constructed by plotting the area under ROC against the duration of target tone.

Results

In total, data from 134 single units were collected from both hemispheres of the caudal part of A1 in three awake cats. The A1 was distinguished from the adjacent auditory fields by the tonotopic organization of BF. The BF of the collected units ranged from 320 to 15,872 Hz (mean ± SD, 6.7 ± 4.9 kHz).

Representative examples of different response patterns

A1 units of awake cats showed multiple discharge patterns to tone bursts. Figure 1*A* presents an example unit with a sustained discharge pattern. Raster plots of spike activities responding to tone stimuli with different durations (10, 20, 40, 80, 160, and 320 ms) at 50 dB SPL are shown at the top of each plot. The mean PSTHs averaged over the 20 trials of each stimulus condition are shown at the bottom. The dotted line displays the level of 2 SDs of the background firing rate, which is used as the threshold for detecting the excitatory response. In this example cell, the duration of excitatory responses to 320 ms tones was 291 ms, which was obtained by counting the over threshold bins of the mean PSTH during the stimulus period (Fig. 1, shaded area). The long duration of response to 320 ms tone characterizes the sustained response pattern of this cell. By shortening the stimulus duration,

the response duration is also proportionally shortened. Consequently, this cell showed a long-pass type of duration tuning (Fig. 1*D*), evaluated by the mean driven rate during the 400 ms period after stimulus onset.

Figure 1*B* shows an example unit with a transient discharge pattern synchronizing to the onset and offset of stimuli (On–Off response). This cell showed a sharp response to the onset of the 320 ms tone, which was characterized by the short response duration of 33 ms. Additionally, the 320 ms tone also evoked an excitatory response after sound offset. Shortening the stimulus duration had no obvious effect on the On response, whereas the occurrence time of the Off response shifted ahead progressively; however, when the stimulus duration was 40 ms, the Off response obviously decreased and partly overlapped with the On response. The Off response became undistinguishable when the stimulus duration was 20 or 10 ms. The duration tuning function of this cell also appears to be a long-pass type (Fig. 1*E*) but, because the cell had obvious spontaneous firing and the excitatory responses were limited within short periods, the tuning function based on the mean driven rate cannot obviously surpass the level of 2 SDs of the background firing rate (dotted line), and the tuning slope of this cell is also slower than that of the sustained-response cell (Fig. 1*D*).

Figure 1*C* shows an example unit with only a transient response to the stimulus onset (On response). The response duration during the 320 ms tone was only 31 ms. Varying the tone duration between 320 and 10 ms had little effect on the response of the cell. The duration tuning function evaluated by the 400 ms time window was lower than the 2 SDs of the background firing rate (Fig. 1*F*), suggesting that the tuning function cannot effectively access the characteristics of this cell.

As indicated by Figure 1, A1 neurons of awake cats showed obvious spontaneous firing and multiple response patterns, making it difficult to use a uniform rate function to fully assess their response properties. We therefore applied PSTH-based ROC analysis to examine whether the spike information can be used to discriminate different tone durations (see Materials and Methods). For each neuron, we used spike trains of 10 ms duration tones as the standard signals and those of other durations as the target signals. Using analyses derived from signal detection theory, ROC curves were calculated for each standard-target comparison: 10–320, 10–160, 10–80, 10–40, and 10–20 ms (Fig. 2*A*), and a neurometric discrimination curve was then derived from these ROC curves (Fig. 2*D*) (see Materials and Methods). From Figure 2*D*, it is clear that the sustained-response cell (the same cell shown in Fig. 1*A*) can perfectly discriminate a 320 or 160 ms tone from a 10 ms tone (the area under the ROC is close to 1.0). The discriminating performance gradually decreases as the target duration decreases from 160 to 20 ms and drops below the level of 0.75 (dotted line) at 40 ms.

The results of ROC analysis of the On–Off-response cell are displayed in Figure 2, *B* and *E*. Although the neurometric function of this cell is much lower than that of the sustained-response cell, it can still well discriminate 320, 160, and 80 ms tones from 10 ms tone (the area under the ROC is higher than 0.75). The lower discriminating performance at 40 and 20 ms is consistent with the fact that the Off response of this cell became smaller at 40 ms and disappeared at 20 ms (Fig. 1*B*); thus, this cell may depend on the Off response to discriminate different durations.

As shown in Figure 2, *C* and *F*, the results of ROC analysis indicated that the On-response cell cannot discriminate duration. We then examined whether the discriminating performance depends on the analysis time window. For this, we reconducted the ROC analysis using two different time windows: 0–50 and 50–400 ms, respectively. All the neurometric functions obtained from the 0–50 ms time window are below 0.75 (Fig. 2*G*–*I*), suggesting that On responses contribute less to duration discrimination. However, ROC analysis using the 50–400 ms time window resulted in a similar neurometric function (Fig. 2*J*–*L*) as that obtained from the 0–400 ms window (Fig. 2*D*–*F*). Thus, the sustained-response and On–Off-response cells depend mainly on later responses to discriminate duration. The similarity between the neurometric functions of 50–400 and 0–400 ms also indicates that removing the less useful On responses cannot obviously improve neurometric function. This is reasonable because our ROC analysis is based on the ED between two spike trains. On responses during the 0–50 ms period show high similarity, resulting in a small ED, therefore contributing less to the neurometric function of 0–400 ms.

Population data of different response patterns
Above we presented three representative examples with distinct response patterns. Accordingly, we classified the 134 units into three groups: 47 sustained-response cells, 20 On–Off-response cells, and 70 On-response cells (for the definition, see Materials and Methods). The effect of duration on the different discharge patterns is illustrated by the mean PSTHs averaged over each unit group (Fig. 3*A*–*C*). In the plots, the PSTHs of various stimulus durations are aligned relative to stimulus onset. The population results showed similar properties to those of the representative examples: the duration of sustained responses proportionally changed with the stimulus duration (Fig. 3*A*); the occurrence time of Off responses closely followed the termination of stimulus (Fig. 3*B*); and the On responses were less affected by the change in duration (Fig. 3*C*). The duration tuning function of each cell was also normalized by its maximum and averaged over each group to construct the mean duration tuning function (Fig.
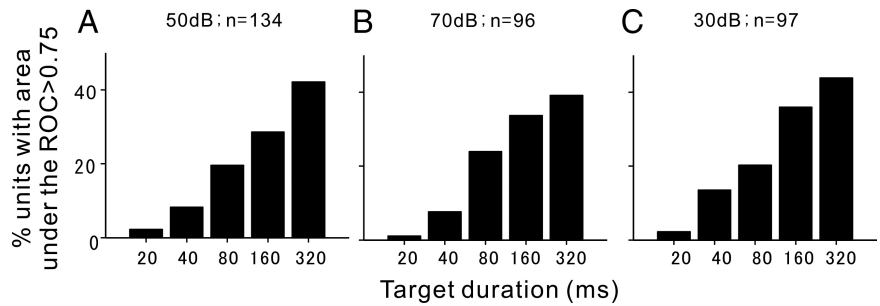

Figure 4. Histograms showing the percentage of units achieving a discrimination performance higher than 0.75, when spike activities are used to discriminate 10 ms tones from 20, 40, 80, 160, and 320 ms tones, respectively.
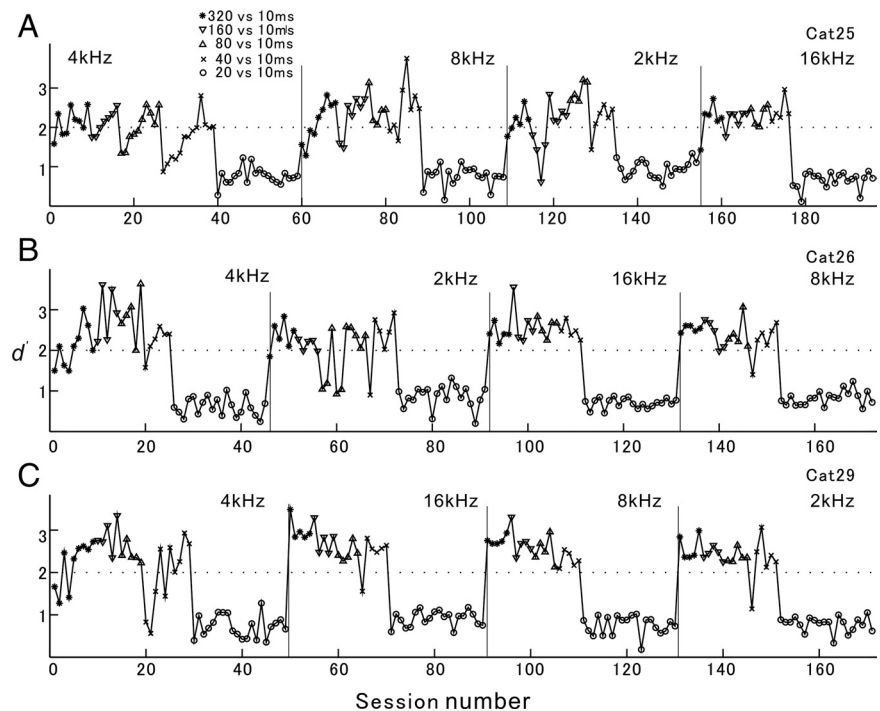

Figure 5. Behavioral performance of three cats during training sessions to discriminate 20–320 ms tones from 10 ms tones. Different marks represent different target durations. Horizontal line shows the criterion ($d' = 2$) of behavioral performance. Vertical lines mark the session when the tone frequency was changed.

3*D*–*F*). Repeating the results of the representative cells, the sustained-response group showed sharp long-pass-type duration tuning (Fig. 3*D*), the On–Off-response group showed slow long-pass tuning (Fig. 3*E*), and the On-response group showed unstable duration tuning (Fig. 3*F*), which was mostly below the mean level of 2 SDs of the background firing rate (dotted line).

The mean and SD of neurometric functions in each unit group are illustrated in Figure 3*G*–*I*, respectively. The 320 ms duration can easily be discriminated from 10 ms duration based on the sustained responses, and the performance gradually decreases with the decrease of target duration (Fig. 3*G*). The population result of On–Off responses is similar in style to that of sustained responses, except that the performance is lower in all discrimination pairs (Fig. 3*H*); thus, different durations can also be discriminated based on On–Off responses, but the reliability is worse. Conversely, the duration can hardly be discriminated based on the On responses alone (Fig. 3*I*), because all the values of the mean and SD are lower than 0.75 (dotted line). The population results collected at tone SPLs of 70 and 30 dB also showed the same tendency (data not illustrated).
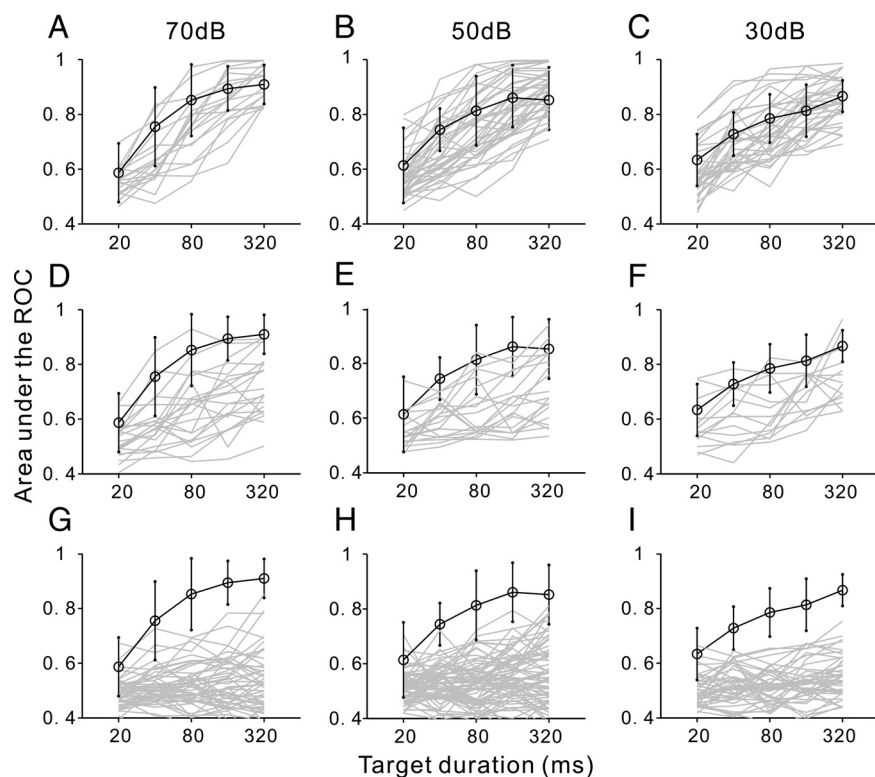
Figure 6. Comparison of psychometric and neurometric functions. **A–C**, Area under the ROC was plotted against the duration of target tone, constructing the neurometric and psychometric functions. Gray lines show the individual neurometrics of sustained-response units. Circles and bars show the mean and SD of the psychometrics, respectively. **D–F**, Neurometrics of On–Off-response units are presented with the psychometrics. **G–I**, Neurometrics of On-response units are presented with the psychometrics.

The effect of the sound pressure level on duration discrimination was evaluated by counting the number of units with a neurometric score higher than 0.75. Figure 4A displays the percentage distribution of such units when tested by 50 dB SPL tones. There were 42.1% of A1 cells that could accurately discriminate 320 and 10 ms tones, whereas the percentage of cells with high performance gradually decreased with shortening of the target duration and dropped to 2.3% at 20 ms. When the sound level was increased to 70 or decreased to 30 dB SPL, the distribution pattern of high-performance units remained unchanged (Fig. 4B,C). Thus, the ability of the A1 discriminating duration may be less affected by the change in sound level.

Psychophysics

To compare the above physiological findings with the actual acoustic percepts, we attempted to train three naive cats to distinguish the duration difference between 10 and 320 ms tones using a go/no-go procedure, wherein cats were rewarded for licking a metal pipe after the presentation of a tone in 320 ms duration (see Materials and Methods). After 5000–6000 trials of behavioral training, cats were able to reliably detect the difference between 10 and 320 ms tones. We then gradually decreased the duration of the target tone from 320 to 20 ms, when the subject's performance exceeds the criterion ($d' \geq 2$) in five successive sessions (100 trials per session). The first segment in Figure 5A shows the result of one cat trained by 4 kHz tones at 50 dB SPL. After establishing reliable discrimination between 10 and 320 ms durations (asterisks), the cat only required two sessions to learn to discriminate 10 and 160 ms durations (downward triangles). As the target duration was decreased to 80 and 40 ms, increasingly more sessions of training were required for the cat to achieve the

performance criterion. Because the cat still could not continuously achieve a d' value >1.0 after it was trained for 2000 trials to distinguish 10 and 20 ms durations (white circles), we changed the tone frequency to 8 kHz and restarted the training from the initial step (10–320 ms). This procedure was repeated four times in each cat using tones of 2, 4, 8, and 16 kHz at 50 dB, respectively. The results of the other two subjects are illustrated in Figure 5, B and C. All the cats soon learned to detect the difference between 10 ms duration and 320, 160, 80, or 40 ms duration and could continuously achieve a high d' value (>2.0), but their performance of 10 and 20 ms discrimination kept fluctuating around the d' value of 1.0. We then tested the three cats using 4 kHz tones at 70 and 30 dB SPLs, and the same results were observed (data not illustrated). In total, 120 sessions (12,000 trials) of 10–20 ms discrimination tasks were conducted on each cat, none of which could stably achieved a d' value >1.0. This is consistent with the above physiological results, wherein the percentage of A1 neurons that could well discriminate 10 and 20 ms durations was quite small ($<\%5$) (Fig. 4).

To construct the psychometric function of duration discrimination, we finally evaluated the cat's behavioral performance by randomly presenting all duration pairs (10–10, 10–20 . . . and 10–320 ms) in the same session. The psychometric function was constructed by plotting the cat's behavioral performance as a function of the target duration (see Materials and Methods). The circles and bars in Figure 6A–C show the mean and SD of the psychometric functions obtained from different cats and tone frequencies at 70, 50, and 30 dB SPL, respectively. The cat's performance gradually decreases by shortening the target duration and drops lower than 0.75 at 20 ms. The psychometric functions of 70, 50, and 30 dB SPL show a similar shape, suggesting again that the duration discrimination is independent of the sound level.

Comparison between neurometric and psychometric functions

In Figure 6, we also plotted the neurometric function of each individual unit (gray lines) to compare with the psychometric functions. Most of the neurometric functions of sustained-response cells were within the 1 SD range of psychometric functions (Fig. 6A–C). In contrast, only a small number of On–Off-response cells showed neurometric functions within the 1 SD range of psychometric functions, and others showed neurometric functions parallel to but lower than the psychometric functions (Fig. 6D–F). As for On-response cells, almost all the neurometrics were not correlated with the psychometrics (Fig. 6G–I). These results suggest that sustained-response neurons contribute greatly to the behavioral performance of duration discrimination, the On–Off-response neurons contribute a little, and the On-response neurons do not.

Although some units have a lower performance individually, they may also provide more or less information for duration

discrimination if they are pooled together. To examine this possibility, we selected the 10 "worst" and "best" units (showing the lowest/highest performance at the discrimination of 320 ms duration) in each unit group, respectively. The pooled rasters and PSTHs of the 10 worst and best sustained-response units are shown in Figure 7, *A* and *B*, respectively. The population neurometric function based on the pooled spike data of 10 worst units (Fig. 7*C*, asterisk line) reaches closer to the psychometric function (circle line) than any individual neurometric functions alone (gray lines). The population neurometric of 10 best units (Fig. 7*D*) outperforms all the individual neurometrics and the mean psychometric, even reaching perfect discrimination (area under the ROC = 1) for almost all durations. The higher performance of the population data may be attributable to the better response continuity of the pooled spikes (Fig. 7*A*,*B*). Pooling the individual spike trains reduces the noise of the responses by filling more spikes (compare Fig. 7 with Fig. 1*A*), that is, the sustained-response pattern becomes more salient.

When the 10 worst On–Off-response units were pooled, the performance was only improved a little at 320 ms duration (Fig. 8*C*), in which the Off response was above the 2 SD level (Fig. 8*A*). In contrast, pooling the 10 best units substantially improved neural performance (Fig. 8*D*). This may be attributable to the salient Off responses of the pooled data (Fig. 8*B*). The population performance of the 10 best units also outperformed the psychophysical performance, suggesting that integrating the Off-response information can also sufficiently discriminate the duration. As for units with only On responses (Fig. 9), the population neurometrics of the 10 worst units does not perform better than the individual neurometrics (Fig. 9*C*); however, that of the 10 best units performs better and reaches closer to the psychometric (Fig. 9*D*). Inspecting the raster and PSTH (Fig. 9*B*), it was found that these best On-response units are actually located on the boundary between the sustained and On-response groups. The pooled data show more sustained-like responses (not higher than the 2 SDs level), which contributes to the improvement of duration discrimination. The results of pooling analysis suggest that the convergence of On responses alone with the downstream auditory centers cannot provide more information useful for duration discrimination. Only integrating the activities of sustained or Off responses can obtain additive information.

Discussion

This study, for the first time, directly compared the neural representation of A1 and behavioral discrimination of sound duration
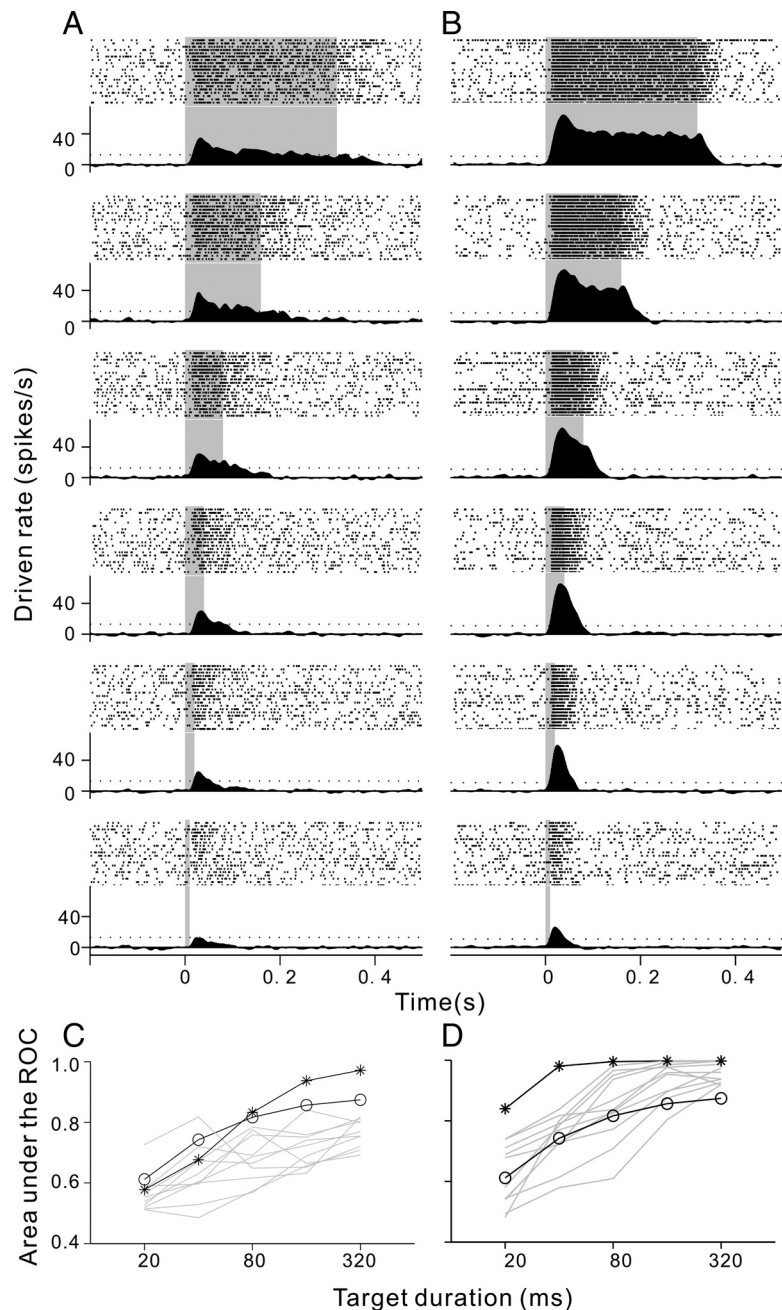

Figure 7. Effect of pooling sustained-response units. *A*, *B*, Raster plots and PSTHs of the spike data pooled by the 10 worst and best units, respectively. Same format as Figure 1. *C*, *D*, Asterisks represent the neurometric function obtained from the pooled data of the 10 worst and best units, respectively. Gray lines show the individual neurometrics of each unit. Circles represent the mean psychometric function.

in the same species. The major findings of our electrophysiological experiments are as follows: (1) the sound duration could be represented by the sustained response extending proportionally with the increase of stimulus duration and the Off response synchronizing with stimulus offset; (2) sustained responses are more effective for encoding duration than Off responses; (3) the duration representations of both sustained and Off responses deteriorate with a decreased stimulus duration, and the percentage of neurons that can discriminate 10 ms duration from longer durations decreases from 40 to 2% as the target duration decreases from 320 to 20 ms; (4) the ability of neural discrimination is relatively stable across sound levels between 30 and 70 dB SPL; and (5) pooling the spike data of individual units can improve the
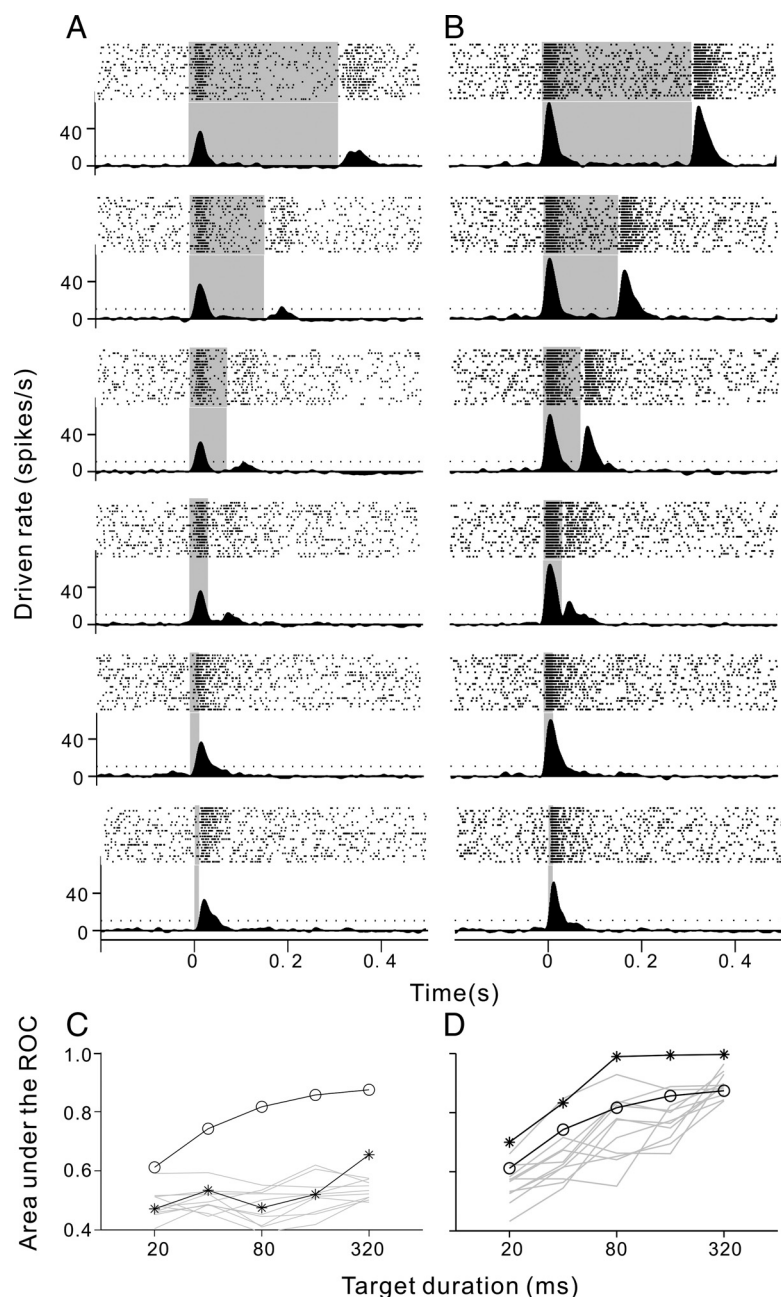
Figure 8. Effect of pooling On–Off-response units. Same format as Figure 7.

1980; Hall and Feng, 1986; Gooler and Feng, 1992). These studies found that a number of "duration-selective neurons" selectively respond to narrow ranges of duration, showing short-pass- or bandpass-type duration tuning. It was suggested that duration-selective neurons may serve to detect behaviorally relevant signals, because their duration selectivity typically approximates the duration of the voice sounds used by these species (Hall and Feng, 1986; Fuzessery, 1994; Ehrlich et al., 1997). For example, the duration of sounds used by bats during echolocation is generally <10 ms, correspondingly, ~88% A1 cells show preferences for stimuli shorter than 10 ms (Galazyuk and Feng, 1997). The percentage of duration-selective neurons varied among species. Neurons with a short-pass or bandpass duration tuning comprised approximately one-third to two-thirds of the cell population of bat IC (Ehrlich et al., 1997; Fuzessery and Hall, 1999; Faure et al., 2003; Mora and Kössl, 2004; Fremouw et al., 2005). In contrast, the total percentage of short-pass and bandpass neurons was only 17% in IC of mice (Brand et al., 2000) and <10% in ICs of rats (Pérez-González et al., 2006) and chinchillas (Chen, 1998).

Because the single-unit record cannot be maintained for long times in awake animals, we briefly tested each unit using six durations ranging from 10 to 320 ms. We used durations longer than 10 ms to test cat A1 neurons, considering that cats usually use long vocalizations to communicate. Moreover, it may be difficult for cats to perceive duration differences shorter than 10 ms, suggested by our psychophysical result that the cat's performance at 10–20 ms discrimination continued to fluctuate around $d' = 1$ even after repetitive training (Fig. 5). As the stimulus duration varied between 10 and 320 ms in this study, the sustained- and On–Off-response neurons generally showed long-pass-type duration tuning (Figs. 1, 3). Strictly speaking, they are not duration selective because they respond to varying rates across broad ranges of stimulus duration. As for neurons with only an On response, it is difficult to evaluate their duration tuning by the mean firing rate because of the effect of spontaneous firing. The absence of duration-selective neurons in our data may be also attributable to the specific choices of stimulation parameters in this study. Because we only tested six durations ranging from 10 and 320 ms, a limited number of duration samples may be unable to fully capture bandpass-type duration tuning in some units. Moreover, we used 500 ms tones to search units, which may result in a sampling bias for long-duration-sensitive cells, in turn leading to an underestimation of neural short-duration discrimination. Even so, we did find

performance of neural discrimination by increasing the salience (decreasing the noise) of sustained- and Off-response patterns.

Compared with the results of psychophysical experiments, we found that the neurometric functions of most sustained-response neurons and some On–Off-response neurons well matched the behavioral performance of cats.

Comparing the results from other species and other auditory centers

Previously, neural representation of sound duration has been well investigated in the central neural system of echolocating bats that emit biosonar pulses of specific duration (Pollak and Schuller, 1981; Pinheiro et al., 1991; Casseday et al., 1994; Fuzessery, 1994; Ehrlich et al., 1997; Galazyuk and Feng, 1997; Faure et al., 2003; Mora and Kössl, 2004) and frogs that produce communication signals of stereotypic duration (Potter, 1965; Narins and Capranica,

some neurons (sustained- and On–Off-response neurons), whose the spike activities can explain the cats' behavior of duration discrimination.

Link between neural response pattern and auditory perception

Previous electrophysiological studies have been conducted on various auditory brain areas of many animals, but to link neural activity to auditory perception, it must also be demonstrated that neural responses correlate with perceptual decisions, as assessed in psychometric experiments. Because psychometric experiments on auditory perception have seldom been conducted on animals, previous studies most often compared human psychophysical performance with neural responses in animals. For the first time, we present both physiological and psychophysical data on duration discrimination in cats and reveal that a cat's auditory perception correlates with the neural responses of A1. In particular, units with a sustained discharge pattern showed a neurometric function well matching the cats' psychometric function (Fig. 6A–C). Units with an On–Off response could also provide information on duration discrimination, although their performance was generally lower than that of sustained responders (Fig. 6D–F). Pooling the spike activities of multiple units improved the performance of neural discrimination, because the sustained- or On–Off-response pattern became more salient after pooling (Figs. 7, 8). Neural performance based on the 10 best sustained- or On–Off-response units obviously outperformed the cat's behavioral performance, even reaching perfect discrimination (Figs. 7D, 8D). This is similar to the findings in middle temporal area MT of monkey, wherein many neurons outperform the behaving animal on a visual discrimination task (Britten et al., 1992). In our data, the population activities of the 10 worst sustained-response units also provided more discriminating information than the individual activity of single units (Fig. 7C). These results suggest that some pooling or averaging is likely to occur in the neural process of sound duration, but where and how the pooling process is conducted in the brain are still open questions.

Temporal resolution of neural response

In this study, we found that cats cannot stably discriminate between 10 and 20 ms durations, even after 12,000 training trials (Fig. 5). This limitation of duration perception is coincident with several other lines of evidence suggesting that the temporal process of acoustic information is limited to time frames of 20 ms or longer. In humans, identifying a separation of two acoustic events requires stimulus onset asynchrony of ~20 ms (Pisoni, 1977; Pastore, 1983; Stevens and Weaver, 2005). Auditory cortical neu-
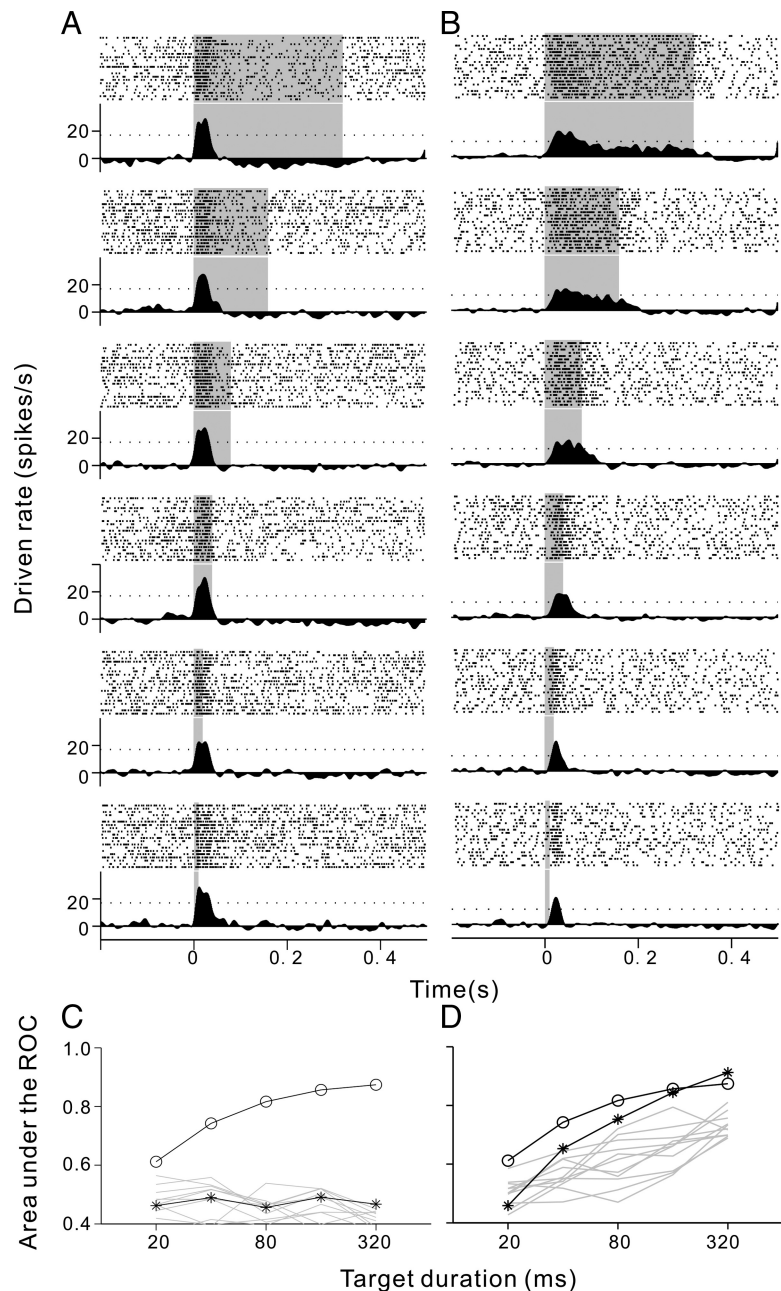

Figure 9. Effect of pooling On-response units. Same format as Figure 7.

rons were also found to have a similar limitation in representing two separated sounds (Eggermont, 1995; Steinschneider et al., 2003, 2005). Moreover, the spike timing of A1 neurons has been proved to convey sufficient information to discriminate slow changes (<50 Hz) in the envelopes of complex sounds, if the spike information is read at temporal resolutions of 20 ms or better (Walker et al., 2008).

Limitations and future directions

Our study has only taken the first steps toward understanding the relationship between neural responses and behaviors for sound duration. One of the limitations of this study is that the neural and behavioral performance measures were obtained from separate cats. Ideally, a more careful comparison between neural and behavioral performance should be performed simultaneously using the same stimuli in the same subject on a trial-by-trial basis.

This is one of the future directions of auditory experiments. Another limitation is that we only measured behavioral performance to discriminate 10 ms from five other durations (320, 160, 80, 40, and 20 ms). Training cats to perform auditory discriminations is notoriously difficult and time consuming, and the standard and target stimuli that can be used are limited to those that have been trained. For this reason, we did not measure the psychometric function in smaller steps, as done in human psychological experiments. Accordingly, the neurometric function was also calculated using limited samples. Thus, care should be taken when generalizing our findings to the untested duration range.

References

Brand A, Urban R, Grothe B (2000) Duration tuning in the mouse auditory midbrain. J Neurophysiol 84:1790–1799.

Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. J Neurosci 12:4745–4765.

Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. Vis Neurosci 13:87–100.

Calford MB, Webster WR (1981) Auditory representation within principal division of cat medial geniculate body: an electrophysiology study. J Neurophysiol 45:1013–1028.

Casseday JH, Ehrlich D, Covey E (1994) Neural tuning for sound duration: role of inhibitory mechanisms in the inferior colliculus. Science 264:847–850.

Chen GD (1998) Effects of stimulus duration on responses of neurons in the chinchilla inferior colliculus. Hear Res 122:142–150.

Eggermont JJ (1995) Representation of a voice onset time continuum in primary auditory cortex of the cat. J Acoust Soc Am 98:911–920.

Ehrlich D, Casseday JH, Covey E (1997) Neural tuning to sound duration in the inferior colliculus of the big brown bat, *Eptesicus fuscus*. J Neurophysiol 77:2360–2372.

Faure PA, Fremouw T, Casseday JH, Covey E (2003) Temporal masking reveals properties of sound-evoked inhibition in duration-tuned neurons of the inferior colliculus. J Neurosci 23:3052–3065.

Fremouw T, Faure PA, Casseday JH, Covey E (2005) Duration selectivity of neurons in the inferior colliculus of the big brown bat: tolerance to changes in sound level. J Neurophysiol 94:1869–1878.

Fuzessery ZM (1994) Response selectivity for multiple dimensions of frequency sweeps in the pallid bat inferior colliculus. J Neurophysiol 72:1061–1079.

Fuzessery ZM, Hall JC (1999) Sound duration selectivity in the pallid bat inferior colliculus. Hear Res 137:137–154.

Galazyuk AV, Feng AS (1997) Encoding of sound duration by neurons in the auditory cortex of the little brown bat, *Myotis lucifugus*. J Comp Physiol A 180:301–311.

Gooler DM, Feng AS (1992) Temporal coding in the frog auditory midbrain: the influence of duration and rise-fall time on the processing of complex amplitude-modulated stimuli. J Neurophysiol 67:1–22.

Green DM, Swets JA (1966) Signal detection theory and psychophysics. New York: Wiley.

Grothe B (1994) Interaction of excitation and inhibition in processing of pure tone and amplitude-modulated stimuli in the medial superior olive of the mustached bat. J Neurophysiol 71:706–721.

Hall J, Feng AS (1986) Neural analysis of temporally patterned sounds in the frog's thalamus: processing of pulse duration and pulse repetition rate. Neurosci Lett 63:215–220.

He J, Hashikawa T, Ojima H, Kinouchi Y (1997) Temporal integration and duration tuning in the dorsal zone of cat auditory cortex. J Neurosci 17:2615–2625.

Kitzes LM, Gibson MM, Rose JE, Hind JE (1978) Initial discharge latency

and threshold considerations for some neurons in cochlear nuclear complex of the cat. J Neurophysiol 41:1165–1182.

Macmillan NA, Creelman CD (1991) Detection theory: a user's guide. Cambridge, MA: Cambridge UP.

Mitani S, Kitama T, Sato Y (2006) Voiceless affricate/fricative distinction by frication duration and amplitude rise slope. J Acoust Soc Am 120:1600–1607.

Mora EC, Kössl M (2004) Ambiguities in sound-duration selectivity by neurons in the inferior colliculus of the bat *Molossus molossus* from Cuba. J Neurophysiol 91:2215–2226.

Narins PM, Capranica RR (1980) Neural adaptations for processing the two-note call of the Puerto Rican treefrog, *Eleutherodactylus coqui*. Brain Behav Evol 17:48–66.

Otto KJ, Rousche PJ, Kipke DR (2005) Microstimulation in auditory cortex provides a substrate for detailed behaviors. Hear Res 210:112–117.

Parker AJ, Newsome WT (1998) Sense and the single neuron: probing the physiology of perception. Annu Rev Neurosci 21:227–277.

Pastore RE (1983) Temporal order judgment of auditory stimulus offset. Percept Psychophys 33:54–62.

Pérez-González D, Malmierca MS, Moore JM, Hernández O, Covey E (2006) Duration selective neurons in the inferior colliculus of the rat: topographic distribution and relation of duration sensitivity to other response properties. J Neurophysiol 95:823–836.

Pinheiro AD, Wu M, Jen PH (1991) Encoding repetition rate and duration in the inferior colliculus of the big brown bat, *Eptesicus fuscus*. J Comp Physiol A 169:69–85.

Pisoni DB (1977) Identification and discrimination of the relative onset time of two component tones: implications for voicing perception in stops. J Acoust Soc Am 61:1352–1361.

Pollak GD, Schuller G (1981) Tonotopic organization and encoding features of single units in inferior colliculus of horseshoe bats: functional implications for prey identification. J Neurophysiol 45:208–226.

Potter HD (1965) Patterns of acoustically evoked discharges of neurons in the mesencephalon of the bullfrog. J Neurophysiol 28:1155–1184.

Qin L, Sakai M, Chimoto S, Sato Y (2005) Interaction of excitatory and inhibitory frequency-receptive fields in determining fundamental frequency sensitivity of primary auditory cortex neurons in awake cats. Cereb Cortex 15:1371–1383.

Qin L, Chimoto S, Sakai M, Wang J, Sato Y (2007) Comparison between offset and onset responses of primary auditory cortex ON-OFF neurons in awake cats. J Neurophysiol 97:3421–3431.

Qin L, Wang JY, Sato Y (2008a) Representations of cat meows and human vowels in the primary auditory cortex of awake cats. J Neurophysiol 99:2305–2319.

Qin L, Wang J, Sato Y (2008b) Heterogeneous neuronal responses to frequency-modulated tones in the primary auditory cortex of awake cats. J Neurophysiol 100:1622–1634.

Repp BH, Liberman AM, Eccardt T, Pesetsky D (1978) Perceptual integration of acoustic cues for stop, fricative, and affricate manner. J Exp Psychol Hum Percept Perform 4:621–637.

Rhode WS, Smith PH (1986) Encoding timing and intensity in the ventral cochlear nucleus of the cat. J Neurophysiol 56:261–286.

Steinschneider M, Fishman YI, Arezzo JC (2003) Representation of the voice onset time (VOT) speech parameter in population responses within primary auditory cortex of the awake monkey. J Acoust Soc Am 114:307–321.

Steinschneider M, Volkov IO, Fishman YI, Oya H, Arezzo JC, Howard MA 3rd (2005) Intracortical responses in human and monkey primary auditory cortex support a temporal processing mechanism for encoding of the voice onset time phonetic parameter. Cereb Cortex 15:170–186.

Stevens AA, Weaver K (2005) Auditory perceptual consolidation in early-onset blindness. Neuropsychologia 43:1901–1910.

Walker KM, Ahmed B, Schnupp JW (2008) Linking cortical spike pattern codes to auditory perception. J Cogn Neurosci 20:135–152.