Linking Cortical Spike Pattern Codes to Auditory Perception

Kerry M. M. Walker, Bashir Ahmed, and Jan W. H. Schnupp

Abstract

■ Neurometric analysis has proven to be a powerful tool for studying links between neural activity and perception, especially in visual and somatosensory cortices, but conventional neurometrics are based on a simplistic rate-coding hypothesis that is clearly at odds with the rich and complex temporal spiking patterns evoked by many natural stimuli. In this study, we investigated the possible relationships between temporal spike pattern codes in the primary auditory cortex (A1) and the perceptual detection of subtle changes in the temporal structure of a natural sound. Using a two-alternative forced-choice oddity task, we measured the ability of human listeners to detect local time reversals in a marmoset twitter call. We also recorded responses of neurons in A1 of anesthetized and awake ferrets to these stimuli, and analyzed these responses using a novel neurometric approach that is sensitive to temporal discharge patterns. We found that although spike count-based neurometrics were inadequate to account for behavioral performance on this auditory task, neurometrics based on the temporal discharge patterns of populations of A1 units closely matched the psychometric performance curve, but only if the spiking patterns were resolved at temporal resolutions of 20 msec or better. These results demonstrate that neurometric discrimination curves can be calculated for temporal spiking patterns, and they suggest that such an extension of previous spike count-based approaches is likely to be essential for understanding the neural correlates of the perception of stimuli with a complex temporal structure. ■

INTRODUCTION

Neurometric analysis has been a powerful tool for linking electrophysiological activity with perception, and has made important contributions to our understanding of vision and somatosensation (Liu & Newsome, 2005; Luna, Hernandez, Brody, & Romo, 2005; Purushothaman & Bradley, 2005; Krug, Cumming, & Parker, 2004; Parker, Krug, & Cumming, 2002; Hernandez, Zainos, & Romo, 2000; Parker & Newsome, 1998; Merchant, Zainos, Hernandez, Salinas, & Romo, 1997; Britten, Shadlen, Newsome, & Movshon, 1992), but has been used to a much lesser extent in auditory research (Narayan, Grana, & Sen, 2006; Javel & Viemeister, 2000; Fay & Coombs, 1992; Relkin & Pelli, 1987). Classically, neurometric studies investigate whether stimulus-related changes in firing rate within a population of sensory neurons could provide the "psychophysical signal" on which sensory discriminations are based. In practice, firing rates are measured over relatively long periods and are subjected to receiver operating characteristic (ROC) analysis to construct neurometric discrimination functions which can be directly compared to psychometric functions measured in two-alternative forced-choice (2AFC) behavioral experiments. This approach has proved useful for the study of stimuli which vary along a single perceptual dimension. However, many natural stimuli are characterized by their complex and individual temporal structure, which cannot easily be represented by a single spike count variable. For example, the characteristic temporal features of many animal vocalizations are encoded by time-locked responses in the auditory nerve. Early stages of the auditory pathway faithfully preserve this temporal pattern code, using special adaptations such as unusually large and secure synapses (Trussell, 2002). As information about a stimulus ascends the auditory pathway toward the primary auditory cortex (A1), individual neurons are increasingly less able to synchronize their discharges to very rapid temporal modulations in their inputs (Cariani, 1999), and the relationship between discharge patterns and temporal features of the stimulus becomes more complicated. Nevertheless, temporal discharge patterns appear to continue to play a role in representing some temporal features of acoustic stimuli in A1 (Narayan et al., 2006; DeWeese, Wehr, & Zador, 2003; Lu, Liang, & Wang, 2001b; Wang, 2000; Winter & Funkenstein, 1973), as well as aspects of complex visual stimuli in visual cortical areas (Eskandar, Richmond, & Optican, 1992). We have recently shown that temporal discharge patterns in A1 of the ferret (*Mustela putorious*) provide considerable amounts of information about the identity of a set of unfamiliar natural sounds

University of Oxford

(natural and time reversed marmoset twitter calls) even though spike rate codes read out over time scales coarser than 40 msec are entirely uninformative (Schnupp, Hall, Kokelaar, & Ahmed, 2006). However, if temporal coding at a time scale of tens of milliseconds in A1 indeed plays an important role in the perception of complex sounds, then we would predict that the temporal patterns should become increasingly harder to discriminate as the stimuli become perceptually more difficult to distinguish. Here we have tested this prediction by developing a "spike pattern neurometric" analysis which enabled us to compare the reliability of ferret A1 spike pattern codes directly with the psychophysical performance of human listeners asked to detect small changes in the temporal structure of a complex natural sound. We found an excellent correspondence between the psychometric and the neurometric results, but only if neural responses were analyzed at temporal resolutions of less than about 20 msec.

METHODS

Stimuli

For this study we required a complex, natural sound with a rich temporal structure. Our aim was to then alter the temporal structure of the stimulus in a manner that would be progressively more difficult to detect perceptually. We used a digitized recording of a marmoset "twitter" call kindly provided by Dr. X. Wang (Wang & Kadia, 2001). We manipulated the temporal structure of this sound using a procedure similar to that previously used in a study of speech recognition (Saberi & Perrott, 1999). Saberi and Perrott locally time-reversed consecutive segments of speech and found that speech intelligibility decreased when the width of the time-reversed segments was 50 msec or longer. Saberi and Perrott's method of simply cutting sound samples into segments, time-reversing each, and rejoining the segments leads to amplitude transient artifacts at the segment boundaries because amplitude transitions at the newly apposed segment boundaries are unlikely to be smooth. In the manipulated speech samples used by Saberi and Perrott, these transient artifacts are audible as a "buzzing" that is superimposed on the manipulated speech. These artifacts do not affect speech intelligibility, but in the study here, we set out to measure whether the temporally manipulated sounds could be distinguished from the original signal, and transient artifacts would have given a strong and rather trivial cue as to which signals had been manipulated by local time reversals. We therefore had to "cross-fade" between consecutive time inverted segments, thus "smoothing out" transient artifacts and discontinuities at the boundaries of the segments. To achieve this, we generated the sound segments to be reversed by first modulating the amplitude of the original signal within each segment with a cosine window function. The sequential cosine windows had 50% temporal overlap so that by simply adding each time-reversed segment across the duration of the stimulus, there was no net amplitude modulation of the original signal. The time-reversed windows were either 10, 20, 40, or 80 msec wide. For decreasing window durations, the locally time-reversed sound more and more closely approximates the original stimulus waveform, converging to the original sound "in the limit" as the window length approaches zero. For brevity, we refer to the temporally manipulated stimuli as "flipped twitters." Figure 1 illustrates these stimuli. The sound files and the Matlab (The MathWorks, Natick, MA) code used to temporally manipulate the original waveform are provided as supplementary material.

Psychophysics

Human listeners ($n = 6$) performed a 2AFC oddity task in a soundproof booth. For the psychophysical experiment, the natural and flipped twitter stimuli were generated in Matlab and presented at a comfortable listening level over headphones (Sony MDR-CD250, Surrey, UK) with an RM1 mobile processor (Tucker-Davis Technologies, Alachua, FL, USA). On each trial, three twitter stimuli were presented in sequence. The sequences, chosen at random, were of the form "ABB," "BBA," "AAB," or "BAA," where "A" stands for the natural twitter and "B" stands for one of the locally time-reversed stimuli. Thus, either the first or the last stimulus in the sequence (but not both) differed from the stimulus in the middle, and the middle stimulus served as the "standard" for comparison. The listeners were instructed to indicate by keypress whether the first or the last stimulus differed from the other two. At least one of the stimuli in the sequence was always a natural twitter stimulus so that the experiment measured the "perceptual distance" of each of the temporally manipulated stimuli from the original sound. Mutual perceptual distances between pairs of different flipped twitter stimuli were not assessed. Catch trials with "0 msec" time reversal windows (i.e., an "AAA" sequence) were also included in this paradigm to determine if first/third guesses were balanced. Subjects completed 10 demonstration trials (two trials per stimulus condition) prior to testing in order to familiarize themselves with the experimental procedure. Training prior to data collection was deliberately kept at a minimum so that the human volunteers, like the animals in the electrophysiology experiments, had very limited exposure to this particular type of stimulus prior to data collection. The subjects then completed four blocks of 10 randomly interleaved test trials per stimulus condition, and were permitted to rest between blocks. Informed consent was received from each subject before they began the experiment and the psychophysical procedures were approved by the local ethical review com-

Figure 1. Natural and flipped twitter stimuli. (A) Schematic of our local time reversals of the acoustic stimulus. In the top panel, the natural stimulus (a marmoset twitter call) waveform is shown in blue with overlapping cosine functions superimposed in red. The ''window'' of the stimulus falling within each period of the waveforms was amplified by the cosine function and locally reversed in time. (B) Spectrograms of the natural stimulus (top), and the ''flipped'' stimuli with 10, 20, 40, and 80 msec flipped window widths.
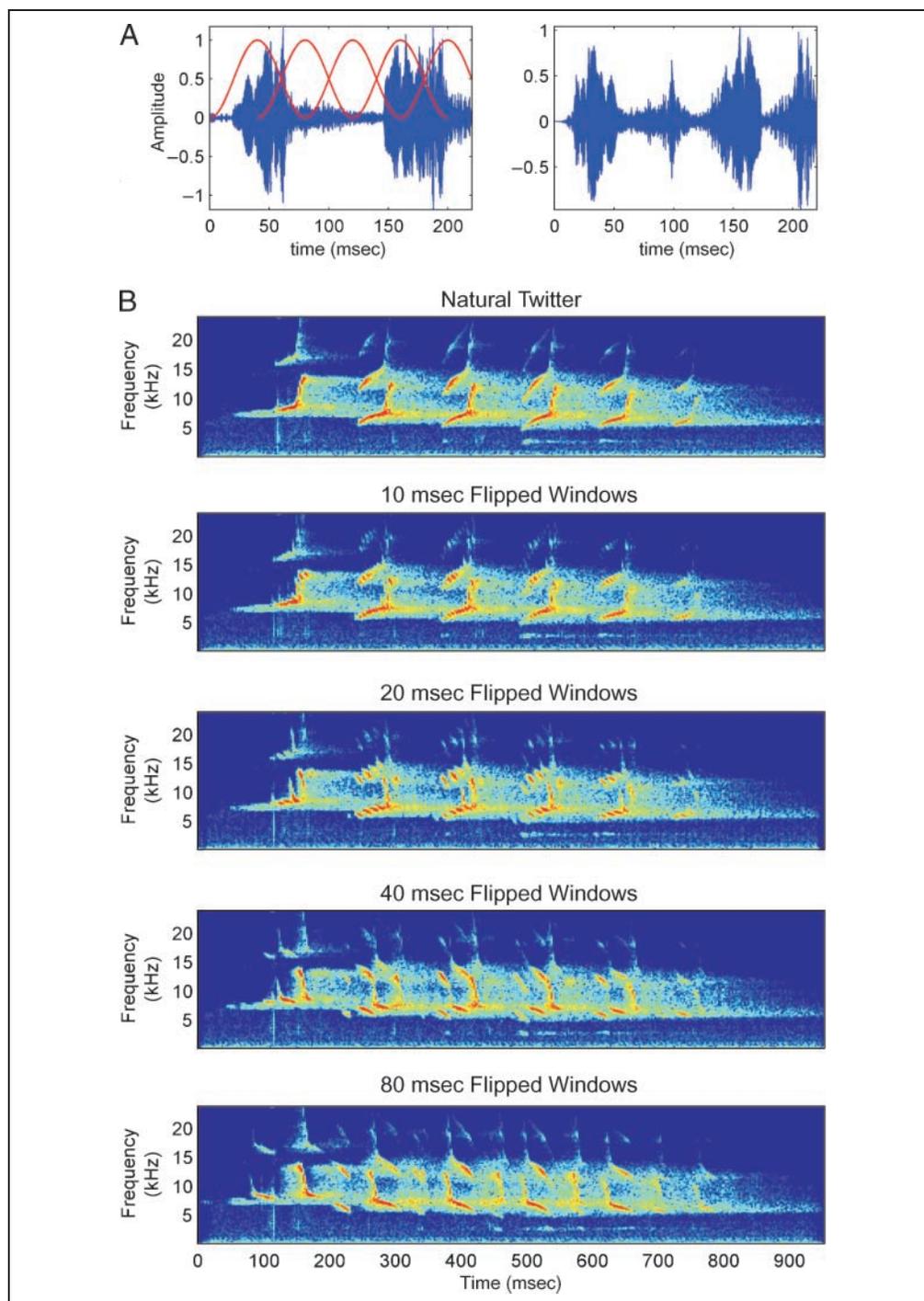
mittee of the Experimental Psychology Department of the University of Oxford.

Extracellular Recordings

All animal experiments were performed under license from the UK Home Office in accordance with the Animal (Scientific Procedures) Act 1986 and were approved by the ethical review board of the Oxford University Animal Care and Use Committee.

Anesthetized Preparation

In three normal adult ferrets, recordings were carried out under local anesthetic. Anesthesia was induced by a 2-ml/kg intramuscular injection of alphaxalone/alphadolone acetate (Saffan; Schering-Plough Animal Health,

Welwyn Garden City, UK). The left radial vein was cannulated, and anesthesia was switched to domitor (0.022 mg/kg/hr) and ketamine (Vetalar, 5 mg/kg/hr) via continuous infusion (Perfusor Secura FT infusor, B. Braun Medical, USA), along with 5 ml/hr of saline supplemented by 5% glucose. The trachea was cannulated and the animal was artificially ventilated (7025 respirator, Ugo Basile, Milan, Italy) throughout the experiment with oxygen-enriched air.

The temporal muscles were retracted to expose the skull, and a metal head holder was fixed to the skull with stainless steel screws and dental acrylic (Simplex Rapid, Austenal Dental, Harrow, UK). The auditory cortex was exposed by craniotomy and removal of the dura. Mineral oil was applied to the exposed cortex to prevent dehydration. Expired CO_2 and ECG were carefully monitored to ensure stable anesthesia throughout (AS/3 monitor, Datex Ohmeda, Finland). All electrophysiological data were recorded in a sound-insulated chamber (Industrial Acoustics Company, Winchester). Spike activity was recorded using 2 MΩ 4 $\times$ 4 silicon array "Michigan probes" (Centre for Neural Communication Technology, University of Michigan, USA). Signals were band-pass filtered (300 Hz to 3 kHz) and digitized at 25 kHz using a TDT multichannel recording system (Tucker-Davis Technologies). BrainWare software (Tucker-Davis Technologies) was used to control stimulus presentation and data collection. To characterize the responses of A1 units, the natural and "flipped" twitter stimuli were presented at 1.5-sec intervals. Each stimulus was presented between 10 and 40 times at a comfortable listening level in a randomly interleaved order. Stimuli were generated using an RP2 real-time processor (Tucker-Davis Technologies) and were presented diotically through miniature earphones (Panasonic RPHV297, Bracknell, UK) mounted on plastic otoscope speculae inserted into both ear canals. The ferrets had no experience of these stimuli prior to the experiment.

Awake Preparation

In one normal adult ferret, extracellular recordings were carried out while the animal was awake and passively listening to the stimuli. This awake ferret preparation was fundamentally similar to that used in the laboratory of Fritz, Shamma, Elhilali, and Klein (2003). A cranial mount was surgically implanted a month before the first recording session. Surgical anesthesia was induced with a domitor (0.022 ml/kg) and ketamine (0.050 ml/kg) intramuscular injection, and was maintained with oxygen-enriched isofluorine. The temporal muscles were retracted and partially removed to expose the skull and the auditory cortex was exposed by craniotomy. An acrylic cranial mount was fixed to the skull (Simplex Rapid). This mount contained a metal well above the craniotomy that allowed access to the dura, and could be sealed with a rubber stop to protect the underlying tissue between re-

cording sessions. It also contained a metal fitting which allowed the head to be restrained by fixing the head mount against a solid recording frame. During the month following implant surgery, the animal was allowed to recover, and was trained with positive reinforcement (food treats) to accept head restraint. Recordings were then carried out on 14 days over the period of 1 month. The animal's head was fixed during recording sessions, and quartz/platinum–tungsten electrodes were lowered through the dura to record auditory cortical activity (Thomas Recording, Giessen, Germany). Up to five electrodes could be lowered into the brain at once, with a single 1–2 MΩ recording site on each electrode. Stimuli were presented in the free-field in an anechoic room via an Audax TWO26M0 speaker (Audax Industries, Chateau du Loir, France) located approximately 80 cm from the animal's head, at 30° contralateral from the midline. Each recording session lasted up to 5 hr, less if the animal exhibited signs of becoming uncomfortable or restless. All other spike signal acquisition and stimulus presentation details were as described for the anesthetized preparation.

Isolating Unit Activity

Neural units were isolated from the digitized signal in BrainWare, using a *k*-means clustering algorithm (Martinez & Martinez, 2005) to group data according to spike features such as amplitude, width, and area. The number of clusters for each site was chosen by assessing cluster separability based upon the similarity of spike shapes across and within the candidate clusters. Clusters with highly homogeneous spike shapes and clear evidence of a refractory period in their interspike-interval histograms were regarded to be single units (i.e., responses from one neuron), whereas all others were termed multiunits (i.e., responses from a small collection of neurons).

Spike-timing data from acoustically responsive units were exported to Matlab for all further analyses. Unit responses to twitter stimuli were visualized as raster plots, which displayed spike events as a function of time for each stimulus presentation (a response "sweep"). Based on a visual inspection of the raster plots, only units that exhibited stimulus-driven responses were included in further analyses. Only the first 1280 msec of response data following onset of each twitter stimulus was analyzed because the vast majority of spikes occurred within the first 800 to 1200 msec of responses, and 1280 was a convenient logarithmic step value for our sampling rate analyses (see below).

Principal Components Analysis

To prepare the neural response data for statistical analysis, poststimulus time histograms (PSTHs) of the responses of each unit were generated so that each 1280 msec long response "sweep" was represented as a

vector of consecutive spike counts. Each vector formed a row in a unit's PSTH matrix. The dimensionality of this vector was dependent upon the bin widths (or temporal resolution) at which spike counts were sampled over the duration of the stimulus presentation. A bin width of 5 msec, which divided the 1280 msec long response into 256 time bins, was used for all analyses, except those which specifically examine the effects of bin sizes. The vector of spike counts across bins can be conceptualized as a point in a 256-dimensional space. Principal components analysis (PCA) was carried out to reduce the dimensionality of this response space. PCA remaps data in a high-dimensional vector space onto a new orthogonal coordinate system such that the smallest number of these new principal component dimensions can explain the largest proportion of the variance in the data (Hernandez et al., 2000). Often, most of the "structure" of the data can be captured by a relatively small number of high-order principal components, whereas many of the low-order principal component dimensions are largely attributable to "noise" and can be discarded. Here, we only used the first five principal components to represent each response pattern. Thus, each response pattern was first quantified as a PSTH of 256 bins, each 5 msec wide, and then represented as a five-dimensional point in principal component space. Finally, each principal component value was multiplied by the corresponding eigenvalue of that principal component, in order to weight the principal components by the amount of response variance they explained.

Receiver Operating Characteristics

The Euclidean distance between two 5-element response vectors in pc-space provided a measure of similarity between these responses because responses with more similar temporal discharge patterns would be closer in pc-space. The Euclidean distance between two responses to the same stimulus was referred to as a *within-category* distance, and *across-category* distances were measured between responses to natural and flipped twitters. For each unit, the distributions of all within-category and across-category Euclidean distances were collected for each flipped twitter condition. In the psychometric oddball task described above, human subjects had to compare one natural and one flipped target stimulus to a standard, and we assume that these stimuli evoked responses in their A1, which are essentially similar to those we recorded in the auditory cortex of the ferrets. The subjects thus had to decide which of the target responses was more similar to the standard, and here we model this decision process by comparing one within-category and one across-category distance drawn at random from the observed distributions. Consequently, a unit should support good psychometric discrimination ability if the across-category distances were reliably greater than the within-category distances

(i.e., if there is little overlap in the within- and across-category distributions). Like previous authors (e.g., Britten et al., 1992), we used ROC analysis to construct neurometric curves for each unit. We constructed ROC curves for each flipped twitter condition by comparing the differences in within-category and across-category distances against a wide range of possible criterion values. The area under the resulting ROC curves provides a measure of discrimination performance (expected proportion of correct responses) that is independent of potential criterion bias and can be directly compared with the performance on our human psychophysics task (Green & Swets, 1974).

The standard error of an ROC curve (Hanley & McNeil, 1982) was calculated as:

$$SE = \sqrt{\frac{A(1-A) + (n_b - 1)(Q_1 - A^2) + (n_w - 1)(Q_2 - A^2)}{n_w n_b}},$$

where A is the area under the curve, n_b and n_w are the number of between- and within-category distances, respectively, and Q_1 and Q_2 were estimated by:

$$Q_1 = \frac{A}{2-A}, \text{ and } Q_2 = \frac{2A^2}{1+A}.$$

RESULTS

Psychophysics

Six human subjects completed a 2AFC oddity task in which they were asked to discriminate a natural marmoset twitter call from locally time-reversed ("flipped") copies of the stimulus (see Methods and Figure 1). Flipped twitters with "0 msec" reversed window widths are natural, unmanipulated twitters. Discrimination scores for each subject across flipped twitter conditions are shown in Figure 2. The performance of human listeners in discriminating natural and flipped twitter stimuli declined as the width of the flipped time bins decreased. Performance on the discrimination task was near perfect for all subjects when the reversed windows were 80 msec wide, and was impaired for most subjects with 40 msec wide flipped windows. When the windows were 20 msec wide, the mean discrimination performance was near chance, and subjects could not discriminate 10 msec flipped from natural twitters ($p > .05$, based on a binomial distribution with 20 trials). Performance did not significantly differ between trials on which the nonstandard stimulus came first in the sequence versus those in which it was the last in the sequence (Wilson score interval, $p > .05$) (Wilson, 1927).

Neurometrics of Individual Units

Multielectrode extracellular recordings of responses to natural and flipped twitters were collected in the left
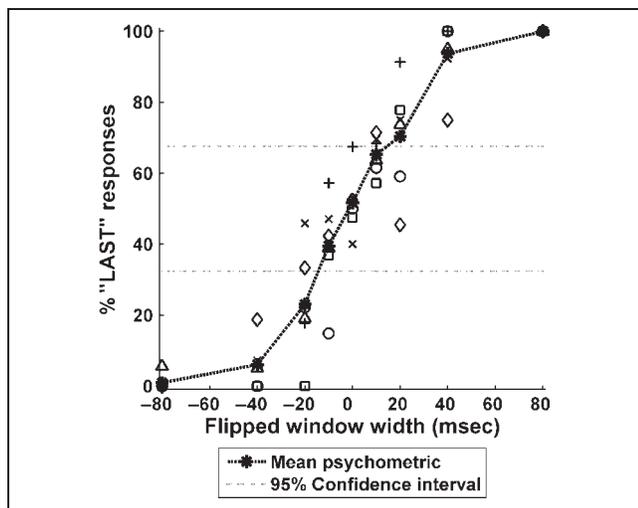
Figure 2. Human discrimination performance. Psychometric curve showing the percentage of correct responses in the 2AFC auditory discrimination task. "Flipped window width" of +20 msec and −20 msec both refer to oddball stimuli which were locally time reversed with 20 msec windows, but a negative value indicates that the oddball stimulus was the first of the three stimuli presented in the trial, whereas positive values indicate that the oddball stimulus was presented last. The six different types of symbols show the mean percentage of "last" responses for each subject as a function of stimulus condition. The continuous line shows the mean across the six subjects.

primary auditory cortex of three adult, anesthetized ferrets. In total, 142 stimulus-driven units were identified: 77 in the first ferret, 28 in the second, and 37 in the third. These included 120 single units and 22 multiunit clusters.

Visual inspection of standard raster plots indicated that a subset of the units in A1 responded to twitter stimuli with reliable and stimulus-specific temporal discharge patterns (Figure 3A and B), as observed previously (Schnupp et al., 2006; Wang, Merzenich, Beitel, & Schreiner, 1995). Although the spike latencies and spike counts of these units varied somewhat as the same stimulus was repeatedly presented, and although their temporal discharge patterns appeared to contain noise from spontaneous activity across trials, this subset of units, nevertheless, exhibited reproducible temporal discharge patterns which, if appropriately decoded, ought to support the discrimination of the natural from some of the flipped twitter calls. PSTHs were generated for each unit, in which the response to each twitter presentation was represented as a vector of spike counts in 5-msec time intervals following stimulus onset. For each unit, PCA was used to reduce the dimensionality of these response vectors (see Methods). The first five principal components together were sufficient to explain between 13% and 60% (mean 27.3%) of the response variance for each unit (Supplementary Figure 1).

A discrimination algorithm was designed to attempt to distinguish natural from flipped twitters based on the information available in the temporal discharge patterns of A1 units (see Methods). If the temporal discharge pattern of a unit was reliably similar for repeated presentations of the same stimulus and sufficiently different when another stimulus was presented, then two responses of that unit to the same twitter stimulus (a "within" twitter comparison) should be closer in pc-space than two responses to different twitter stimuli (an "across" twitter comparison). But if the responses lacked reproducible and stimulus-selective spiking patterns, then their distances in pc-space would be uninformative about

Figure 3. Responses of two ferret A1 units to the presentation of twitter stimuli. (A) Raster plot displays showing the responses of one unit to the natural twitter call (top) and the four locally time-reversed calls (bottom). Each row of dots shows the response to a single presentation of the stimulus, with each dot showing the time of occurrence of an action potential after stimulus onset. (B) Raster plots showing the responses of a different unit.
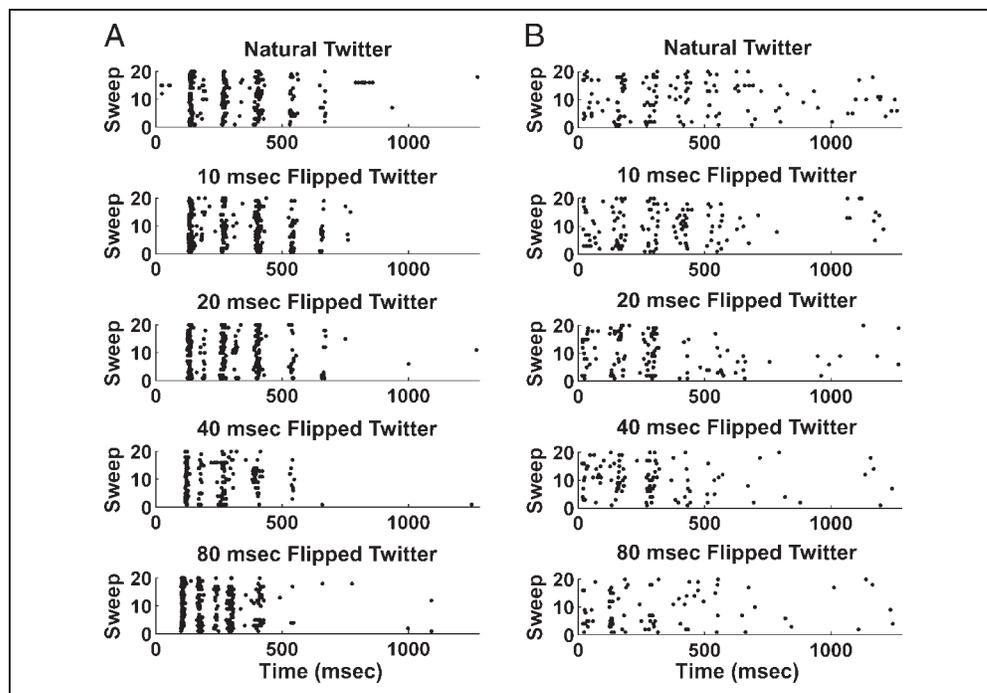
Figure 4. Calculating
neurometrics for two ferret
A1 units. (A, B) Distributions
of within-category response
distances (filled black
circles) and across-category
distances (open gray circles)
for the units shown in
Figure 3A and B,
respectively. (C, D) ROC
curves constructed from the
distributions shown in A and
B, respectively. ROC curves
are shown for the 0 (blue),
10 (green), 20 (yellow),
40 (magenta), and 80 msec
(red) stimulus conditions.
(E, F) Neurometric
functions (black circles) for
the two units, constructed
by plotting the area under
the ROC curves shown in
C and D. The average
psychometric performance
of the human volunteers
with standard error bars is
also shown for comparison
(black asterisks).
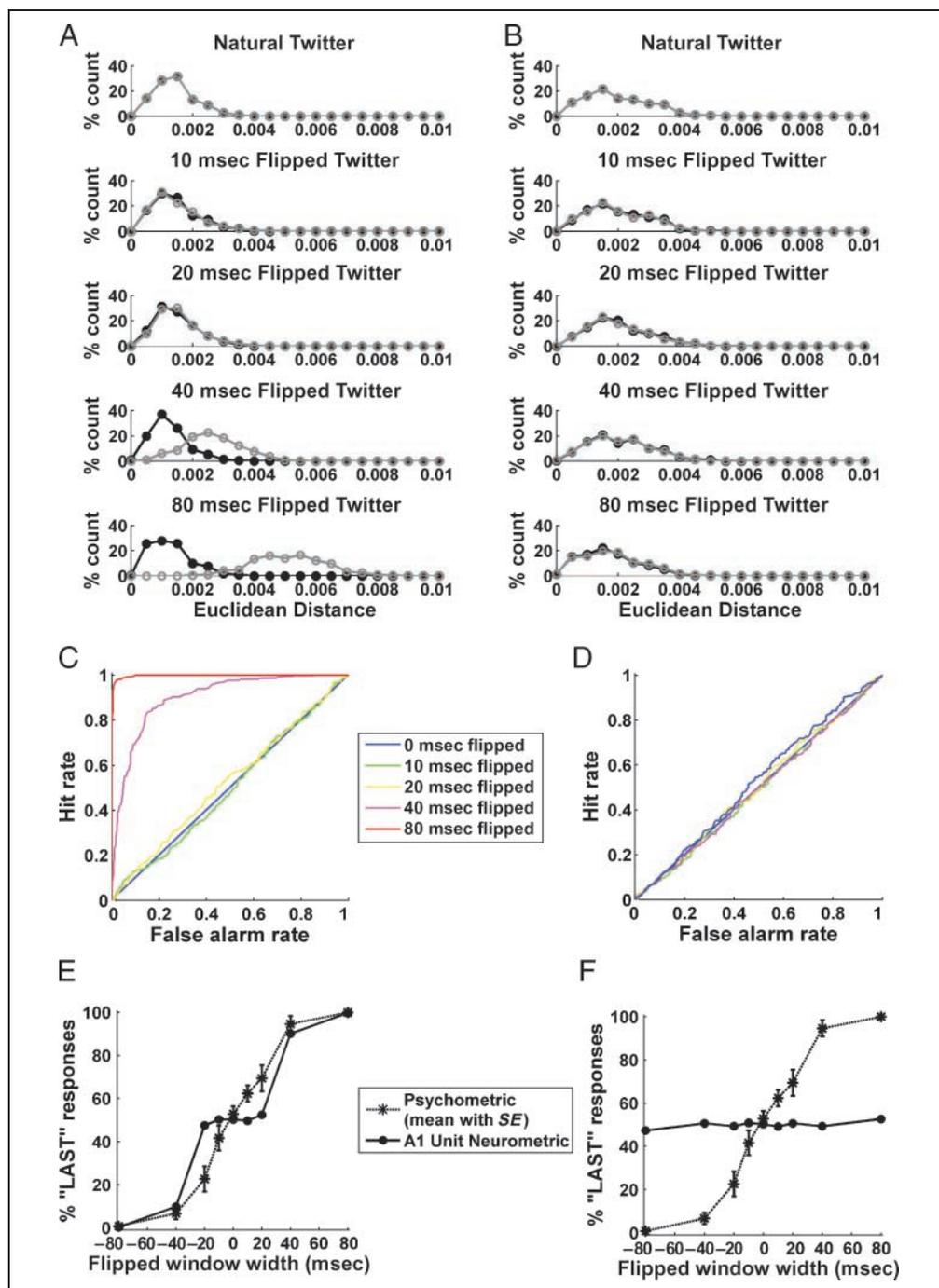
the categorization of the twitter stimuli to which they
were responding. Histograms of within- and between-
category distance distributions for each flipped twitter
condition for the two units illustrated in Figure 3A and B
are shown in Figure 4A and B, respectively. Using analyses
derived from signal detection theory, ROC curves were
calculated for each flipped twitter condition (Figure 4C
and D) (Green & Swets, 1974), and a neurometric dis-
crimination curve for each unit was then derived from
these ROC curves (Figure 4E and F; see Methods).

The neurometric functions of all 142 units are plotted
along with the psychometric function in Figure 5A. From

this figure, it is clear that neurometric discrimination
curves varied widely from unit to unit. Neurometric per-
formance at each flipped twitter condition ranged from
chance performance up to psychometric means. As is il-
lustrated in Figure 6, the neurometric scores for many
units were significantly above chance (standard error of
the ROC curve, $p < .05$), and the discrimination perfor-
mance of these units improved as flipped window
widths increased.

Figure 5B shows the results of the same neurometric
algorithm when a single spike rate variable, calculated
over the whole response period, is used instead of

Figure 5. Neurometric functions for a population of A1 units in anaesthetized ferrets. (A) Temporal pattern neurometrics for all 142 units in our sample are plotted superimposed (gray circles). The average psychometric performance (with standard error bars) of the human listeners is also shown for comparison (black asterisks). (B) Spike rate neurometrics for the same 142 units (gray circles) are plotted, along with the average psychometric curve (black asterisks with standard error bars). (C) Spike rate neurometrics calculated for each of the 142 units from the spike rates observed in a single, optimally chosen, 5-msec-wide time bin. (D) Difference between temporal pattern neurometric curves shown in A and ''best bin'' neurometric curves shown in C, shown for all those units whose neurometric performance reaches at least 70% in either analysis for any stimulus parameter.
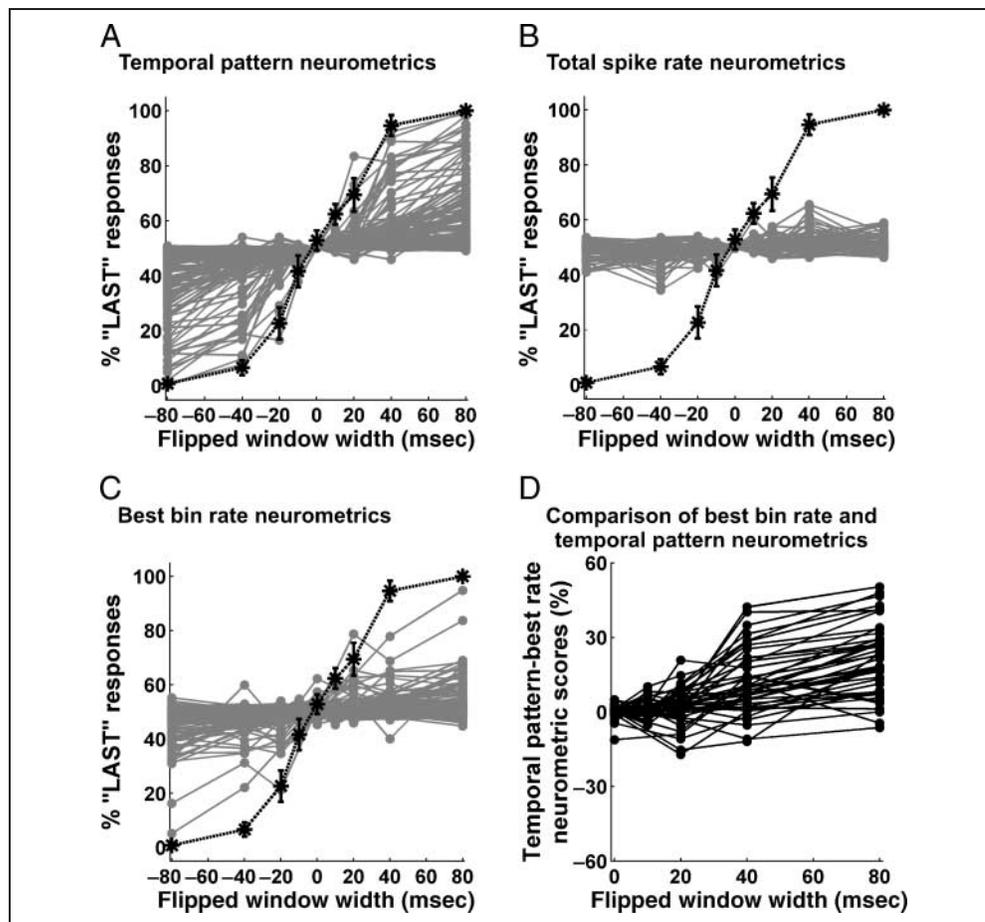
temporal pattern information. These neurometrics did not reach psychometric performance values, and they did not show systematic improvement in discrimination as the flipped windows widened. Of course, one might ask whether the much poorer performance of the sin-
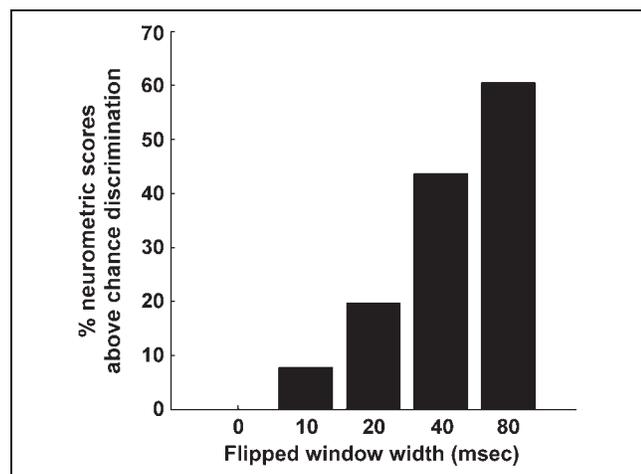

Figure 6. More A1 units show significant neurometric discrimination for wider flipped windows. Histogram showing the number of units achieving a discrimination performance significantly above chance (at $p < .05$, standard error of the ROC curve) as a function of stimulus condition.

gle spike rate neurometric compared with the spike pattern neurometric might not simply be attributable to the fact that the temporal bin width of 5 msec used for the temporal pattern analysis is a more fortuitous choice then the 1280-msec-wide window used in the spike rate analysis shown in Figure 5B. In other words, if one were to use a single spike rate measure obtained from a single, suitably chosen 5-msec interval, would one be able to match the neurometric performance reached by the pattern analysis? To test this, we calculated single spike rate neurometrics for each 5-msec-wide time bin, from 0 to 1280 msec in 5-msec-wide steps, and for each unit chose the 5-msec-wide time bin that gave the best neurometric performance (i.e., the greatest area under the neurometric curve). The neurometrics obtained in this way for each of the 142 units are shown in Figure 5C. Although these neurometrics based on the single best 5 msec bin for each unit are considerably better than the average spike rate neurometrics shown in Figure 5B, they, nevertheless, fail to reach the same level of performance as the full pattern neurometrics shown in Figure 5A, nor do they approximate human psychometric performance as well as the pattern neurometrics do. Figure 5D plots the difference between the spike pattern neurometric and the best 5-msec bin neurometric for all units that reach at least 70% performance in either

neurometric curve for any of the stimuli. Such a comparison is of little interest for units which perform poorly with all neurometric algorithms, and they were therefore not plotted in Figure 5D in order not to clutter the graph unnecessarily. These differences are positive (i.e., pattern neurometric performance exceeds single best bin neurometric performance) in the large majority of cases. We also repeated this analysis using 10, 20, and 40 msec wide optimally chosen time bins and obtained very similar results (not shown). Full-pattern neurometrics consistently outperformed single best bin neurometrics, and single best bin neurometrics never closely matched the human psychometric curve. These results suggest that in A1, the discharge pattern, and not simply the number of spikes in any one time window, carries important information about the identity of temporally varying complex sounds.

In the neurometric algorithm above, the temporal discharge patterns across 256 time bins was quantified using only the first five principal components. It is possible that incorporating more information about response variation into the decision algorithm by including more principal components would improve the neurometric performance. An investigation of this question demonstrated the importance of the present method of weighting the principal components by their eigenvalues. When neurometrics were calculated using the raw principal component values, performance varied considerably depending upon the number of principal components used to represent the response. Initially, using more than one principal component in the algorithm improved neurometric performance for some units, but increasing the number of principal components beyond the first three caused a decline in performance for most units. This interpretation of neural responses would require an observer to consider an optimum number of principal components for each cell in order to maximize discrimination performance. In contrast, when the principal components were multiplied by their eigenvalues to account for their relative importance in explaining the response variance, using only the first five principal components yielded near-maximum discrimination performance for all neurons (Supplementary Figure 2).

From Figure 5A, it is clear that most units' individual neurometric curves were considerably flatter than the psychometric curve, and only a small proportion of neurometric curves approached the psychometric curve closely. One might expect that the perceptual decision-making process operating in an observer engaged in a sensory discrimination task might pool or combine information across a number of neurons rather than just rely on one single neuron in the auditory cortex, and it is of interest to ask what sort of pooling mechanisms might be able to explain psychometric performance. In past studies, there have been at least two commonly used methods of combining information across units: ''enveloping'' of neurometrics and pooling of responses

(Parker & Newsome, 1998). Here, both techniques were used to examine the neurometric performance of groups of units.

Enveloped Neurometrics

If a distinction between several stimuli is required, behavioral decisions might be based only on the outputs of those neurons that respond most selectively to these stimuli (Barlow, 1972). The consequences of this process of listening to the outputs of only the ''highest performing'' neurons on each discrimination task can be represented graphically by drawing an optimum or ''best'' envelope along the contour of the best neurometric curves (Parker & Newsome, 1998). By applying this best envelope approach to our data, we found that the three ''best'' units in our sample proved sufficient to obtain a neurometric curve that closely matched psychometric discrimination performance (Figure 7).

In the past, enveloping has most often been used in investigating more peripheral areas of the nervous system (Parker & Newsome, 1998). Generally speaking, how the brain might implement an algorithm that is equivalent to the enveloping of neurometric functions is as yet unclear. Thus, perhaps a more appropriate alternative to enveloping best neurometrics is to pool the responses of several simultaneously active neurons to determine how the algorithm would perform using the temporal patterns of a summed spike input. This is especially relevant in A1, where complex stimuli are
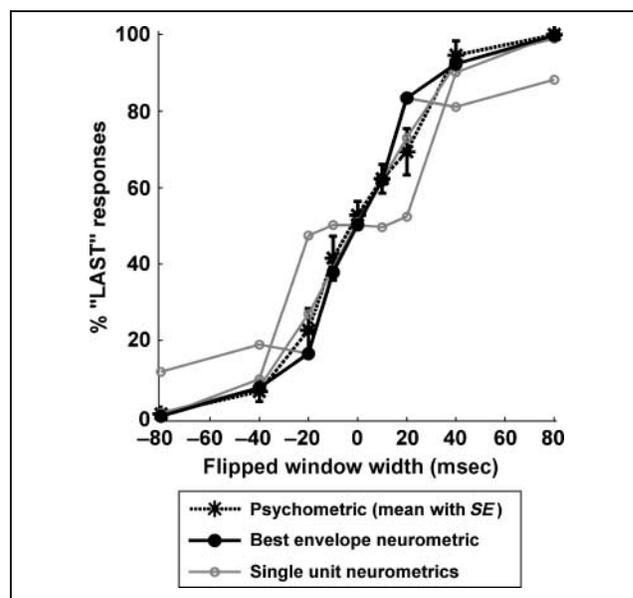

Figure 7. Upper limit of neurometric performance curves across individual A1 units. ''Best envelope'' neurometric curve (black circles) constructed from the three most informative units in our sample is shown, along with the individual neurometrics of these three units (gray circles). The average psychometric performance (with standard error bars) of the human listeners is also shown (black asterisks).

thought to be encoded as the distributed firing across populations of units with varied frequency tuning (Wang et al., 1995).

Pooled Neurometrics

The 4×4 Michigan electrodes used in the anesthetized ferret experiment can simultaneously collect responses from several neural units at each of 16 electrode recording sites during each penetration into the brain. This allowed us to examine the temporal spiking patterns of neural populations by pooling the responses of neural units in each penetration. To preserve spike timing information, the 256-element response vectors for each trial were summed across all units. Once the responses had been pooled for the complete set of trials, they were transposed using PCA and temporal discharge neurometrics were calculated as for individual units. Neurometrics of pooled units across 14 multielectrode penetrations are shown in Figure 8. The neurometric curves of each individual unit are also shown for comparison, as are the human psychophysics data.

In all 14 penetrations, the pooled neurometric discriminated twitters better than most individual units alone, demonstrating that the contribution of simultaneous activity from multiple units tends to improve neurometric performance even if some of these units themselves have noisy, poorly informative responses to the stimuli. The performance of the pooled neurometrics varied considerably from site to site. In Penetration 2, neurometrics based on each individual unit performed near chance, but the neurometric performance of the pooled activity of all these units closely matched the human psychometric curve. In four other penetrations in which the individual neurometrics performed poorly, even the activity of pooled units performed at chance. The collection of pooled neurometrics varied in performance from chance up to human psychometric values (Figure 9A), demonstrating again a correspondence between the discrimination capacity of A1 responses and the perceptual performance of human listeners.

In some previous neurometric investigations of spike rate codes (e.g., Britten et al., 1992), the responses of single units recorded on separate occasions have been pooled by summing their spike counts. Correlations in firing across neurons may limit the amount of independent information that a further neuron can add to the pool, so these studies incorporated an interneuronal correlation term into their pooling models to account for this effect. When neural responses are pooled for the purposes of examining temporal spiking patterns, the effects of interneuronal correlations are less clear. Each spike in a response may be a part of the neural "signal" that represents the sound if its timing is determined by the stimulus. Or it may contribute to the neural "noise" if its timing is unrelated to the sound being presented.

Pooling information across a neural population can enhance the signal-to-noise ratio and lead to better neurometric performance only if noise correlations in the pooled population are small, so that each neuron in the pool gives a more-or-less independent look at the stimulus. To directly investigate whether noise correlations affected the pooled neurometric performance obtained from our data, we removed correlations between neural responses in the pool by "shuffling" responses obtained on separate stimulus presentations, and compared these neurometrics to those derived from pooling only synchronously recorded activity. The rationale behind this approach is similar to that motivating the use of shift-predictors in cross-correlation histogram analysis (Moore, Perkel, & Segundo, 1966). For each unit in a pool, the order of responses was randomized within each stimulus condition prior to summing the responses across units. Therefore, the response sweeps of each neuron were added to responses of other neurons in the pool that were evoked by the same stimulus, but at different points in time. The neurometric performances of these "shuffled" pools for the 14 penetrations are shown in Figure 9B. The pooled neurometric performance for most penetrations improved with shuffling, indicating that noise correlations across units do reduce the benefit obtained from pooling simultaneously recorded responses. The upper envelope of these pooled neurometrics was largely similar to the psychometric curve, with the exception of the 10-msec condition where the neurometric of one penetration performed better than human listeners.

Effects of Temporal Resolution and Integration Windows

Given that the discrimination algorithm used to calculate neurometric curves makes use of temporal pattern information, one would expect its performance to decline if the temporal resolution at which these patterns are read out is inappropriate. Binning spike data can reduce noise that results from temporal jitter in the stimulus-locked responses of auditory neurons (Hill, Stange, & Mo, 1989) and provide a more efficient encoding of information about the stimulus, but if the binning is too coarse, important temporal patterning can be lost. Therefore, there should be a range of resolutions at which A1 responses can be binned to provide the most efficient spike pattern code without disrupting discrimination performance. To explore the possible effects of sampling rate, enveloped neurometrics were constructed as for Figure 7, but the spike patterns were binned using a range of bin sizes in logarithmic steps from 2.5 to 1280 msec. The results of this analysis, shown in Figure 10, indicated that the discrimination algorithm performed best when the spike data were sampled at the finest temporal resolutions, as might be expected because these contain the most information
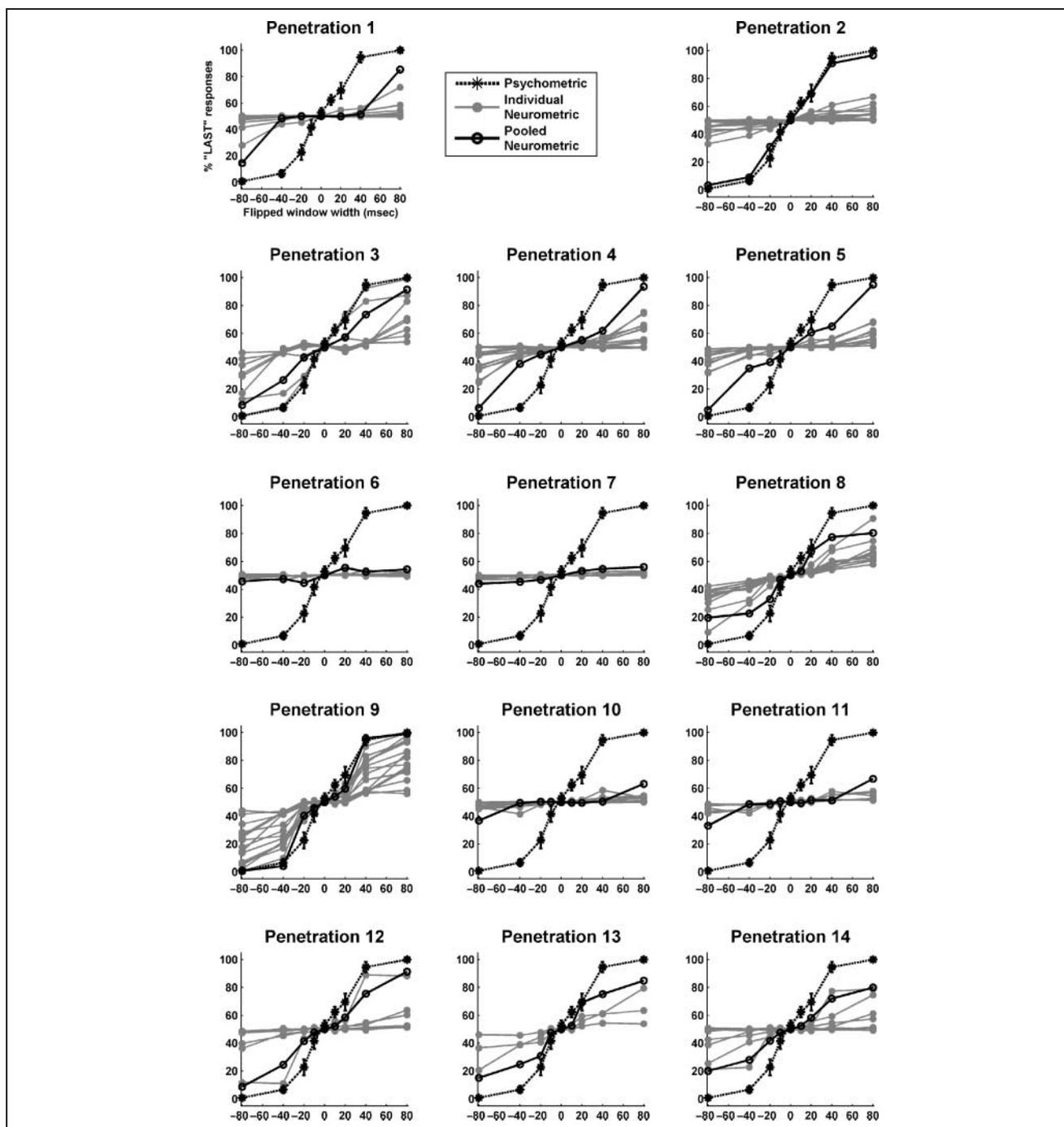
Figure 8. Neurometric curves of pooled activity within each electrode penetration. Each plot shows the neurometric curve obtained from the pooled simultaneous responses of all the units recorded in a given electrode penetration (black circles), along with the individual neurometrics of each of these units (gray circles). The average psychometric performance (with standard error bars) of the human listeners is also shown (black asterisks).

about spike timing. However, Figure 10 also shows that neurometric performance remained high when the resolution was dropped to 10 msec. For resolutions between 20 and 40 msec, neurometric performance remained near perfect for the easiest discrimination task (i.e., 80 msec flipped twitters), but was impaired for more difficult twitter discriminations. At still coarser resolutions,

performance was severely impaired across the entire stimulus set.

Analyzing the responses at an appropriately fine temporal resolution seems clearly important, but it may not be necessary to subject the whole neural response to neurometric analysis. Analyzing only the first few 100 msec of the response after stimulus onset might suffice to

Figure 9. Summary of pooled neurometric curves across all penetrations. (A) Neurometric curves derived by pooling simultaneously recorded activity in each multielectrode penetration are shown (gray circles) along with the human psychometric curve (black asterisks with standard error bars). (B) Neurometric curves derived from the same data, but activity is pooled after randomly shuffling the order of responses for each unit to disrupt stimulus-independent trial-to-trial correlations in the activity in each penetration are shown (gray circles), where the order of each unit's response sweeps was scrambled prior to pooling. The human psychometric curve is shown for comparison (black asterisks with standard error bars).
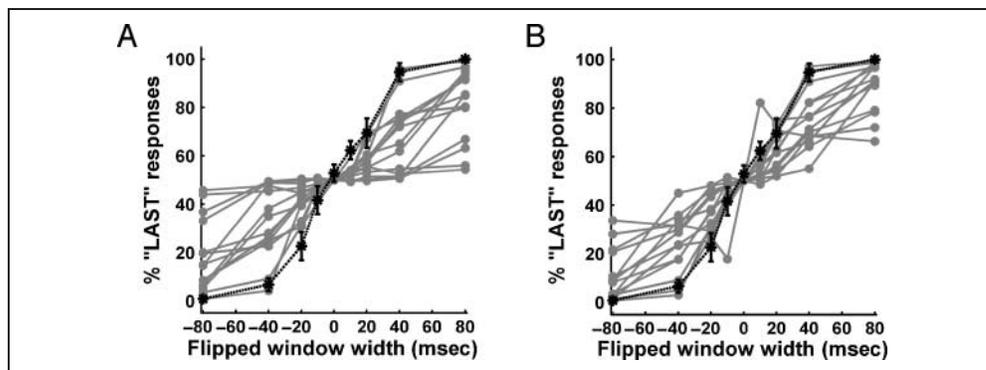
achieve adequate neurometric performance. To investigate this possibility, we repeated the neurometric analysis described above at a 5-msec temporal resolution, but we progressively shortened the length of the analyzed response from the full 1280 msec to the first 800, 400, 200, 100, and 50 msec. The result of this analysis is shown in Figure 11. Reducing the analyzed response duration to 400 msec can be seen to have little impact on the neurometrics, and even if only the first 200 msec of the response are analyzed, a proportion of the units are able to maintain good neurometric performance. However, when the analyzed response periods were reduced further to 100 msec or less, neurometric performance collapsed. Our results are therefore broadly similar to those observed by Narayan et al. (2006), who reported optimum temporal resolutions for the analysis of song bird auditory forebrain responses to complex vocalizations of about 10–

20 msec, and best temporal integration windows on the order of several hundred milliseconds (597 msec median). It would be interesting for future studies to investigate how these temporal integration windows might relate to reaction times in listeners.

Neurometrics in the Awake Animal

If the above similarities between psychometric and A1 neurometric curves indicate that perceptual performance is limited by the resolution of temporal spiking patterns in A1, these similarities should hold when we use the same neurometric algorithm to interpret the auditory cortical responses of an awake animal. An unanesthetized animal more closely resembles the behavioral state of the human listeners. Figure 12 shows the neurometric performance of 29 auditory cortical single units recorded in an awake ferret. As shown in the anesthetized ferret, neurometrics based on the temporal discharge patterns of auditory cortical neurons in the awake ferret vary from chance performance up to human psychometric values, such that the best envelope of these curves closely matches the psychometric curve (Figure 12A). In contrast, neurometrics based on the spike rates of these responses do not reach behavioral performance levels (Figure 12B).

A Note on the Stimulus–Response Relation

Many previous studies have investigated the temporal and spectral characteristics of neural firing in response to complex stimuli (Narayan et al., 2006; Theunissen, Sen, & Doupe, 2000; Wang et al., 1995). Although significant progress has been made in this area, some aspects of the stimulus–response relation remain unclear. In the present study, rather than focus on the nature of the stimulus–response relation, we have focused on the relation between neural responses and psychophysical performance. However, a simple visual inspection of the pooled temporal discharge patterns of A1 units in
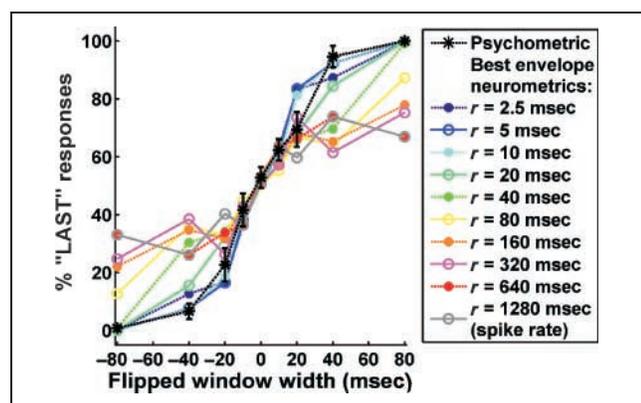
Figure 10. Sensitivity of the neurometric curves to changes in temporal resolution. Best-envelope neurometrics were constructed as for Figure 7 (colored circles), but the temporal resolution (bin width) used to represent the raw response pattern vectors was varied as shown in the legend. Note that some lines are partially occluded by others. The average psychometric performance (with standard error bars) of the human listeners is also shown (black asterisks). Neurometrics are comparable to psychometric performance only at relatively fine temporal resolutions.

Figure 11. Effect of reducing the total analysis time window on neurometric performance. Neurometric curves calculated at 5 msec temporal resolution and plotted as in Figure 5A, but the duration of the analyzed response period is progressively reduced from the full 1280 msec duration as indicated above each panel.
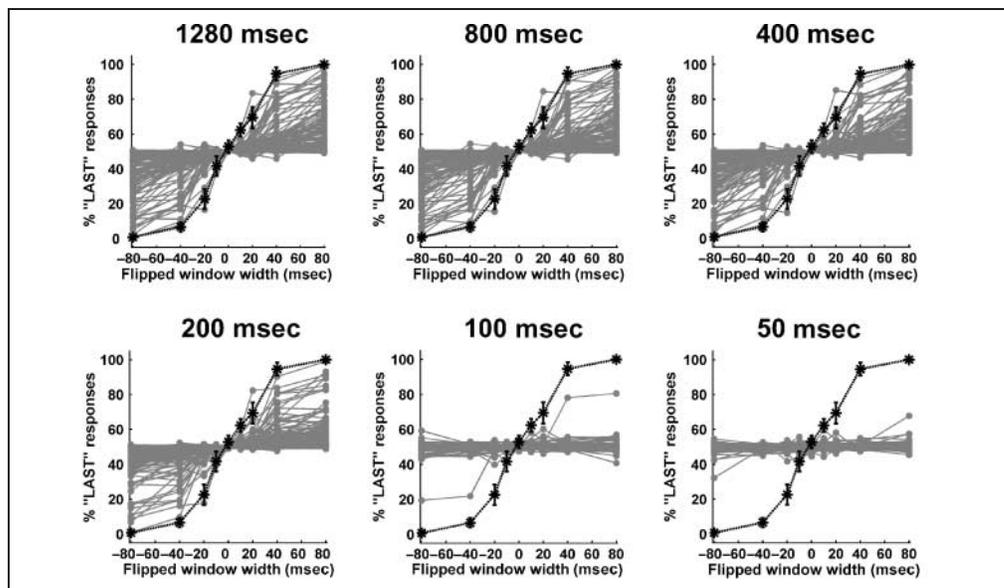
anesthetized ferrets suggests that the cortical population as a whole responded with correlated firing to discrete events in the spectro-temporal acoustic pattern of the stimulus (Figure 13A). The pooled PSTH was similar for auditory cortical neurons recorded in the awake animal, but with stronger responses to the onset and offset of the sound (Figure 13B). This type of synchronization of auditory cortical firing with transients in complex vocalizations has been previously described in other species (Wang et al., 1995; Winter & Funkenstein, 1973). At the single-unit level, the stimulus–response relationship may be more complex, but a more complete description of this relation was not attempted here.

DISCUSSION

Previous studies that have compared neurometric functions in the sensory cortex and psychometric perfor-

mance have focused almost exclusively on spike rates (Liu & Newsome, 2005; Krug et al., 2004; Britten et al., 1992; Tolhurst, Movshon, & Dean, 1983). To our knowledge, only one previous study (Hernandez et al., 2000) has explored potential links between neurometrics based on temporal discharge patterns and psychometric performance. Hernandez and colleagues investigated the respective roles that spike rate and periodic discharge patterns in the primary somatosensory cortex might play in encoding the frequency of a vibrating tactile stimulus. They found that spike rate-based neurometrics were adequate to account for psychophysical thresholds, whereas neurometrics derived from the periodicity of neural discharge resulted in thresholds that were much lower than those obtained behaviorally. These results served to reinforce the notion that, in the sensory cortex, mean spike rates correlate with perception, whereas temporal discharge patterns do not. Our

Figure 12. Neurometric functions for a population of A1 units in an awake ferret. (A) Temporal pattern neurometrics for all 29 units in our sample are plotted superimposed (gray circles). The average psychometric performance (with standard error bars) of the human listeners is also shown for comparison (black asterisks). (B) Spike rate neurometrics for the same 29 units (gray circles) are plotted, along with the average psychometric curve (black asterisks with standard error bars).
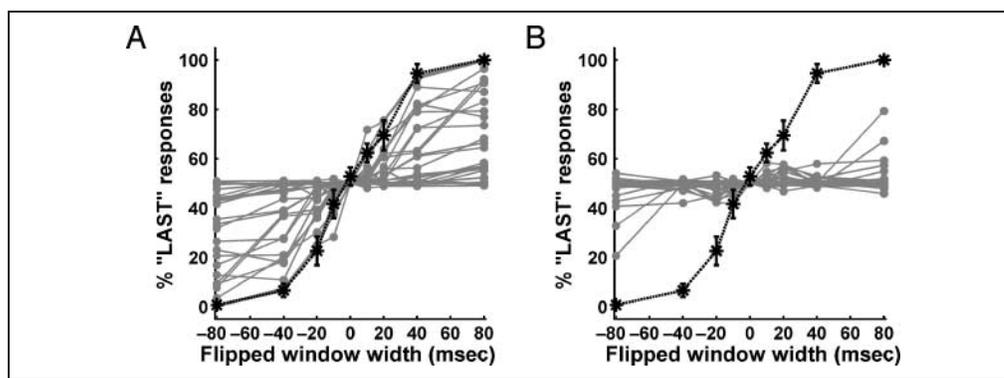
Figure 13. Pooled neural responses are locked to discrete stimulus events. (A) The normalized poststimulus time histogram (PSTH) of the pooled responses of 142 units from the anesthetized dataset is shown (black line) superimposed on the waveform of the natural twitter stimulus (gray line) to which it is responding. Stimulus onset occurs at 250 msec. (B) Normalized pooled PSTH for the 29 units recorded in the awake dataset.
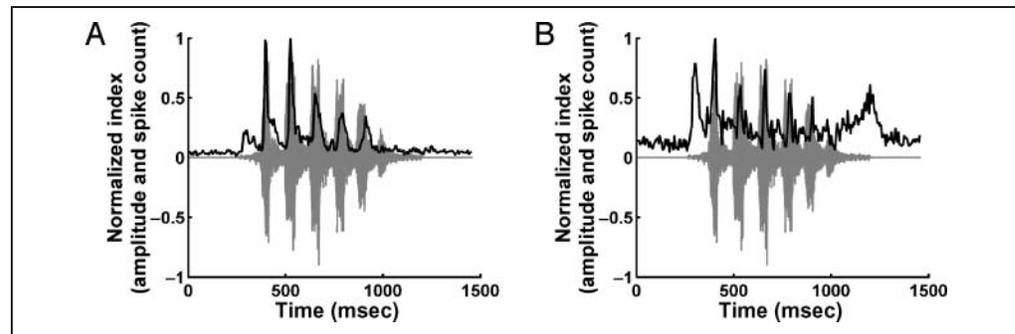
present results suggest that this may not be the general case for all stimuli. We found that spike rate-based neurometrics failed completely to account for behavioral performance, whereas the spike pattern-based neurometric analysis we devised here closely matched the psychometric curve. There are, of course, many important differences between our study and that of Hernandez et al. (2000), the most important of which is likely to be the nature of the temporal change in the stimuli that is to be detected. Hernandez et al. investigated a tactile frequency discrimination task, whereas in our case, subjects were required to detect a change in the fine structure of a complex natural sound, which might be best described perceptually as a change in the ''rhythm'' of the sound. In addition, the neurometric analysis used by Hernandez et al. was limited to the examination of temporal discharge patterns that are periodic, whereas our approach allows one to investigate temporal discharge patterns of any form.

We envisage that the adapted neurometric method which we introduced here may prove to be a valuable tool in future studies of auditory cortex function. Despite dramatic progress in our understanding of the functional organization of the auditory cortex in recent years, many aspects of the roles played by the auditory cortex in perception remain uncertain. Early lesion studies on monkeys suggested that damage to the auditory cortex impaired an animal's ability to perform cognitive tasks involving sound localization (Heffner & Masterton, 1975) or the discrimination of con-specific vocalizations, but these studies also noted that cortical lesions did not impair reflex-like orientation toward unexpected sounds. More recent observations on a human patient who had suffered extensive bilateral auditory cortex damage appear to confirm that such lesions lead to a total loss of auditory awareness, a ''cortical deafness,'' even though accurate reflexive orienting behavior to sudden sounds remains intact (Garde & Cowey, 2000). These observations suggest that the auditory cortex may play a critical role in every aspect of

the conscious perception of sound, even if sophisticated subconscious auditory processing seems to be possible without it. In recent years, a number of researchers have investigated whether particular aspects of auditory perception can be ascribed to particular auditory cortical areas. The most commonly used approach investigates stimulus-induced changes in discharge patterns in various cortical areas (Narayan et al., 2006; Wallace, Rutkowski, & Palmer, 2005; Recanzone, Guard, Phan, & Su, 2000; Middlebrooks, 1999; Rauschecker, 1998), but to link this cortical neural activity to perception, one must also demonstrate that the neural responses correlate with perceptual decisions as assessed in psychometric experiments. Ideally, this should be done on a trial-by-trial basis in an animal engaged in the perceptual task. Studies which have managed to achieve this in the visual (Britten et al., 1992) and somatosensory (Hernandez et al., 2000) systems have been very influential, but similar experiments have not yet been attempted in the auditory system, perhaps, in part, because the simple rate coding assumption inherent in conventional neurometric analysis appears inappropriate for the auditory cortex. Many recent studies have emphasized the importance of the temporal patterning of responses in the auditory cortex in the encoding of acoustic stimuli (Narayan et al., 2006; Schnupp et al., 2006; Nelken, Chechik, Mrsic-Flogel, King, & Schnupp, 2005; Furukawa & Middlebrooks, 2002; Gehr, Komiya, & Eggermont, 2000; Brugge, Reale, & Hind, 1996). Consequently, the neurometric approach may have to be adapted to be sensitive to differences in spike pattern rather than in spike rate before it can usefully be applied to the auditory cortex. An important contribution made by the present study is to show that this can, indeed, be done successfully.

The psychophysical task used here probes a form of temporal pattern judgment that is sometimes necessary to discriminate complex natural sounds. Listeners discriminated natural from ''flipped'' twitters with near-perfect accuracy when local time reversals extended

over 80 msec but performed near chance with time reversals of 20 msec or below. There are several other lines of evidence to suggest that the ordering of auditory events is limited to time frames of 20 msec or longer. In humans, accurate temporal order judgments of click trains with varying amplitude require interclick intervals greater than 20 msec (Hirsh, 1959), and those of two pure tones presented simultaneously but with asynchronous onsets require stimulus onset asynchronies of approximately 20 msec (Stevens & Weaver, 2005; Pastore, 1983; Efron, 1963). Interestingly, it has been suggested that the range of rates of change of stimulus events in spoken language, which is in the order of 2–50 Hz (Rosen, 1992), contains temporal modulation frequencies that the auditory system can detect. The limited temporal resolution of these perceptual phenomena may be attributable to limitations in the accuracy of temporal pattern codes in the underlying neural circuitry. Our analysis of temporal discharge patterns of ferret A1 units is consistent with this hypothesis. As time-reversed windows become narrower, inaccuracies in stimulus-locked firing result in a failure to reflect the fine differences in the temporal structure of the stimuli. Other previous studies on the responses of cat, monkey, and human auditory cortical neurons have observed similar limitations in the temporal resolution of cortical response patterns to discrete auditory events. For example, responses to stop-consonant syllables exhibit characteristic temporal response patterns when voice onset times are longer than about 20–40 msec, but not for shorter voice onset times (Steinschneider et al., 2005; Steinschneider, Fishman, & Arezzo, 2003; Sharma, Marsh, & Dorman, 2000; Eggermont, 1995). Here, the temporal discharge patterns of a subset of A1 units were found to discriminate natural from flipped twitters if the latter contained temporal manipulations that were at least 20 msec long. In addition, discrimination based on the stimulus-synchronized firing of A1 units approximated human psychometric performance only when the responses were read out in time bins that were 20 msec or finer. Schnupp et al. (2006) found similar optimum temporal resolutions for sampling A1 discharge patterns for twitter classification, in their case, between 10 and 50 msec. Taken together, these results suggest that psychophysical thresholds for ordering sequential acoustic events may often lie near 20 msec because the temporal discharge patterns that auditory cortical neurons use to encode these stimuli are unable to represent finer time scales accurately.

It is important to note, however, that this limited temporal resolution may not apply in higher-order cortical areas or in the context of other types of auditory tasks, such as pitch judgments, which, at the level of the auditory cortex, may be performed through a specialized set of rate coding pitch neurons (Bendor & Wang, 2005), or gap detection, where very short gaps will result in transients that affect the spectrum of the signal. Nor might it apply to the perception of transients within a

frequency channel, such as ramped and damped tones (Lu, Liang, & Wang, 2001a). Some physiological results suggest that auditory processes operating at temporal time scales either below or above the "lower pitch limit" (i.e., ca. 30–50 Hz) may be processed by separate subgroups of A1 neurons. More rapid (i.e., >50 Hz) transients within an auditory perceptual channel (i.e., within a pure-tone frequency) are encoded in the spike rates of some A1 neurons (Lu et al., 2001a, 2001b), whereas slower modulations elicited stimulus-synchronized discharges in other A1 neurons (Lu et al., 2001b). This is consistent with the present findings, where among a sample of A1 units that were driven by our complex stimuli, only a subset exhibited temporal discharge patterning that supported discrimination of the stimuli. Therefore, although temporal pattern-based neurometrics have been shown to perform better than spike rate neurometrics on a task of discriminating relatively slow (<50 Hz) changes in the envelopes of complex stimuli, spike count-based neurometrics might be expected to perform better on tasks that require the discrimination of very fast (>50 Hz) transients within a single-frequency channel. This idea does not require that these two types of encoding be carried out in separate neural populations, although the work of Lu et al. (2001b) suggests that this might be the case.

It is also important to stress that a 10- to 20-msec limit in the temporal resolution at which sequential auditory events can elicit spikes in A1 neurons would not imply that the temporal precision of the spikes in representing the onset of each of these auditory events is limited to 50–100 Hz. On the contrary, the temporal jitter in evoked spikes across repeated presentations of a stimulus can be considerably less than 10 msec for some A1 neurons (Elhilali, Fritz, Klein, Simon, & Shamma, 2004; Wehr & Zador, 2003).

We have shown that the temporal discharge patterns of some A1 units were sufficiently distinguishable that the output of the three best discriminating units was sufficient to match human psychophysical performance. Therefore, the activity of a small number of cortical units is, in principle, sufficient to account for psychophysical data, provided these units are preselected to be the most informative. However, it is far from clear whether perceptual decision-making processes do, indeed, rely on only a small number of units, and how the most appropriate neurons for the particular task at hand would be chosen. Also, results such as those of Britten et al. (1992) challenge this "best envelope" notion. They show that many middle temporal (MT) neurons have spike rates that outperform the behaving animal on a visual discrimination task, and psychometric performance, in this case, is closely approximated by the median of neurometric performance across units. Therefore, approaches that attempt to interpret the activity of pooled populations of simultaneously active neurons may be a preferable way to interpret cortical data.

Using a simple, linear pooling of simultaneously active neurons, we have shown that the responses of even independently uninformative neurons can provide a temporal activation pattern that supports discrimination of complex sounds that matches behavioral measures. Further, the temporal spiking patterns of A1 neurons appears to be quite robust in that the neurometric performance of informative neurons was largely unimpaired when their responses were pooled with noisy units. It would certainly be possible to conceive of alternative and potentially more powerful pooling algorithms than the one used here, however, the algorithm we have chosen is a biologically plausible one. A convergence of connections with simple synaptic summation would achieve this type of pooling of raw spike responses. Past neurometric studies have often used highly complex rules to pool responses in order to match psychometric findings. For instance, Purushothaman and Bradley (2005) offered a detailed examination of the effects of varying pool size, weighting of neural inputs into the pool, and pooling nonlinearity on MT neurometric performance. Furukawa, Xu, and Middlebrooks (2000) found that the responses of at least 128 randomly chosen A2 cat units are needed to reduce the median errors in a sound localization task to psychoacoustic thresholds. Although these analyses are interesting in their own right, the interpretation of such models can be problematic. For any one task, optimal pooling will require only a few cells, and optimization of performance on a slightly different task will likely require a different pool. This returns to a problem with enveloping the responses of a few best neurons, namely, that it is unclear how the brain might isolate the activity of these pools from task to task. We have shown that neurometrics based on the pooled activity of A1 neurons using a simple linear summation can closely approximate behavioral performance, provided that the neural activity is recorded simultaneously.

One might expect response properties of auditory cortical neurons to differ in anesthetized and awake animals, but there are as yet no published studies which have carried out carefully controlled comparisons of awake and anesthetized recordings to demonstrate clear and fundamental differences in response properties. In the present study, we recorded responses of auditory cortex neurons to the same stimulus set in both anesthetized and awake ferrets. Although the awake and the anesthetized datasets were not ''fundamentally different,'' there, nevertheless, appeared to be subtle differences. For example, there seemed to be a somewhat higher proportion of neurons with good neurometric performance in the awake dataset (compare Figures 5 and 12). We also observed that the PSTHs of pooled firing across units appeared to show stronger onset and offset response in the awake cortex (Figure 13). However, when we recorded the awake data, we had to switch to a different type of recording electrode as the

silicon probes used in the anesthetized prep cannot penetrate the dura. We can therefore not be certain to what extent any quantitative differences in the two datasets may reflect differences in anesthetic state, or to what extent they might be attributable to differences in the sampling of the neural population which might be inherent in the slightly different recording equipment and methodology used in these two sets of electrophysiological experiments. This potential confound makes quantitative differences in our awake and anesthetized datasets difficult to interpret, which is why we did not examine these differences here in greater detail, but focused instead on the clear and unambiguous similarities in the two datasets. In both the anesthetized and the awake data, the best envelope and pooled neurometrics matched human psychometric performance. Stimulus-locked temporal discharge patterns are present in both states, and these firing patterns are able to support the discrimination of complex sounds where a simple spike rate-based analysis fails. The fact that these features are observed regardless of anesthetic state suggests that the temporal discharge patterns described here are a fundamental aspect of complex sound encoding in the auditory cortex.

Acknowledgments

REFERENCES

Barlow, H. B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology? *Perception, 1,* 371–394.

Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature, 436,* 1161–1165.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *Journal of Neuroscience, 12,* 4745–4765.

Brugge, J. F., Reale, R. A., & Hind, J. E. (1996). The structure of spatial receptive fields of neurons in primary auditory cortex of the cat. *Journal of Neuroscience, 16,* 4420–4437.

Cariani, P. (1999). Temporal coding of periodicity pitch in the auditory system: An overview. *Neural Plasticity, 6,* 147–172.

DeWeese, M. R., Wehr, M., & Zador, A. M. (2003). Binary spiking in auditory cortex. *Journal of Neuroscience, 23,* 7940–7949.

Efron, R. (1963). Temporal perception, aphasia and déjà vu. *Brain, 86,* 403–424.

Eggermont, J. J. (1995). Representation of a voice onset time continuum in primary auditory cortex of the cat. *Journal of the Acoustical Society of America, 98,* 911–920.

Elhilali, M., Fritz, J. B., Klein, D. J., Simon, J. Z., & Shamma, S. A. (2004). Dynamics of precise spike timing in primary auditory cortex. *Journal of Neuroscience, 24,* 1159–1172.

Eskandar, E. N., Richmond, B. J., & Optican, L. M. (1992). Role of inferior temporal neurons in visual memory: I. Temporal encoding of information about visual images, recalled images, and behavioral context. *Journal of Neurophysiology, 68,* 1277–1295.

Fay, R. R., & Coombs, S. L. (1992). Psychometric functions for level discrimination and the effects of signal duration in the goldfish (*Carassius auratus*): Psychophysics and neurophysiology. *Journal of the Acoustical Society of America, 92,* 189–201.

Fritz, J., Shamma, S., Elhilali, M., & Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neuroscience, 6,* 1216–1223.

Furukawa, S., & Middlebrooks, J. C. (2002). Cortical representation of auditory space: Information-bearing features of spike patterns. *Journal of Neurophysiology, 87,* 1749–1762.

Furukawa, S., Xu, L., & Middlebrooks, J. C. (2000). Coding of sound-source location by ensembles of cortical neurons. *Journal of Neuroscience, 20,* 1216–1228.

Garde, M. M., & Cowey, A. (2000). "Deaf hearing": Unacknowledged detection of auditory stimuli in a patient with cerebral deafness. *Cortex, 36,* 71–80.

Gehr, D. D., Komiya, H., & Eggermont, J. J. (2000). Neuronal responses in cat primary auditory cortex to natural and altered species-specific calls. *Hearing Research, 150,* 27–42.

Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics.* New York: Kreiger.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143,* 29–36.

Heffner, H. E., & Masterton, B. (1975). Contribution of auditory cortex to sound localization in the monkey (*Macaca mulatta*). *Journal of Neurophysiology, 38,* 1340–1358.

Hernandez, A., Zainos, A., & Romo, R. (2000). Neuronal correlates of sensory discrimination in the somatosensory cortex. *Proceedings of the National Academy of Sciences, U.S.A., 97,* 6191–6196.

Hill, K. G., Stange, G., & Mo, J. (1989). Temporal synchronization in the primary auditory response in the pigeon. *Hearing Research, 39,* 63–73.

Hirsh, I. (1959). Auditory perception of temporal order. *Journal of the Acoustical Society of America, 31,* 759–767.

Javel, E., & Viemeister, N. F. (2000). Stochastic properties of cat auditory nerve responses to electric and acoustic stimuli and application to intensity discrimination. *Journal of the Acoustical Society of America, 107,* 908–921.

Krug, K., Cumming, B. G., & Parker, A. J. (2004). Comparing perceptual signals of single V5/MT neurons in two binocular depth tasks. *Journal of Neurophysiology, 92,* 1586–1596.

Liu, J., & Newsome, W. T. (2005). Correlation between speed perception and neural activity in the middle temporal visual area. *Journal of Neuroscience, 25,* 711–722.

Lu, T., Liang, L., & Wang, X. (2001a). Neural representations of temporally asymmetric stimuli in the auditory cortex of awake primates. *Journal of Neurophysiology, 85,* 2364–2380.

Lu, T., Liang, L., & Wang, X. (2001b). Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nature Neuroscience, 4,* 1131–1138.

Luna, R., Hernandez, A., Brody, C. D., & Romo, R. (2005). Neural codes for perceptual discrimination in primary somatosensory cortex. *Nature Neuroscience, 8,* 1210–1219.

Martinez, W. L., & Martinez, A. R. (2005). *Exploratory data analysis with Matlab.* London: Chapman & Hall/CRC Press UK.

Merchant, H., Zainos, A., Hernandez, A., Salinas, E., & Romo, R. (1997). Functional properties of primate putamen neurons during the categorization of tactile stimuli. *Journal of Neurophysiology, 77,* 1132–1154.

Middlebrooks, J. C. (1999). Cortical representations of auditory space. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 425–436). Cambridge: MIT Press.

Moore, G. P., Perkel, D. H., & Segundo, J. P. (1966). Statistical analysis and functional interpretation of neuronal spike data. *Annual Review of Physiology, 28,* 493–522.

Narayan, R., Grana, G., & Sen, K. (2006). Distinct time scales in cortical discrimination of natural sounds in songbirds. *Journal of Neurophysiology, 96,* 252–258.

Nelken, I., Chechik, G., Mrsic-Flogel, T. D., King, A. J., & Schnupp, J. W. (2005). Encoding stimulus information by spike numbers and mean response time in primary auditory cortex. *Journal of Computational Neuroscience, 19,* 199–221.

Parker, A. J., Krug, K., & Cumming, B. G. (2002). Neuronal activity and its links with the perception of multi-stable figures. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 357,* 1053–1062.

Parker, A. J., & Newsome, W. T. (1998). Sense and the single neuron: Probing the physiology of perception. *Annual Review of Neuroscience, 21,* 227–277.

Pastore, R. E. (1983). Temporal order judgment of auditory stimulus offset. *Perception & Psychophysics, 33,* 54–62.

Purushothaman, G., & Bradley, D. C. (2005). Neural population code for fine perceptual decisions in area MT. *Nature Neuroscience, 8,* 99–106.

Rauschecker, J. P. (1998). Cortical processing of complex sounds. *Current Opinion in Neurobiology, 8,* 516–521.

Recanzone, G. H., Guard, D. C., Phan, M. L., & Su, T. K. (2000). Correlation between the activity of single auditory cortical neurons and sound-localization behavior in the macaque monkey. *Journal of Neurophysiology, 83,* 2723–2739.

Relkin, E. M., & Pelli, D. G. (1987). Probe tone thresholds in the auditory nerve measured by two-interval forced-choice procedures. *Journal of the Acoustical Society of America, 82,* 1679–1691.

Rosen, S. (1992). Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 336,* 367–373.

Saberi, K., & Perrott, D. R. (1999). Cognitive restoration of reversed speech. *Nature, 398,* 760.

Schnupp, J. W., Hall, T. M., Kokelaar, R. F., & Ahmed, B. (2006). Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. *Journal of Neuroscience, 26,* 4785–4795.

Sharma, A., Marsh, C. M., & Dorman, M. F. (2000). Relationship between N1 evoked potential morphology and the perception of voicing. *Journal of the Acoustical Society of America, 108,* 3030–3035.

Steinschneider, M., Fishman, Y. I., & Arezzo, J. C. (2003). Representation of the voice onset time (VOT) speech parameter in population responses within primary auditory cortex of the awake monkey. *Journal of the Acoustical Society of America, 114,* 307–321.

Steinschneider, M., Volkov, I. O., Fishman, Y. I., Oya, H., Arezzo, J. C., & Howard, M. A., III (2005). Intracortical responses in human and monkey primary auditory cortex support a temporal processing mechanism for encoding of the voice onset time phonetic parameter. *Cerebral Cortex, 15,* 170–186.

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

Theunissen, F. E., Sen, K., & Doupe, A. J. (2000). Spectral–temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *Journal of Neuroscience, 20,* 2315–2331.

Tolhurst, D. J., Movshon, J. A., & Dean, A. F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research, 23,* 775–785.

Trussell, L. O. (2002). Modulation of transmitter release at giant synapses of the auditory system. *Current Opinion in Neurobiology, 12,* 400–404.

Wallace, M. N., Rutkowski, R. G., & Palmer, A. R. (2005). Responses to the purr call in three areas of the guinea pig auditory cortex. *NeuroReport, 16,* 2001–2005.

Wang, X. (2000). On cortical coding of vocal communication sounds in primates. *Proceedings of the National Academy of Sciences, U.S.A., 97,* 11843–11849.

Wang, X., & Kadia, S. C. (2001). Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. *Journal of Neurophysiology, 86,* 2616–2620.

Wang, X., Merzenich, M. M., Beitel, R., & Schreiner, C. E. (1995). Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: Temporal and spectral characteristics. *Journal of Neurophysiology, 74,* 2685–2706.

Wehr, M., & Zador, A. M. (2003). Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature, 426,* 442–446.

Wilson, E. B. (1927). Probable inference, the law of succession, and statistical Inference. *Journal of the American Statistical Association, 22,* 209–212.

Winter, P., & Funkenstein, H. H. (1973). The effect of species-specific vocalization on the discharge of auditory cortical cells in the awake squirrel monkey (*Saimiri sciureus*). *Experimental Brain Research, 18,* 489–504.