

25. K. W. Farrell, M. A. Jordan, H. P. Miller, L. Wilson, *J. Cell Biol.* **104**, 1035 (1987).
26. We thank R. Burns, M. A. Jordan, K. Middleton, B. Goode, and L. Wilson for critical reading of the manuscript and K. Sullivan for synthesizing the

mutagenic oligonucleotides. Supported by NIH grant GM 41751 (to K.W.F.) and Department of Defense fellowship DAAL03-90-G0155 (to C.R.S.).

8 December 1993; accepted 17 March 1994

A Panoramic Code for Sound Location by Cortical Neurons

John C. Middlebrooks,* Ann E. Clock, Li Xu, David M. Green

By conventional spike count measures, auditory neurons in the cat's anterior ectosylvian sulcus cortical area are broadly tuned for the location of a sound source. Nevertheless, an artificial neural network was trained to classify the temporal spike patterns of single neurons according to sound location. The spike patterns of 73 percent of single neurons coded sound location with more than twice the chance level of accuracy, and spike patterns consistently carried more information than spike counts alone. In contrast to neurons that are sharply tuned for location, these neurons appear to encode sound locations throughout 360° of azimuth.

It often is assumed that dimensions of perception are represented in the brain by orderly maps containing sharply tuned neurons, and such maps are well known in the visual and somatosensory cortices. Maps of auditory space have been demonstrated at the level of the midbrain (1), but physiological studies of the auditory cortex have failed to demonstrate any evidence of a space map containing sharply tuned neurons. This has been puzzling, given that auditory cortex lesions in human patients and in experimental animals result in prominent deficits in sound localization behavior. In the inferior temporal and striate visual cortices, information-theoretic analysis has demonstrated that the spike patterns of neurons, including both spike count and spike timing, can carry nearly twice the stimulus-related information of spike counts alone (2). We studied spatial coding by auditory neurons in the cat's anterior ectosylvian sulcus cortical area (area AES). We used an artificial neural network to classify the spike patterns of single neurons according to sound source location. We found that single neurons can code for sound locations throughout 360° of azimuth.

Data were obtained from 67 single units on the posterior bank of the anterior ectosylvian sulcus of eight chloralose-anesthetized cats (3). Noise bursts, 1 to 300 ms in duration, were presented in an anechoic room from 18 loudspeakers spaced in 20° steps of azimuth around the horizontal plane. Stimulus locations were varied pseu-

randomly to acquire 40 responses to stimuli at each of 18 sound source locations (for a total of 720 responses). Units responded with latencies greater than 10 ms, and their responses typically were restricted to the first 50 ms after stimulus onset; thus, nearly all the driven spikes fell within a time interval of 40 ms (4). To train the artificial neural network, we used the responses of single units on odd-numbered trials (20 trials per stimulus location). To test the network performance, we used the even-numbered trials from the same unit. This is a form of cross validation, in which training and testing utilized independent data sets.

The spike pattern elicited by a noise burst usually contained only a small number of spikes. The average spike counts at optimal loudspeaker locations were around three spikes per trial, yet even at such optimal locations, many stimulus presentations elicited no spikes at all. For that reason, we averaged the responses of a single neuron across multiple presentations of a given stimulus to estimate a spike density function. One way of computing this density function would have been to form, for each of the 18 stimulus locations, a single average of all 20 spike patterns in the training set and a single average of all 20 patterns from the test set. Under that condition, however, the network would have overtrained to idiosyncrasies in the set of 18 average training patterns. Moreover, the single set of test patterns would have provided only one pattern per stimulus location with which to evaluate network performance.

We adopted a different approach that utilized a "bootstrapping" procedure (5). We generated each bootstrapped training or test pattern by averaging across M patterns from the original data set of 20 training or test

patterns, drawn randomly with replacement. Thus, a given bootstrapped pattern might incorporate zero, one, or more copies of any particular spike pattern. Repeating this process, we generated N_{train} bootstrapped training patterns from the training set and N_{test} bootstrapped test patterns from the test set. Network performance improved with increasing N_{train} , nearing an asymptote for values of N_{train} between 10 and 20, so we used a value of $N_{\text{train}} = 20$. We used $N_{\text{test}} = 100$ so that we could assess performance in terms of percent correct. In preliminary studies, we tested values of M (the number of spike patterns included in each bootstrap) between 1 and 32 and found that network performance improved monotonically with increases in the size of M ; in the results presented here, we used $M = 20$. At each of the 18 stimulus locations, we generated a set of 20 bootstrapped training patterns and 100 bootstrapped test patterns.

The artificial neural network was a one-layer linear perceptron with 40 inputs, each corresponding to the estimated probability of a spike occurring in a 1-ms poststimulus time bin. Two output units formed a weighted sum of the 40 inputs. We used the Widrow-Hoff learning rule (6) to adjust the weights so that the output units produced the sine and cosine of the stimulus azimuth (7); the sine and cosine are proportional to the distance of the stimulus from the midsagittal plane and the interaural plane, respectively. For convenience in presenting network output, we computed the arctangent of the two outputs, which gave a continuously varying output in degrees of azimuth. We favored this simple network over more complex multilayer nonlinear nets because we found that it performed as well as more complicated nets and was less prone to overtraining.

By traditional measures of tuning based on spike counts, all of the units showed broad spatial tuning. For example, all but three of the units responded across at least 180° of azimuth with spike counts that were within 50% of their maxima, and the responses of 42% of the units never fell below 50% of their maxima, regardless of stimulus location. Despite this broad tuning of spike counts, the temporal spike patterns of most neurons varied systematically with sound location. Examples of the spike patterns elicited from one unit are shown in a raster plot in Fig. 1. Patterns varied both in spike count and spike timing. For this unit, the main feature of spike timing that varied with location was the overall latency of the burst, but the intertrial variability in the timing of spikes within patterns frustrated any effort to identify patterns simply on the basis of first-spike latency. The performance of the network, which used the timing of spikes throughout each burst to classify

J. C. Middlebrooks, A. E. Clock, L. Xu, Departments of Neuroscience and Otolaryngology, University of Florida Brain Institute, Gainesville, FL 32610-0244, USA. D. M. Green, Department of Psychology, Psychoacoustics Laboratory, University of Florida, Gainesville, FL 32611, USA.

*To whom correspondence should be addressed.

responses of this unit, is shown in Fig. 2. The network outputs clustered near the correct stimulus locations. The neuron could code sound location across all stimulus locations, not just within a particular best area. For this unit, the median error between stimulus and network response was 21.0° across all stimulus locations. In our sample of 67 units, median errors ranged from 15.5° to 73.1° and averaged 39.5° . By comparison, random chance performance would give a median error of 90° .

The ability of these single units to code sound location fell far short of the performance of a behaving cat. For instance, a cat can walk to the source of a noise burst with accuracy approaching 100% when the targets are separated by 30° (8). Nevertheless, given that localization coding by single neurons is well above chance levels, one would expect the accuracy of location coding to increase with increases in the size of the neuronal pool. The improvement that could be achieved by invoking the contri-

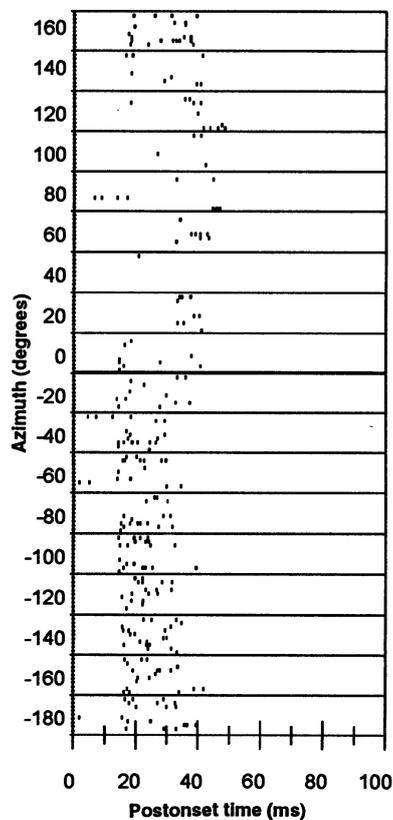


Fig. 1. Spike patterns of a single unit. Patterns are shown in a raster format in which each row of marks represents the times of occurrence of spikes elicited by one stimulus presentation at the azimuth indicated along the vertical axis. Ten such rows, representing spike patterns elicited by 10 of the 40 stimulus presentations, are illustrated at each stimulus location. The stimulus duration was 100 ms. Negative azimuths indicate the animal's left side, contralateral to the recording site.

butions of more neurons is limited by the degree of correlation in location coding among multiple neurons in the pool and by the efficiency with which information is combined across neurons (9). The absence of data on correlation between neurons prohibits estimating the size of the neuronal pool that could account for the cat's localization behavior.

The burst of spikes elicited at the onset of a noise burst showed relatively little influence of the duration of the noise burst. That is, a 1-ms noise burst elicited a 30- to 40-ms burst of spikes that was similar to that elicited by a 100-ms noise burst. For 25 units, we obtained responses to 1-, 10-, and 100-ms noise bursts. When we trained the network with responses to a single duration and tested with responses

to stimuli of the same duration, network performance was comparable across the three durations tested; across 25 units, median errors averaged 46.1° , 42.5° , and 42.4° for durations of 1, 10, and 100 ms, respectively. When trained with 100-ms responses and tested with 1- or 10-ms responses, performance was still well above chance, with average median errors increasing by only 7.8° and 10.8° , respectively.

It is a common observation that the spatial tuning of auditory units in the cortex broadens considerably when the stimulus sound pressure level is increased (10). Nevertheless, network performance was relatively robust when we trained and tested with responses to stimuli that roved between 20 and 40 dB above neuronal thresh-

Fig. 2. Network performance. The artificial neural network was trained with a set of average spike patterns bootstrapped from half of the recorded spike patterns. Each cross represents the network output based on one pattern bootstrapped from an independent test set; 100 network outputs are shown for each stimulus location. The solid line represents the circular centroids of the network responses at each stimulus location; the circular centroid is the azimuth of the resultant formed by vector addition of the individual responses and represents the central tendency of the data. The dashed line represents perfect performance.

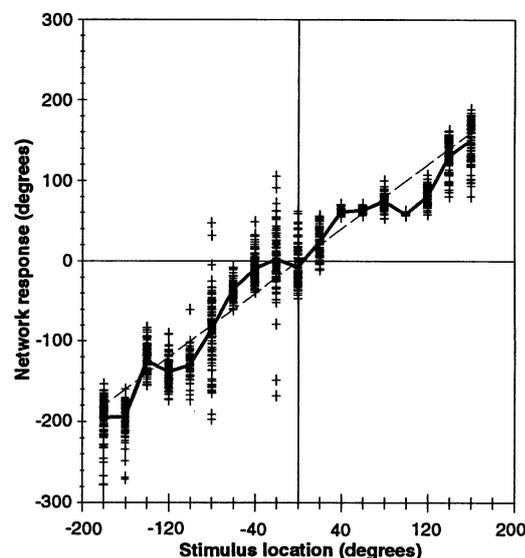
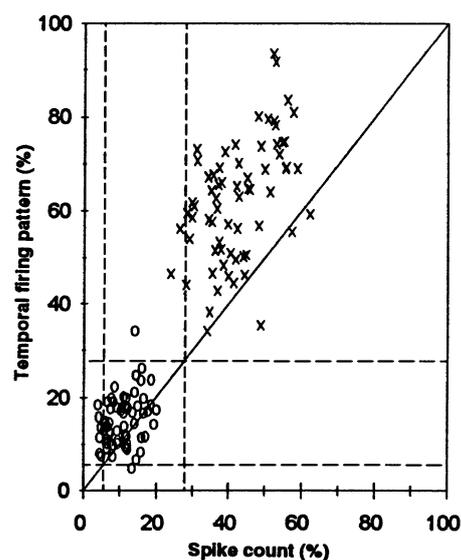


Fig. 3. Classification of spike patterns compared with classification based on spike counts. A maximum likelihood estimator was used to classify spike counts according to location. As with the network classification, we bootstrapped 20 training and 100 test patterns for each stimulus location. To compare the neural network output, which was continuously graded in azimuth, with the discrete output of the maximum likelihood estimator, we quantized the network output in 20° steps corresponding to the 20° loudspeaker separation. Each symbol represents results from one unit. The position along the abscissa represents the performance of the maximum likelihood estimator in classifying spike counts, and the position along the ordinate represents the performance of the artificial neural network in classifying spike patterns. Open circles represent the percent of bootstrapped trials in which the response was within $\pm 10^\circ$ of the stimulus location. The lower and the left-most dashed lines indicate random chance performance (5.6%). Crosses represent the percent of trials in which the response was within $\pm 50^\circ$ of the stimulus location, that is, no more than two loudspeaker positions away from the correct loudspeaker. The upper and the right-most dashed lines indicate chance performance for the 50° case (27.8%). The solid diagonal line indicates equal performance by count and pattern classifiers.



old. In a sample of the 40 units for which there is complete data, the median errors in the roving-level condition increased by an average of only 6.3° above the larger of the median errors in the two constant-level conditions. Network performance was substantially worse, however, when we attempted to train with responses to 20-dB stimuli and test with 40-dB responses.

We compared the classification based on spike patterns with a classification based on spike counts alone. The one-layer perceptron was poorly suited to classify counts because the spike count measure has only a single dimension. Instead, we used a maximum likelihood estimator, which can be shown to be an optimal classifier (11). In that procedure, spike counts from the bootstrapped training sets were used to accumulate the spike probability distribution for each stimulus location, then counts from the test sets were classified by assigning to each count the most probable stimulus location. We could not use the maximum likelihood procedure to classify temporal spike patterns because the probability distribution for that 40-dimensional measure (corresponding to the 40 1-ms time bins) could not be estimated. The network classification of spike patterns generally outperformed any classification of spike counts. A comparison of the performance of the network in classifying spike patterns with the performance of a maximum likelihood estimator in classifying spike counts alone is shown in Fig. 3. The spike count classifier was fairly successful at identifying the correct loudspeaker (open symbols), although its performance was below the random chance level for 6 of 67 units. In contrast, the pattern classifier performed better than chance for all but one unit, and the pattern classifier surpassed the spike count classifier for 50 of 67 units (74.6%). Moreover, when the spike count classifier erred, it tended to make larger errors than did the pattern classifier, so the pattern classifier gave more responses within 50° of the correct loudspeaker for all but three units.

In the optic tectum and superior colliculus, sound locations appear to be represented by a place code in which, by virtue of the spatial tuning of single neurons, a particular sound source location maps onto a particular place within the tectal or collicular map (1). No such auditory map has been demonstrated in the cerebral cortex. Our results suggest an alternative view, in which a stimulus at nearly any location activates a diffuse population of neurons. In this view, each of the active neurons signals, with its temporal firing pattern, the approximate location of the sound source, and the precise sound source location is signalled by the concerted activity of many such neurons. Each neuron is panoramic in

that it can carry information about locations throughout 360° of azimuth. Although there are many instances in sensory physiology in which particular stimuli can be shown to be represented by the levels of activity of tuned neurons, it is unrealistic to imagine that every possible stimulus dimension is represented by a corresponding class of tuned neurons. Sound location might prove to be one example of a stimulus dimension that is represented in the cortex by a temporal firing code rather than by a place code. The temporal pattern classification that we have applied in this instance may be applicable to a variety of problems in sensory coding.

REFERENCES AND NOTES

1. E. I. Knudsen, *J. Neurosci.* 2, 1177 (1982); A. R. Palmer and A. J. King, *Nature* 299, 248 (1982); J. C. Middlebrooks and E. I. Knudsen, *J. Neurosci.* 4, 2621 (1984).
2. L. M. Optican and B. J. Richmond, *J. Neurophysiol.* 57, 162 (1987); B. J. Richmond and L. M. Optican, *ibid.* 64, 370 (1990).
3. The animal protocol was approved by the University of Florida Institutional Animal Care and Use Committee. Anesthesia was induced with isoflurane in 70% N₂O/30% O₂, then an intravenous infusion of α -chloralose (25 mg/ml in propylene glycol) was begun. The cat was transferred from the isoflurane to the α -chloralose; typically 125 to 150 mg of α -chloralose was needed. Anesthesia was maintained with supplemental injections of
4. α -chloralose (~0.5 cm³ per hour). We monitored the level of anesthesia by monitoring heart rate with an esophageal stethoscope.
5. B. Efron and R. Tibshirani, *Science* 253, 390 (1991).
6. B. Widrow and M. E. Hoff, *IRE WESCON Con. Rec.* pt. 4, 96 (1960). Reprinted in J. A. Anderson and E. Rosenfeld, *Neurocomputing: Foundations of Research* (MIT Press, Cambridge, MA, 1988).
7. Efforts to configure the network with a single output unit that would generate estimated stimulus azimuth were confounded by the discontinuity in the nominal azimuth scale for loudspeakers located behind the animal. There, two loudspeakers that are only 20° apart, at -180° and +160°, correspond to a nominal difference in azimuth of 340°. In contrast, sine and cosine of azimuth are continuous across the rear midline.
8. W. M. Jenkins and R. B. Masterton, *J. Neurophysiol.* 47, 987 (1982).
9. K. H. Britten, M. N. Shadlen, W. T. Newsome, J. A. Movshon, *J. Neurosci.* 12, 4745 (1992).
10. J. C. Middlebrooks and J. D. Pettigrew, *J. Neurosci.* 1, 107 (1981); T. J. Imig *et al.*, *J. Neurophysiol.* 63, 1448 (1990); R. Rajan *et al.*, *ibid.* 64, 872 (1990).
11. J. Neyman and E. S. Pearson, *Philos. Trans. R. Soc. London Ser. A* 231, 289 (1933); D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics* (Wiley, New York, 1974).
12. Supported by NIH (National Institute on Deafness and Other Communicative Disorders) and the Office of Naval Research.

3 November 1993; accepted 24 March 1994

Specification of Pore Properties by the Carboxyl Terminus of Inwardly Rectifying K⁺ Channels

Maurizio Tagliatela, Barbara A. Wible, Roberta Caporaso, Arthur M. Brown*

Inwardly rectifying potassium (K⁺) channels (IRKs) maintain the resting membrane potential of cells and permit prolonged depolarization, such as during the cardiac action potential. Inward rectification may result from block of the ion conduction pore by intracellular magnesium (Mg_i²⁺). Two members of this family, IRK1 and ROMK1, which share 40 percent amino acid identity, differ markedly in single-channel K⁺ conductance and sensitivity to block by Mg_i²⁺. The conserved H₅ regions were hypothesized to determine these pore properties because they have this function in voltage-dependent K⁺ channels and in cyclic nucleotide-gated channels. However, exchange of the H₅ region between IRK1 and ROMK1 had no effect on rectification and little or no effect on K⁺ conductance. By contrast, exchange of the amino- and carboxyl-terminal regions together transferred Mg_i²⁺ blockade and K⁺ conductance of IRK1 to ROMK1. Exchange of the carboxyl but not the amino terminus had a similar effect. Therefore, the carboxyl terminus appears to have a major role in specifying the pore properties of IRKs.

The predicted topology of IRKs is very different from that of voltage-dependent K⁺ channels. For IRKs, hydropathy plots suggest only two potential transmembrane segments, called M₁ and M₂ (1-4); for voltage-dependent K⁺ channels, six transmembrane segments (S₁ to S₆) have been postulated (5). Segments M₁ and M₂ are separated by a region called H₅ because of

its extensive sequence similarity with a conserved region linking the fifth and sixth transmembrane segments of voltage-dependent K⁺ channels (also called SS₁-SS₂) (6). Segment H₅ is thought to line the pore of voltage-dependent K⁺ (7), Na⁺ (8), and Ca²⁺ channels (9), as well as cyclic nucleotide-gated channels (10). Because of sequence homology, H₅ has