Supporting Information

Miller and Recanzone 10.1073/pnas.0901023106

SI Text

Maximum Likelihood Estimator Errors. The maximum likelihood estimator had unsigned errors on the order of 120° and 160° in ipsilateral space, which is significantly greater than chance (90°). This is because of the relatively small sample size and the strong influence of neurons that have high activity in non-preferred directions and best locations in contralateral space. Using the maximum likelihood estimator, neurons with high spontaneous firing rates will drive the estimates to always be at the location of the best direction for that neuron. While this is a slight problem for the model, it is unlikely to be a real issue in the brain for at least 3 reasons. First, these high firing rate neurons will likely have a lesser impact when the larger population size is in use, as opposed to the relatively sparse sampling that occurs here. Second, it is not clear that these neurons form the output that influences any type of integrator/processor/discriminator at other levels of the cerebral cortex that form the percept from where the sound came. Finally, a related issue is that these neurons may in fact be inhibitory, and although they are necessary to sculpt the sharp spatial tuning functions of neurons representing contralateral space, they do not directly contribute to the percept of acoustic space.

Other Population Models Tested. In addition to the maximum likelihood estimator, we also tested 2 other population models to determine if the total spike rate could account for sound localization ability. The first was similar to that described in the motor system by Georgopolous and colleagues (1, 2). This analysis was restricted to neurons with 12 trials per stimulus (over 95% of the total sample). The average number of spikes over 300 ms was taken, and the stimulus location that provided the greatest number of spikes regardless of the stimulus intensity was assigned as the best direction. The averaged response to the remaining 63 stimuli was then normalized by this value.

The neurons were then assigned to a particular direction based on the best direction of the cell. The probability that an ideal observer would predict any particular location given that the stimulus came from one specific location can be explained by the following example. If the stimulus came from -157.5°, the probability that the ideal observer would guess location -157.5° is the averaged response of all neurons that had their best direction at -157.5°, which was usually near 1.0. The probability that the ideal observer would guess location -135° is the averaged response of the neurons with -135° as their best direction when the stimulus was presented at -157.5°. This continued on for all possible stimulus locations. The result was a relative probability that the population vector would predict each of the 16 different locations for any given actual stimulus location. An example of these functions for A1 neurons is shown in Fig. S1. Each plot shows these probability functions superimposed for ipsilateral locations (*Left*, locations -157.5° to -22.5°) and contralateral locations (*Right*, locations 0° to 180°). These functions are best illustrated for contralateral space at 75 dB SPL (Fig. S1*D Right*). Here, there is a clear peak of each individual function that corresponds to the correct stimulus location (the leftmost peak corresponds to stimuli at 0°). As the stimulus intensity decreases, the salience of these individual peaks is reduced. In contrast, in ipsilateral space, the regularity of these peaks is not as apparent at any intensity.

To convert these probabilities to actual estimates requires that the ideal observer incorporate some sort of threshold criterion. For example, if the ideal observer was to use a winner-take-all strategy, as is the case for the maximum likelihood estimator, it would simply select the peak of the probability function as the estimate. In contrast, the ideal observer could use the entire probability function and estimate different locations in proportion to the probability function values. Fig. S2 shows the mean unsigned errors of these estimates based on the threshold criterion used. In these plots, the x axis shows the threshold criterion based on the percent of the best direction prediction (the peak of the functions shown in Fig. S1). Thus, 0% indicates that the ideal observer used the entire probability function, 50% corresponds to the ideal observer considering only probabilities that were within half of the best direction prediction, and 100% corresponds to the winner-take-all strategy whereby only the peak was selected. These graphs show that, overall, this population vector model was better at estimating contralateral locations compared with ipsilateral locations at all stimulus intensities tested. However, the threshold had a strong influence on the accuracy of the observer. If the winner-take-all threshold was used, the performance was much better than the human observers, whereas when the entire probability function was used, the performance was much worse. This indicates that the ideal observer must select a particular threshold to be best correlated with the psychophysical performance, in contrast to the maximum likelihood estimator, which does not have this limitation. However, it should be noted that there are several similarities between these 2 models; for example, the estimates are more accurate for contralateral locations compared with ipsilateral locations, and area CL neurons provide the most predictive power and area MM the least.

The second population model was to use a linear pattern discriminator model that has been used previously, for example in modeling the discrimination of monkey vocalizations (3). In this case, the responses to the 16 different locations at each intensity were analyzed independently and the entire spike train was used (300-ms bins). For each trial, a single trial was selected (the test trial) and the spike rate for this trial was taken. This spike rate was then compared with the average spike rate of the remaining trials for that stimulus location, and with the average spike rate of all trials for each of the other 15 locations. The difference between the test trial and these 16 averaged locations was then calculated, and the linear pattern estimate was selected to be the location with the smallest absolute difference. In the rare instances that 2 or more locations had the same difference, each was taken as the estimate. This was repeated for all locations, and the averaged accuracy for each of the 16 locations was taken as the percentage of times that the estimator was correct given all of the times that the stimulus was presented. The logarithm of these values was then taken and summed across all neurons from that particular cortical area. The results from this analysis for the 25 dB sound pressure level stimulus are shown for each cortical area in Fig. S3. These plots are similar to those seen in Fig. S1, and even more exaggerated. In these cases, in all areas except (potentially) areas MM and R, there are clear peaks in these functions corresponding to the stimulus location. Again, a threshold has to be imposed to convert these probabilities to estimates, and if a winner-take-all strategy is used, this discriminator is virtually never wrong for locations in both ipsilateral and contralateral space (data not shown). It should be noted that the ability to discriminate these functions is increased with increases in stimulus intensity, and these plots represent the worst performance of the linear pattern discriminator. Thus, a disadvantage of this model is that, like the population vector model, a

threshold has to be imposed, and in addition it does not capture the contralesional deficits seen in lesion studies (the discriminator's performance is the same for ipsilateral and contralateral space). It is consistent with the other 2 models, however, in that the model's accuracy based on neurons in cortical area MM is significantly worse than the other cortical areas, and the accuracy based on CL neurons is the best.

Performance Along the Midline. One issue is that the model clearly has less accurate estimates along the midline compared with the psychophysical results. This is expected if both cortical hemispheres contribute to localization along the midline. While unilateral lesions have not shown a consistent deficit for midline locations, they have not been tested immediately following the lesion, and an initial deficit could have been overcome following plastic compensation by the intact hemisphere. This model predicts that there will be deficits in localizing midline locations immediately after deactivation of caudal auditory cortex, a prediction that can be tested using lesion, pharmacological, or cooling techniques. A second issue is that we are comparing human psychophysics to monkey cortical neurons. This is probably not critical, as sound localization performance between monkeys and humans using this class of stimuli is equivalent, which is particularly evident when the same laboratory, stimulus, and apparatus is used (e.g., refs. 4, 5). Finally, the neurons that were used in these simulations were not restricted to those with a particular spatial tuning profile, and many of these neurons were not spatially tuned (6). This seemed to be a reasonable approach, as any type of interpreter of these spike rates does not know a priori that the stimulus would come from 0° in elevation. It may be that some of the ''un-tuned'' cells were actually well tuned to spatial locations above or below the interaural axis, and thus would be expected to contribute to any population processing of spatial location. The relatively small number of neurons necessary to give accurate performance, 100 to 200 neurons, is similar to that calculated in extrastriate cortex (7) and is consistent with investigating only one elevation of acoustic space.

It should also be mentioned that this study is not definitive for area MM, where the sample was limited to 42 neurons. This sparse sampling could easily contribute to the relatively weak localization performance of these neurons. These data were included for 2 reasons. First, this region has only rarely been sampled electrophysiologically (8), so these data represent a significant proportion of the current published work in this area. Second, this area is quite small, so the numbers of neurons reported here are in line with the relative sizes of the different cortical areas explored in this study (e.g., refs. 9, 10). Thus, the poor localization ability by the population code may well be in line with the relatively coarse spatial representation in this area.

1. Georgopoulos AP, Schwartz AB, Kettner RE (1986) Neuronal population coding of movement direction. *Science* 233:1416–1419.
2. Georgopoulos AP, Kettner RE, Schwartz AB (1988) Primate motor cortex and free arm movements to visual targets in three-dimensional space. II. Coding of the direction of movement by a neuronal population. *J Neurosci* 8:2928–2937.
3. Russ BE, Ackelson AL, Baker AE, Cohen YE (2007) Coding of auditory-stimulus identity in the auditory non-spatial processing stream. *J Neurophysiol* 99:87–95.
4. Su TK, Recanzone GH (2001) Differential effect of near-threshold stimulus intensities on sound localization performance in azimuth and elevation in normal human subjects. *J Assoc Research Otolaryngol* 2:246–256.
5. Recanzone GH, Beckerman NS (2004) Effects of intensity and location on sound location discrimination in macaque monkeys. *Hear Res* 198:116–124.
6. Woods TM, Lopez SE, Long JH, Rahman JE, Recanzone GH (2006) Effects of stimulus azimuth and intensity on the single-neuron activity in the auditory cortex of the alert macaque monkey. *J Neurophysiol* 96:3323–3337.
7. Shadlen MN, Britten KH, Newsome WT, Movshon JA (1996) A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J Neurosci* 16:1486–510.
8. Kosaki H, Hashikawa T, He J, Jones EG (1997) Tonotopic organization of auditory cortical fields delineated by parvalbumin immunoreactivity in macaque monkeys. *J Comp Neurol* 386:304–316.
9. Hackett TA, Pruess TM, Kaas JH (2001) Architectonic identification of the core region in auditory cortex of macaques, chimpanzees and humans. *J Comp Neurol* 441:197–222.
10. Jones EG (2003) Chemically defined parallel pathways in the monkey auditory system. *Ann N Y Acad Sci* 999:218–233.

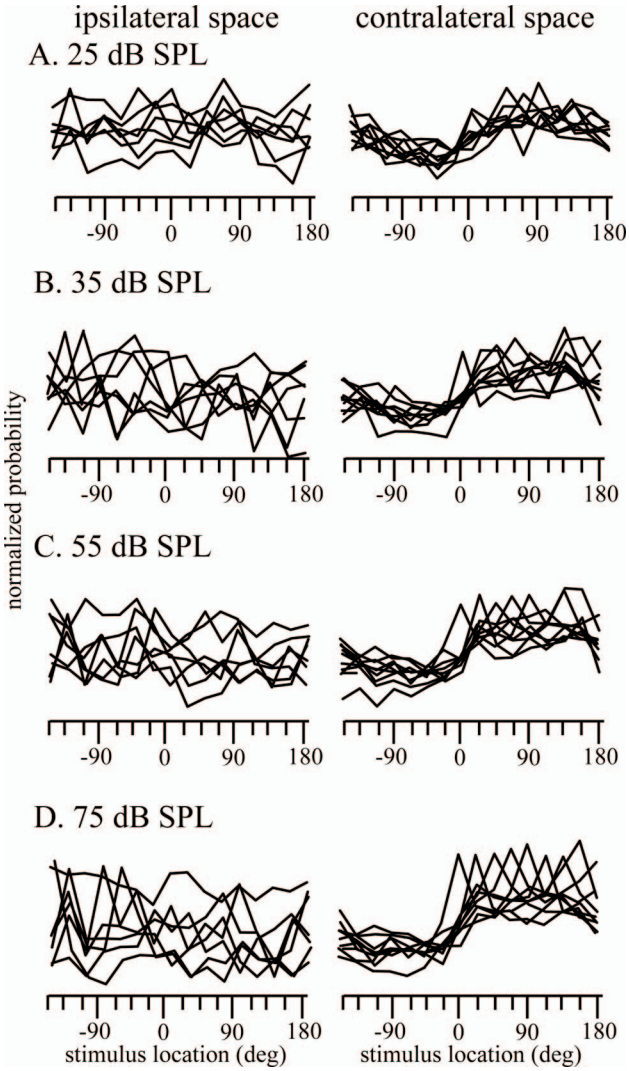

Fig. S1. Probability distributions for ipsilateral (*Left*) and contralateral (*Right*) locations using the population vector model. Each line represents the probability that an ideal observer would select each of the 16 possible locations when the stimulus actually was presented at a particular location. Peaks in these probability functions correspond to the actual stimulus location in most cases, particularly at 75 dB sound pressure level in contralateral space.

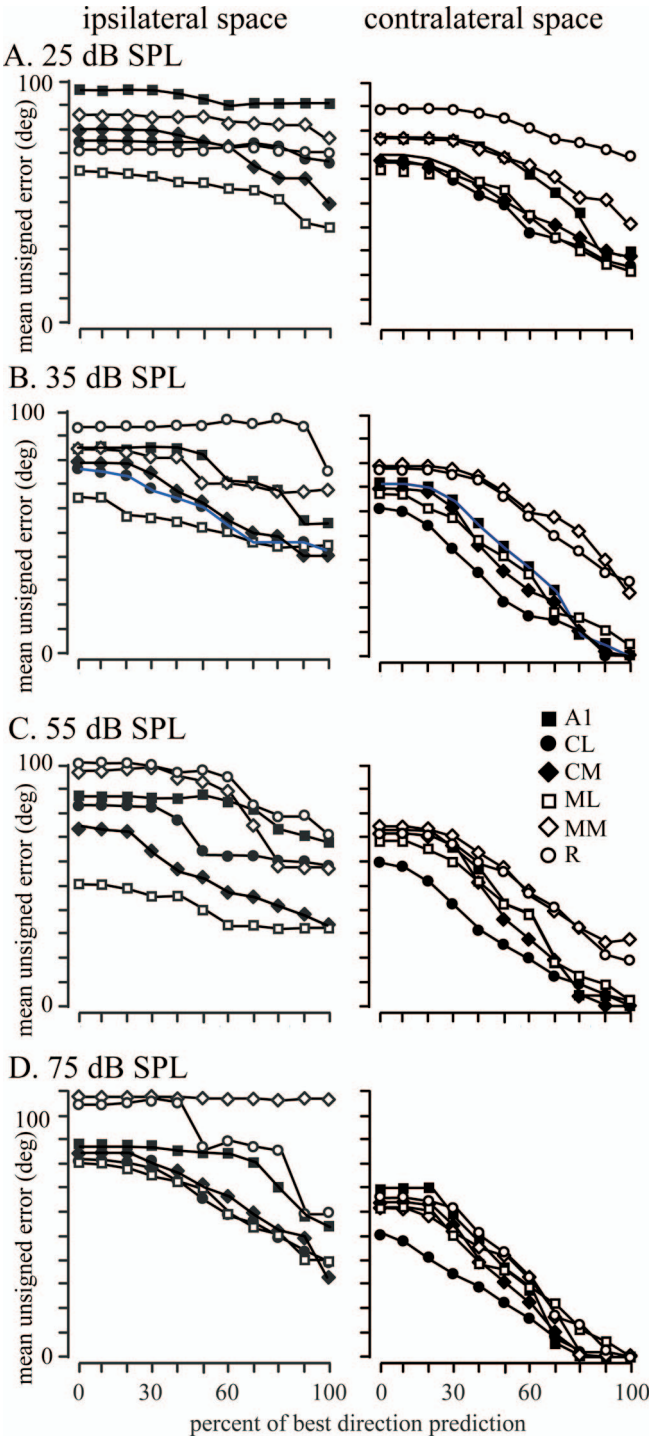

Fig. S2. Mean unsigned error as a function of the threshold based on the probability functions shown in Fig. S1. The *x* axis shows the threshold as a function of the percentage of the peak of the probability function, with 0 indicating that the entire probability function was used, 50 indicating that values within 50% of the peak were used, and 100 indicating that only the peak was included (i.e., winner-take-all). Estimates decreased dramatically with more stringent thresholds. Estimates of locations in ipsilateral space were consistently worse than those for contralateral space.

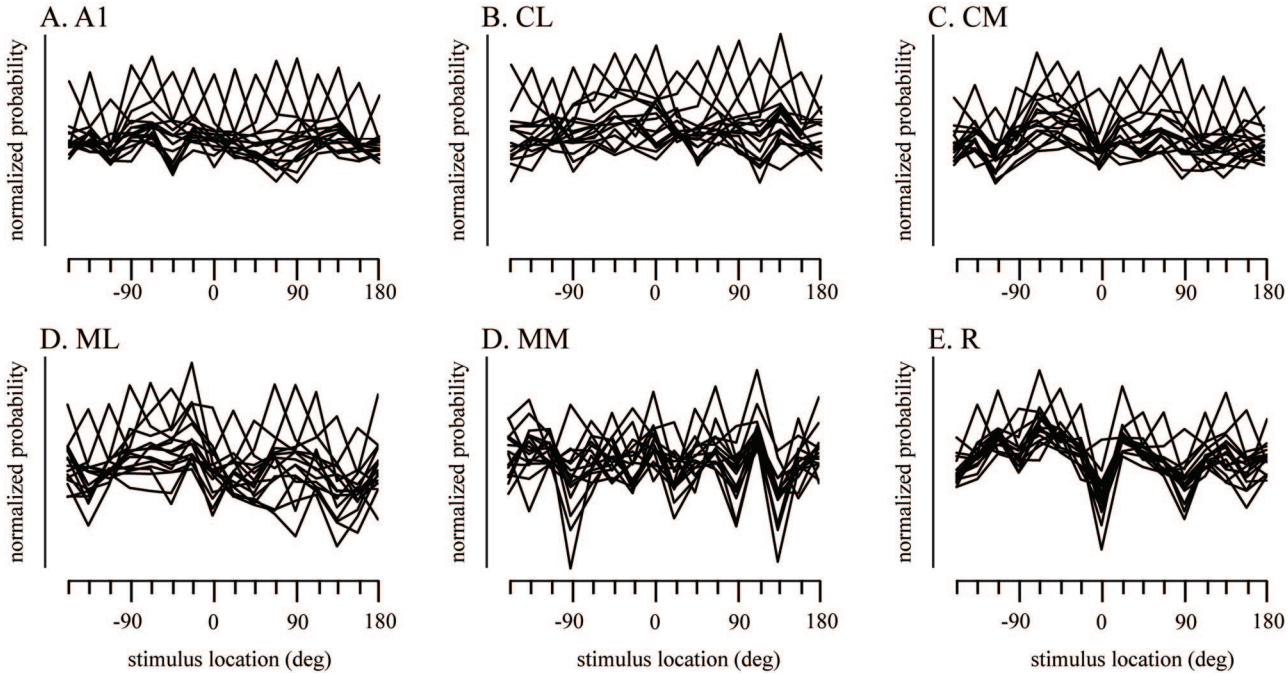

Fig. S3. Probability functions using the linear pattern estimator model. Each panel shows the results from a different cortical area, and each line shows the mean unsigned error for each of 16 different stimulus locations, similar to Fig. S1. All data were from trials using 25 dB sound pressure level stimuli. For most cortical areas, the peaks are at the actual stimulus location. A winner-take-all criterion would result in perfect or near-perfect performance for both ipsilateral and contralateral locations in all cortical areas except MM and R.