Difference in response reliability predicted by STRFs in the cochlear nuclei of barn owls

Louisa J. Steinberg & Jose Luis Pena, Department of Neuroscience, Albert Einstein College of Medicine, Bronx, NY

Abstract

The brainstem auditory pathway is obligatory for all aural information. Brainstem auditory neurons must encode the level and timing of sounds, as well as their time-dependent spectral properties, the fine structure and envelope, which are essential for sound discrimination. This study focused on envelope coding in the two cochlear nuclei of the barn owl, nucleus angularis (NA) and nucleus magnocellularis (NM). NA and NM receive input from bifurcating auditory nerve fibers and initiate processing pathways specialized in encoding interaural time (ITD) and level (ILD) differences, respectively. We found that NA neurons, though unable to accurately encode stimulus phase, lock more strongly to the stimulus envelope than NM units. The spectrotemporal receptive fields (STRFs) of NA neurons exhibit a pre-excitatory suppressive field. Using multilinear regression analysis and computational modeling, we show that this feature of STRFs can account for enhanced across-trial response reliability, by locking spikes to the stimulus envelope. Our findings indicate a dichotomy in envelope coding between the time and intensity processing pathways as early as at the level of the cochlear nuclei. This allows the ILD processing pathway to encode envelope information with greater fidelity than the ITD processing pathway. Furthermore, we demonstrate that the properties of the neurons' STRFs can be quantitatively related to spike timing reliability.

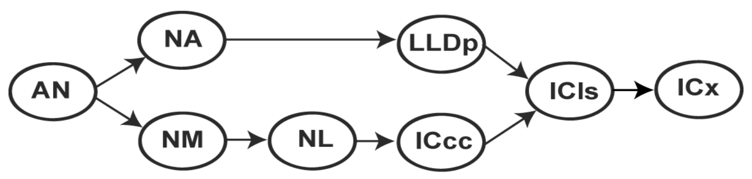

Figure 1: The ITD and ILD pathways diverge from the auditory nerve (AN) at the level of the cochlear nuclei, nucleus magnocellularis (NM) and nucleus angularis (NA), respectively. Nucleus laminaris (NL), which projects to the central core of the inferior colliculus (ICcc), and the pars posterior of the lateral lemniscus (LLDp) are the first binaural nuclei in their respective pathways. The ITD and ILD pathways converge on the lateral shell of the central core of the inferior colliculus (ICls), whose neurons project to the external nucleus of the inferior colliculus (ICx), giving rise to space-specific neurons and a map of auditory space.

Methods

Owls (*Tyto alba*) were anesthetized by intramuscular injection of xylazine and ketamine. The level of anesthesia was monitored by toe-pinch and additional injections of anesthetics were given as needed. Recording electrodes were positioned using stereotaxic coordinates for each nucleus. Neurons of NA (n=23) were recorded using tungsten electrodes. Neurons of NM (n=49) were recorded by loose-patch method with glass electrodes. NA and NM neurons' phase locking was measured using up to 300 repetitions of pure tones at each neurons best frequency, until approximately 4000 spikes had been collected. Structures were distinguished based on their tuning characteristics or phase locking abilities, as well their response latency in addition to stereotaxic location. STRFs were measured using the unfrozen noise protocol for which 500 ms segments of de novo synthesized broadband noise (1000-12000 Hz) are presented. To measure the variability and reproducibility of a neuron's response, we used the frozen noise protocol, presenting a single 500 ms broadband white noise segment repeatedly. For each of these conditions stimuli were presented until approximately 3000-4000 spikes had been collected.

Analysis

To compute the STRF, the pre-event stimulus ensemble (PESE) was compiled from all 15 ms segments of noise preceding each spike. Each of these noise segments is passed through a gammatone filter bank (91 linearly spaced filters, 1000-10000 Hz) (Keller and Takahashi, 2000; Christianson and Pena, 2007). For each frequency channel, the envelope of each noise segment is extracted using the Hilbert transform. These envelopes are averaged, giving a phase-insensitive estimate of the average power of each frequency channel preceding a spike. These values were baseline corrected by subtracting averaged envelopes derived from a shuffled PESE, called simply the stimulus ensemble (SE), for which spike trains and 500 ms noise segments were decorrelated. The STRF was used to predict the response of a neuron to a novel stimulus by convolving each frequency channel in the spectrogram of the coefficient over 500 ms of stimulus. Both PSTH and predicted firing rate were first convolved by a Gaussian kernel with 0.5 ms SD. stimulus by the corresponding channel in the STRF and then averaging across frequency channels. Comparison was done between the predicted response and the actual poststimulus time histogram (PSTH) of the neuron for frozen noise by computing the correlation

The variability and reproducibility of a neurons' response was quantified using the shuffled-autocorrelogram SAC, as described by Joris (2006).

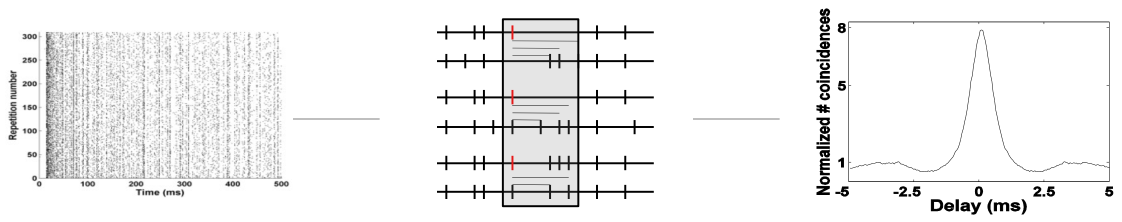

Figure 2: Computation of the shuffled auto-correlogram (SAC). For each presentation of a frozen noise segment, we recorded the neural response, a spike train. For each spike occurring in each spike train the forward time interval is measured between one spike and all spikes in all other spike trains. There are no intra-train comparisons; therefore, the effects of the refractory period of the neuron on the autocorrelogram are eliminated. The resulting SAC is normalized for firing rate, bin width of the correlogram and the duration of the stimulus, as well as the possible number of spike-train comparisons.

NA and NM units exhibit gross differences in their response properties

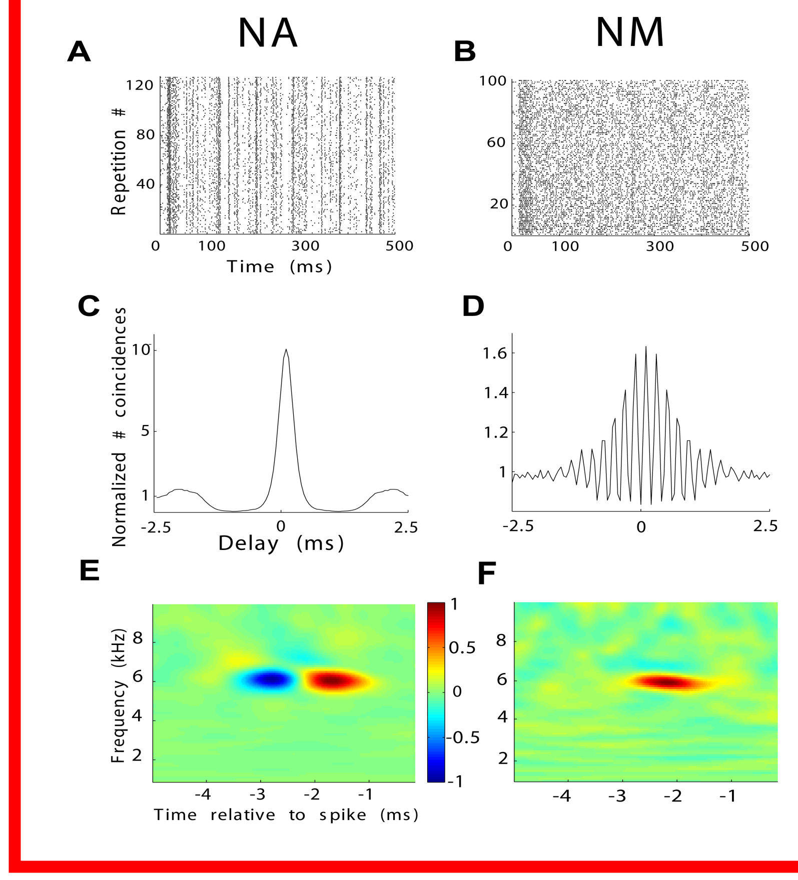

Figure 3: Representative rasters are shown for the responses of NA (A) and NM (B) neurons to repeated presentations of the same stimulus (frozen noise). (C) SAC corresponding to the NA unit in (A) with a peak height of 10. (D) SAC corresponding to the NM unit in (B) with a peak height of 1.6. The high-frequency structure in D is a result of phase locking in NM. (E) A representative example of NA unit's STRF. STRF*bf* = 6.1 kHz, STRF*bw* = 800 Hz, STRF*tw* = 0.9 ms. (F) A representative example of a NM unit's STRF. STRF*bf* = 6 kHz, STRF*bw* = 500 Hz, STRF*tw* = 0.9 ms. Notice that the NA unit's excitatory subfield is preceded by a suppressive field (blue), which is absent in NM.

Response reliability is correlated with $STRF_{sf}$

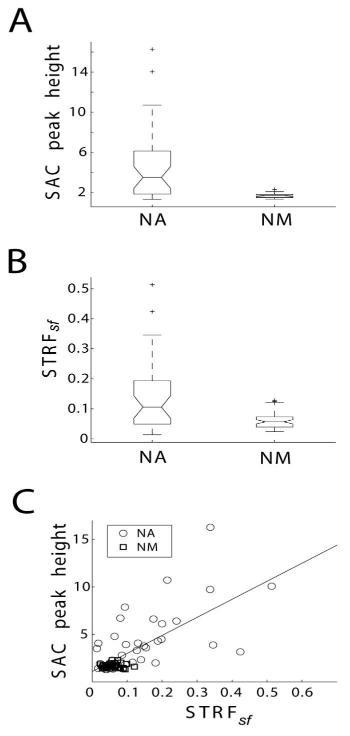

Figure 4: (A) SAC peak heights (in units of normalized number of coincidences) are significantly larger in NA units (medians significantly different by Kruskal-Wallis test, p < 0.001), indicating that responses are more reliable in NA compared to NM; (B) STRF*sf* is significantly larger in NA units than in NM units (medians significantly different by Kruskal-Wallis, p<0.001). Boxes extend from lower quartile to upper quartile of the sample, with the center line marking the median. Outliers (+) are data points 1.5 times the interquartile range of the sample. (C) STRF*sf* and SAC peak height are correlated in NA (regression: 17x + 1.9, R=0.62, p < 0.001).

STRFs more accurately predict the response of NA units than NM units

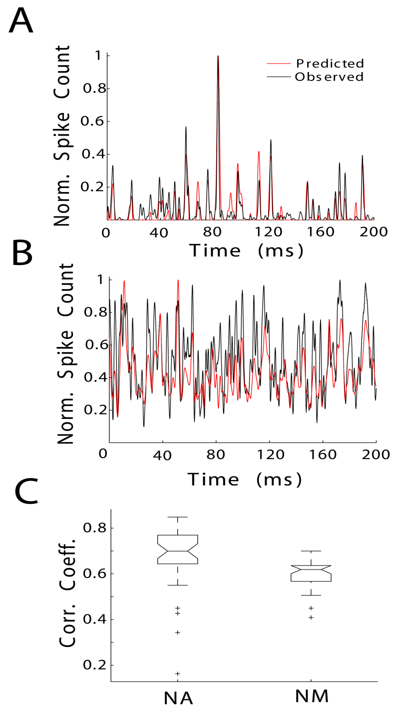

Figure 5: Example of observed (black) and predicted (red) PSTHs, given in units of normalized spike counts (Norm. Spike Count), from NA (A, correlation coefficient = 0.85) and NM (B, correlation coefficient = 0.7). (C) The correlation between the predicted and observed PSTHs (Corr. Coeff.) was compared across NA and NM. NA STRFs yielded better predictions of neural responses (medians significantly different by Kruskal-Wallis, p < 0.001). Boxes extend from lower quartile to upper quartile of the sample, with the center line marking the median. Outliers (+) are data points 1.5 times the interquartile range of the sample.

Multilinear regression analysis shows that $STRF_{sf}$ is the greatest contributor to response reliability

Coefficients of multilinear regression analysis

	Reg. coeff. (NA)	*p* (NA)	Reg. coeff. (NM)	*p* (NM)
$STRF_{sf}$	10.16	<0.001	- 0.09	0.57
$STRF_{bw}$	2.57	0.28	0.17	0.57
$STRF_{bf}$	-7.95	<0.001	-1.4	<0.001

Table 1: Regression coefficients (Reg. Coeff.) and p-values for the multilinear regression analysis performed on STRF*sf*, STRF*bw*, STRF*bf* and SAC peak height data from NA and NM.

Modeling the effects of $STRF_{sf}$ on response reliability.

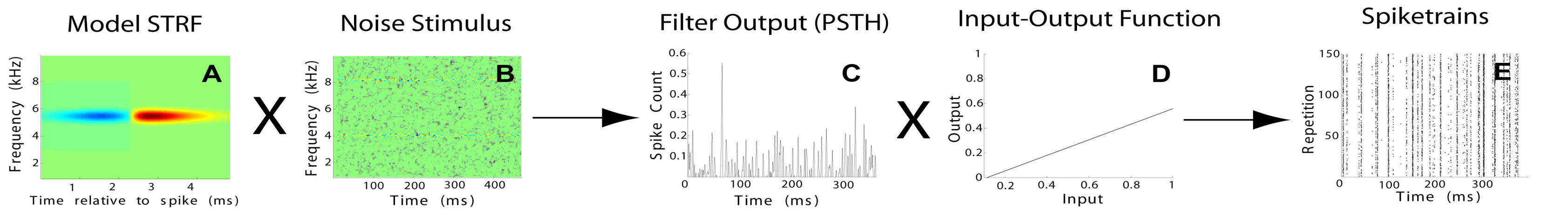

Figure 6: Schematic description of the model. Model STRFs (A) were convolved with a noise stimulus (B). This PSTH-like filter output (C) was put through a linear input-output function estimated from the *in vivo* data (D). The resulting output was used to generate spike trains (E), which were analyzed using the SAC metric.

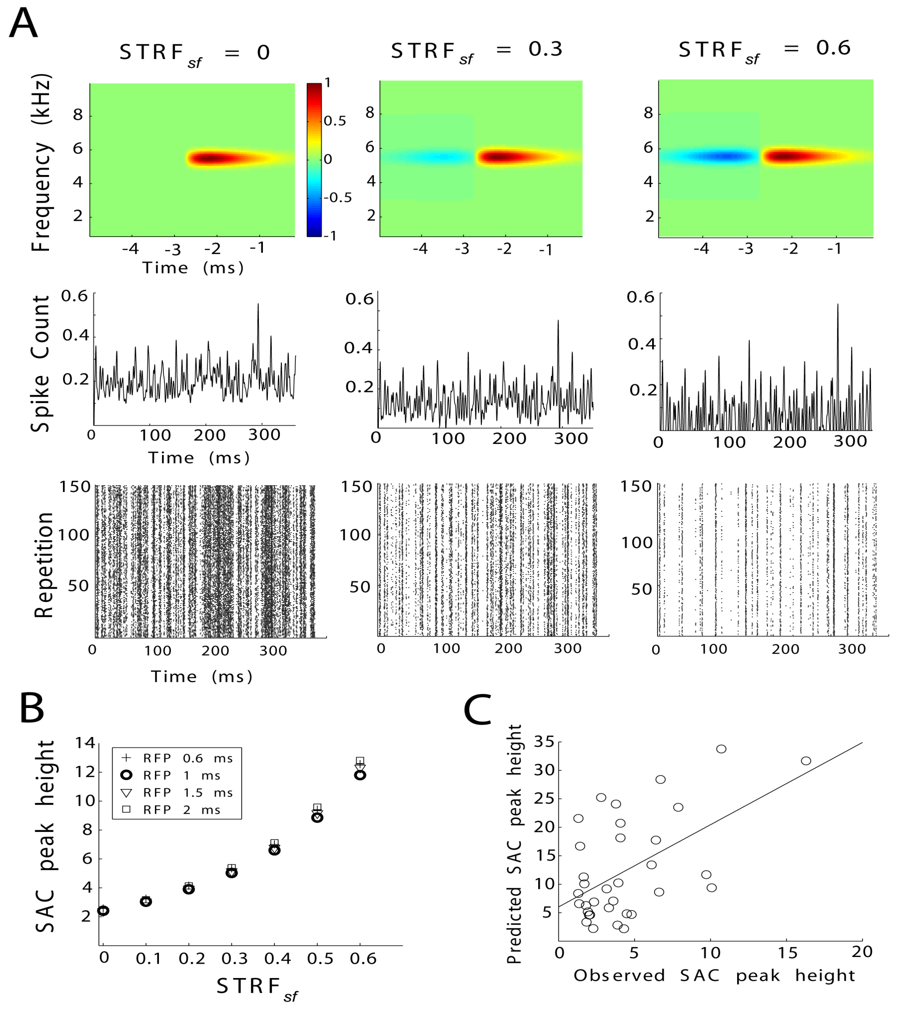

Figure 7: (A) Representative STRFs (top), PSTHs (middle) and corresponding rasters (bottom) at suppressive field magnitudes (STRF*sf*) equal to, from left to right, 0, 0.3, and 0.6. PSTHs are shown as normalized predicted spike count. Quantification reveals that pre-excitatory suppressive fields enhance response reliability (SAC peak height, B, in units of normalized number of coincidences). Different refractory periods (RFP) were tested (indicated in B). Datapoints for RFP = 1 ms are enhanced for clarity. A refractory period of 1 ms most closely approximates the estimated refractory periods across NA and NM. (C) The SAC peak height predicted by the model is correlated with the *in vivo* observed SAC peak height (R = 0.54, p = 0.001).

Conclusions

- The cochlear nucleus angularis (NA) shows greater envelope locking than its counterpart in the ITD pathway, nucleus magnocellularis (NM).
- STRFs in NA showed significantly larger pre-excitatory suppressive fields than NM, indicating that NA has more selective spectrotemporal tuning compared to NM.
- Modeling shows that greater pre-excitatory suppressive fields in the STRF enhance response reproducibility
- The first processing stage of the ILD pathway conveys spectrotemporal information with greater fidelity than the ITD pathway