

Distinct Roles for Onset and Sustained Activity in the Neuronal Code for Temporal Periodicity and Acoustic Envelope Shape

Yi Zheng¹ and Monty A. Escabi^{1,2}

¹Biomedical Engineering and ²Electrical and Computer Engineering, University of Connecticut, Storrs, Connecticut 06269-1157

Auditory neurons are selective for temporal sound information that is important for rhythm, pitch, and timbre perception. Traditional models assume that periodicity information is represented either by the discharge rate of tuned modulation filters or synchrony in the discharge pattern. Compelling evidence for an invariant rate or synchrony code, however, is lacking and neither of these models account for how the sound envelope shape is encoded. We examined the neuronal representation for envelope shape and periodicity in the cat central nucleus of the inferior colliculus (CNIC) with modulated broadband noise that lacks spectral cues and produces a periodicity pitch percept solely based on timing information. The modulation transfer functions of CNIC neurons differed dramatically across stimulus conditions with identical periodicity but different envelope shapes implying that shape contributed significantly to the neuronal response. We therefore devised a shuffled correlation procedure to quantify how periodicity and envelope shape contribute to the temporal discharge pattern. Sustained responses faithfully encode envelope shape at low modulation rates but deteriorate and fail to account for timing and envelope information at high rates. Surprisingly, onset responses accurately entrained to the stimulus and provided a means of encoding repetition information at high rates. Finally, we demonstrate that envelope shape information is accurately reflected in the population discharge pattern such that shape is readily discriminated for repetition frequencies up to ~ 100 Hz. These results argue against conventional rate- or synchrony-based codes and provide two complementary temporal mechanisms by which CNIC neurons can encode envelope shape and repetition information in natural sounds.

Key words: inferior colliculus; periodicity; repetition; amplitude modulation; temporal code; phase locking; phasic; tonic

Introduction

Speech, music, and many natural acoustic signals contain periodic waveforms that contribute to pitch and rhythm perception, and that are important for auditory object recognition, music identification, and source segregation tasks (Bregman, 1990; Moore, 1997). In humans, rhythmic context is perceived at modulation frequencies < 20 Hz, whereas pitch is dominant at modulation frequencies above this range (Moore, 1997). Pitch and rhythm perception is not exclusive to humans but is also conserved across species including nonhuman primates, birds, and cats (Chung and Colavita, 1976; Heffner and Whitfield, 1976; Colavita, 1977; Cynx and Shapiro, 1986; Tomlinson and Schwarz, 1988). The ability to faithfully encode timing information in the CNS is therefore a fundamental attribute of the hearing process.

How pitch is represented in the nervous systems has been the subject of an ongoing debate dating back more than a century (Seebeck, 1841; Ohm, 1844; Helmholtz, 1885). Central to the

debate is the question of whether pitch perception is a temporal (Schouten, 1940; Licklider, 1951) or spectral phenomenon (Goldstein, 1973; Wightman, 1973; Terhardt, 1974; Bilsen, 1977). Spectral or place theories propose that the perceived pitch is accounted for by the harmonic structure of the sound and the resulting population activity pattern on the cochlea (Shamma and Klein, 2000; Oxenham et al., 2004; Cedolin and Delgutte, 2005). Various classes of sounds, however, lack resolved harmonics, and in such cases pure spectral theories fail to account for pitch perception. This occurs, for instance, when a broadband noise is temporally modulated because spectral harmonics are not present in such sounds (Pollack, 1969; Burns and Viemeister, 1981). In this case, the temporal structure contains relevant information and the pitch is proportional to the repetition frequency of the sound.

Precise spike timing has long been recognized as a critical factor in the neuronal code for periodic sound information. In the auditory nerve and brainstem, neurons can convey information about the stimulus repetition by phase-locking to the modulation cycle at rates exceeding 1 kHz (Kim et al., 1990; Joris and Yin, 1992; Rhode and Greenberg, 1994; Rhode, 1995; Cariani and Delgutte, 1996; Cedolin and Delgutte, 2005; Dreyer and Delgutte, 2006). However, the ability to follow high modulation rates is systematically reduced at higher levels of the auditory pathway (Langner, 1992; Joris et al., 2004). For example, in the central

Received June 23, 2008; revised Nov. 11, 2008; accepted Nov. 12, 2008.

This work was supported by National Institutes of Health Grant DC006397. We thank Dr. S. Kuwada and D. Kim for numerous discussions and reviewing this manuscript, and H. L. Read for help with surgical procedures.

Correspondence should be addressed to Monty A. Escabi, University of Connecticut, Electrical and Computer Engineering, 371 Fairfield Way, U2157, Storrs, CT 06269. E-mail: escabi@engr.uconn.edu.

DOI:10.1523/JNEUROSCI.2882-08.2008

Copyright © 2008 Society for Neuroscience 0270-6474/08/2814230-15\$15.00/0

nucleus of the inferior colliculus (CNIC), the upper cutoff for synchronized activity is ~ 300 Hz (Rees and Møller, 1987; Langner and Schreiner, 1988; Rees and Palmer, 1989; Krishna and Semple, 2000; Nelson and Carney, 2007). This reduced synchrony may be related to a transformation from temporal to place-rate code in the CNIC (Langner and Schreiner, 1988).

Periodicity information is not the only temporal attribute essential to the hearing process. Neurons throughout the auditory pathway are also exquisitely sensitive to the temporal structure of the acoustic pressure waveform such that the timing of the first spike is precisely locked to the onset of a sound and contains information about the sound envelope (Heil and Irvine, 1997; Heil, 2001). Psychophysically, the temporal characteristics of the rising phase of a sound envelope contribute to the perceived intensity and timbre of a sound (Iverson and Krumhansl, 1993; Irino and Patterson, 1996). Within the context of music, the “attack” and “decay” of the rising and falling phases of musical notes contribute to the familiarity and quality of musical instruments (Risset and Wessel, 1982; Paquette and Peretz, 1997). Yet it is not clear how the shape of the envelope is simultaneously encoded with repetition information in periodic stimuli. Specifically, in the classic sinusoidal amplitude modulation (SAM) stimuli that are commonly used to study amplitude modulation (AM) coding, the envelope shape covaries with modulation frequency. At low rates, the modulation envelope has a slow rise and decay time, whereas at high rates the rise and decay times are faster. If envelope shape and repetition coding involve independent mechanisms, envelope shape information would be intermixed with periodicity information in the neuronal response, making it difficult to characterize the precise structure of the temporal code.

Here, we examined how envelope shape and periodicity information in broadband sounds are concurrently encoded in the CNIC. We use a novel stimulus paradigm combined with a shuffled correlation procedure that allows us to decompose the neuronal response pattern into a periodicity and envelope shape component. The three test sounds consist of sinusoidal amplitude modulated noise (SAMN) that contains both temporal periodicity and a well defined envelope shape, periodic noise bursts (PNB) that contain strictly periodicity information, and sine ramp noise (SRN) that contains shape information but no periodicity. Direct comparisons between these orthogonal stimulus condition demonstrates that envelope shape and periodicity appear to be encoded differentially by sustained and onset components of the neuronal output. This result argues against the concept of a modulation tuning filter on the basis of rate or synchrony alone, and suggest that information regarding the stimulus shape and periodicity are directly encoded by the temporal activity pattern.

Materials and Methods

Animal preparation and recording. Animals were housed and handled according to approved procedures by the University of Connecticut Animal Care and Use Committee and in accordance with guidelines set by the National Institutes of Health, the United States Department of Agriculture, and the American Veterinary Medical Association. All efforts were made to minimize the number of animals in this study, and alternatives were considered for the experimental and surgical procedures. Experiments were performed in an acute recording setting (48–72 h). Surgery was performed under sodium pentobarbital (25–30 mg/kg) and acepromazine (0.28 mg/kg) on adult cats ($N = 5$) with a clean outer ear canal and clear middle ears. An endotracheal tube was inserted to minimize breathing artifacts and respiratory noise. The pinnae were retracted and the animal was placed in a stereotaxic assembly with hollow earbars. A craniotomy was performed over the anterior fossae, and the overlying

occipital cortex was aspirated. The bony tentorium was then removed to fully expose the inferior colliculus. After the surgical procedure, the animal was maintained in an areflexive state via continuous infusion of ketamine HCl (6–10 mg \cdot kg $^{-1}$ \cdot h $^{-1}$) and diazepam (0.4–0.7 mg \cdot kg $^{-1}$ \cdot h $^{-1}$) in lactated Ringer’s solution. The infusion rate was adjusted according to physiologic criteria (heart rate, breathing, and reflexes). Every 12 h, the animal was given dexamethasone (1.2 mg/kg) to prevent brain edema and atropine (0.04 mg/kg) to reduce salivation.

Neuronal recordings from 142 units [65 single units (SUs); 77 multiunits (MUs)] were obtained from the central nucleus of the cat inferior colliculus. All of the data analysis herein consists of single-unit data, with the exception of the rate and synchrony analysis of Figure 4, which also contains small multiunit clusters. Epoxy-coated tungsten electrodes (4–8 M Ω at 1 kHz; A-M Systems) or glass-coated tungsten electrodes (5–15 M Ω ; 1–2 μ m tip diameter) were advanced with a Burleigh microdrive (ULN 6000; Burleigh Instruments) at an angle of $\sim 30^\circ$ relative to the sagittal plane and approximately orthogonal to the CNIC frequency band lamina (Oliver and Morest, 1984). The central nucleus was identified physiologically by presenting pure tone-pips to identify a low- to high-frequency gradient (Merzenich and Reid, 1974). Recording tracts and corresponding units that did not follow this gradient were presumed to be outside the central nucleus and were excluded from this study. Approximately two units were recorded per penetration tract. A concerted effort was made to fully sample the CNIC by sampling data along the medio-lateral and rostral-caudal aspects of the CNIC. Neuronal traces were amplified and passed through a window discriminator. Spike event times were then acquired digitally at a rate of 12.2 kHz (RA16PA, RX6; Tucker-Davis Technologies). After a unit was isolated, a frequency-response area was measured and the best frequency (BF) determined. Best frequencies spanned 1.5–40.8 kHz (mean BF, 13.9 ± 9.4 kHz; median BF, 11.1 ± 16.9 kHz).

Acoustic stimuli and delivery. All experiments were performed inside a sound isolation chamber to reduce external acoustic noise (IAC). Sounds were delivered via a closed binaural speaker system (EC1 electrostatic diaphragms; Tucker-Davis Technologies) connected to hollow ear bars (Kopf Instruments). The delivery system was calibrated for frequencies 1–47 kHz [± 3 dB sound pressure level (SPL)] with a 400 sample finite impulse response inverse filter (implemented on a Tucker-Davis Technologies RX6).

Acoustic stimuli were generated digitally at a rate of 96 kHz in MATLAB (Mathworks) and delivered via a professional audio card (RME 9652 sound card). Sounds were delivered binaurally with no interaural time difference or interaural level difference (diotically) through calibrated ECI headphones (Tucker-Davis Technologies) connected to hollow ear bars. Three stimuli were generated that allowed us to decompose the neuronal response pattern into a periodicity and envelope shape component. First, 5 s segments of PNB were generated by periodically gating uniformly distributed noise with a brief b-spline window. This sound contained strong periodicity information, although the envelope shape of the noise burst was fixed for all repetition conditions. The gating window duration was 0.25 ms with 0.05 ms rise and decay time. In two experiments, PNB repetition rates spanned 5–500 Hz ($n = 38$ recording sites) in logarithmic increments (1.4 octave steps; 15 modulation frequencies), whereas in the latter three experiments ($n = 104$ recording sites) the upper repetition frequency was extended to 1341 Hz (18 modulation frequencies). Five second segments of SAMN were generated by modulating uniformly distributed noise with a sinusoid envelope at identical modulation rates as PNB. Here, the SAMN contained periodicity and shape information; however, the envelope shape covaried with the modulation frequency of the sound. Both SAMN and PNB were presented in random interleaved order (10 repeats for each stimulus condition) with 500 ms pause between consecutive conditions. Finally, the SRN consisted of one cycle of the sinusoid envelope that was used to modulate uniformly distributed noise. SRN therefore contained shape but no repetition information. SRN stimuli were presented every 200 ms (100 randomly interleaved repeats for each condition) at the equivalent modulation frequencies as for PNB and SAMN (see Fig. 1A). All sounds were presented at a fixed level of 80 dB peak SPL. For all three stimuli,

unfrozen noise carriers were used so that only the envelope waveform was preserved from cycle-to-cycle (for PNB and SAM) or trial-to-trial.

Modulation transfer functions. Rate (rMTF) and temporal (tMTF) modulation transfer functions (MTFs) were derived for each stimulus condition using conventional methods. For the PNB and SAMN stimuli, the dataset was screened for temporal adaptation across all recordings. Adaptation was not present in the data beyond 500 ms; therefore, the initial 500 ms was discarded for all of the subsequent analysis.

The temporal precision of the synchronized response to PNB and SAMN was characterized for each unit by measuring the vector strength (VS). At each modulation frequency, the VS (Goldberg and Brown, 1969; Joris et al., 2004) was computed according to the following:

$$VS = \frac{1}{N} \sqrt{\sum_{n=1}^N \cos^2(\theta_n) + \sum_{n=1}^N \sin^2(\theta_n)}, \quad (1)$$

where $\theta_n = 2\pi t_n/T$, N is the total number of recorded action potentials, t_n is the time of occurrence of the n th spike, and T is the modulation period. The VS or synchrony index measures the degree of temporal synchronization in the neuronal response at each repetition frequency with a maximum value of 1. When the neuronal responses are plotted over a period of the modulation frequency, a VS of 1 corresponds to all the responses falling at a particular phase of the modulation period, whereas a value of zero corresponds to responses distributed evenly across all phases. The VS was tested for significance at each modulation frequency ($p < 0.001$) with the Raleigh statistic (Buunen and Rhode, 1978).

Shuffled-correlation analysis. A modified shuffled correlation analysis was developed that efficiently removes noise from the neuronal spike train and allowed us to measure the stimulus-evoked response for SAMN and PNB. The shuffled correlation (SC) was obtained as the root-mean pairwise correlations between nonoverlapping four-cycle segments of the neuronal spike train as follows:

$$\Phi_{\text{shuffled}}(\tau) = \sqrt{\frac{1}{N(N-1)} \sum_{k=1}^N \sum_{l \neq k} \Phi_{kl}(\tau)}, \quad (2)$$

where N is the number of consecutive four-cycle stimulus segments, $\Phi_{kl}(\tau)$ is the circular cross-correlation between the k th and l th segments, and $N(N-1)$ is the total number of segment pairs that are correlated. The rastergram for each modulation condition were first segmented into nonoverlapping four-cycle segments as illustrated in Figure 5B. A circular correlation was used in Equation 2 (see Fig. 5C,D) because all of the responses are contiguous at the boundaries because the stimuli are periodic. This prevents edge artifacts that would occur if a standard correlation was performed. A fast implementation of Equation 2 was obtained by noting that (see Appendix):

$$\Phi_{\text{shuffled}}(\tau) = \sqrt{\frac{1}{N(N-1)} \left(N^2 \cdot \Phi_{\text{PSTH}}(\tau) - \sum_{k=1}^N \Phi_{kk}(\tau) \right)}, \quad (3)$$

where Φ_{PSTH} is the circular autocorrelation function for the poststimulus time histogram (PSTH) over four stimulus cycles:

$$\text{PSTH}(t) = \frac{1}{N} \sum_{k=1}^N s_k(t), \quad (4)$$

and $s_k(t)$ is the neural spike train for the k th four-cycle segment. This algorithm resulted in a marked reduction in the computational load [$N + 1$ correlations compared with $N(N-1)$], which was necessary to efficiently jackknife the data across trials (Efron, 1981). Also note that $\Phi_{kl}(\tau) = \Phi_{lk}(-\tau)$, so that Φ_{shuffled} has even symmetry.

The SC analysis was performed at each modulation frequency and for each stimulus condition. To assure that the analysis resolution was conserved across all modulation frequencies, the neuronal response was sampled at a rate of 10 samples per stimulus period. The square root normalization was performed because, for SAMN, Φ_{shuffled} closely

matched a cosine model $\Phi_{\text{model}}(\tau) = (\lambda_{\text{DC}} + \lambda_{\text{AC}} \cdot \cos(2\pi f_m \tau))$, where λ_{DC} is the average driven spike rate and λ_{AC} is driven response component (see Fig. 5E). Second, this normalization conveniently expresses the neuronal output in units of spikes/second.

There are several advantages to our SC approach over direct averaging using a PSTH procedure as recently performed by Malone et al. (2007) to study the neural discharge pattern for AM sounds. First, the SC requires that we correlate spike train data across trials. This trial-to-trial correlation smooths the spike train output at the natural time scale of the temporal response before averaging, which helps remove estimation noise (see Fig. 5C). This initial “smoothing” guarantees that the correlation functions for each trial are not binary functions (0 or 1, as for the spike train data). Second, the procedure performs $N(N-1)$ averages (compared with N for a PSTH) which further helps diminish estimation noise from the response metric. The effects of this averaging and smoothing are clearly evident in the resulting SC functions for single units using SAMN (see Fig. 5D), which are often noise-free and resemble a cosine pattern for individual units. Finally, the SC removes the group delay from the neuronal response. This is useful for computing the population average SC (see Fig. 9) because it guarantees that all of the responses from different units are aligned to a common phase of zero before averaging. A simple population average PSTH would not retain this property and the distribution of response group delays would be reflected in the population response pattern (Malone et al., 2007). Removal of the group delay is particularly important in the present study because the distribution of group delays can be significantly wider than the stimulus period, especially for the high modulation rates tested (e.g., >300 Hz). If the group delay were not removed, the population SC would be blurred by the group delay distribution. Such blurring would artificially distort the population response because it would diminish the strength of phase-locked activity across the neuronal population, especially at the high rates.

Shuffled correlation modulation transfer function. At each condition, several metrics were obtained from the SC, and these were then plotted as a function of modulation frequency. First, we estimated the modulation index (MI) of the neuronal output directly from the SC as $\beta = 1 - \min(\Phi_{\text{shuffled}})/\max(\Phi_{\text{shuffled}})$. The MI was measured at each modulation frequency, and we thus represented the neuronal response as a modulation index MTF (miMTF). This allowed us to characterize the strength of the phase-locked neuronal output across stimulus conditions.

Next, we measured the match between the envelope waveform and the neuronal response estimated from the SC. This provided a measure of the timing fidelity of the neuronal response and how accurately the neuron (or population) encoded the envelope waveform shape. The envelope shape index (ESI), as follows:

$$ESI(f_m) = \frac{\langle \Phi_{\text{shuffled}}, \Phi_{\text{stimulus}} \rangle}{\sqrt{\|\Phi_{\text{shuffled}}\| \|\Phi_{\text{stimulus}}\|}}, \quad (5)$$

was estimated as the correlation coefficient between the shuffled correlation and the stimulus envelope autocorrelation function. Because the DC component of the correlation function does not provide any information about the envelope information, it was removed before computing the ESI. The SAMN ESI was obtained by using the all-order shuffled and stimulus correlations (DC component removed), whereas the PNB ESI was obtained by using the higher-order correlation functions in Equation 5 (first harmonic and DC components removed). As a rationale, note that the PNB higher-order harmonics provide the pertinent shape information that distinguishes the PNB stimulus from SAMN, and presumably its neuronal response pattern. If a neuron produces an approximately sinusoidal response pattern to PNB (e.g., as was often observed for high modulation frequencies), the direct ESI metric produces a positive ESI because the fundamental component of the stimulus correlation would be highly correlated with the neuronal SC. Thus, we modified the ESI metric for PNB by removing the fundamental component so that it more efficiently reflects disparities between SAMN and PNB shape. The validity of this approach was tested by using the all-order ESI for PNB. When the PNB response pattern exhibits a sinusoidal like pattern, the higher-order ESI was close to zero whereas the all-order ESI was positively biased as expected from the additional fundamental compo-

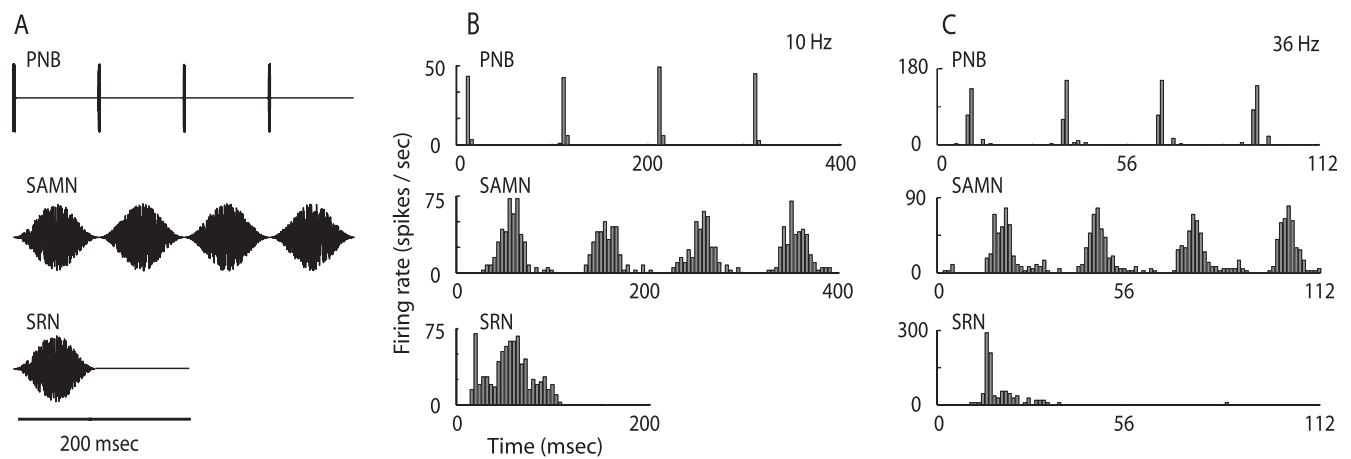


Figure 1. *A*, Proposed stimulus paradigm for decomposing the neuronal response of CNIC neurons into orthogonal stimulus dimensions of envelope shape and repetition. PNB consist of brief broadband noise bursts (250 μ s), which contain repetition information and a fixed envelope shape for all conditions. The PNB sound was varied as a function of repetition frequency, while the burst amplitude and duration remain constant. In comparison, SRN contains envelope shape information but no repetition cues because the sound is presented in isolation every 200 ms. The SRN envelope was varied in random order across the corresponding repetition frequencies used in PNB. Thus, the SRN and PNB can be thought of as orthogonal acoustic conditions. Finally, SAMN contains both shape and repetition cues and therefore shares the acoustic attributes from the SRN and PNB. Note that, mathematically, the envelope for the SAMN can be expressed as the convolution between the PNB and SRN envelope. *B*, The response of a typical mixed unit to PNB, SAMN, and SRN at 10 Hz modulation frequency. *C*, The response of the same unit to PNB, SAMN, and SRN at 36 Hz modulation frequency.

ment. In both cases, the observed trends and general conclusions were identical; however, the higher-order ESI enhanced the disparity between SAMN and PNB shape as expected from the above argument. In both cases, an ESI value of 1 indicated that the neuronal response closely matched the stimulus envelope correlation function, whereas a value near zero indicates a lack of association between the stimulus and response. To compare the shape encoding fidelity across stimulus conditions, the data were represented as an ESI modulation transfer function (esiMTF).

The data for each neuron were jackknifed (Efron, 1981) and significance estimates were obtained for the MI and ESI metrics by requiring that the parameters exceeded the levels expected for a Poisson neuron of identical spike rate. To do this, simulated Poisson rastergrams were generated at each modulation condition by requiring that the Poisson model firing rate was identical with the measured firing rate. This requirement guaranteed that the rMTF of the Poisson neuron was identical with the real neuron rMTF; however, all of the interspike interval information that would allow for a temporal code was discarded. The shuffled correlation was then computed for the Poisson neuron and jackknifed across response trials. A *t* test was then performed on the MI and ESI metrics between the real data and Poisson neuron at a significance level of $p < 0.05$.

Periodic extension shuffled correlation. To compare SRN responses with those from SAMN and PNB, we synthetically modified the SRN response and performed a “periodic” shuffled correlation (PSC). Noting that the SRN corresponds to one cycle of the SAMN, a straightforward way to predict the SAMN output with the SRN response is to assume that it consists of periodic copies of the SRN response. This procedure for predicting the SAMN response from the SRN strictly holds if the neuronal response is stationary as expected for a linear time invariant neuron (i.e., no adaptation). Although we recognize that there will be adaptation to consecutive periods in the response to SAMN, a similar adaptation pattern should in theory also be present in the PNB response. To assure that adaptation did not adversely affect our results, the first 500 ms of SAMN and PNB data were removed from all of the SC analysis. Adaptation was not visible in the data beyond this 500 ms point and both PNB and SAMN exhibited similar adaptation timescale. In contrast to PNB and SAMN, the SRN responses should contain strictly information regarding the envelope shape. We therefore sought to compare these three stimulus conditions to independently characterize periodicity and envelope shape factors of the neuronal output.

To implement the PSC, the neuronal responses to the SRN were first periodically extended by appending single response trials together into four period segments for each modulation condition (see Fig. 7). We

then used Equation 2 to generate the shuffled correlation for the periodically extended responses. To quantify how onset and sustained activity contribute to neuronal discharge pattern for SAMN and PNB, we used the SRN response to predict the neuronal response patterns observed for SAMN and PNB. To do this, the PSC procedure was performed separately for onset and sustained activity, and these were then compared with the PNB and SAMN responses. For neurons that exhibited a mixed response (onset and sustained component), it was necessary to break up the spike train into each of the contributing components (see Fig. 7) before computing the PSC. To do this, we computed the distribution of the first- and second-spike latencies for SRN and used the minimum observed second spike latency as a decision boundary to classify the onset and sustained response (see Fig. 7B). These two components were then periodically extended to four cycles before separately computing the PSC for the onset and sustained components (see Fig. 7C,D). Quantitative comparisons between the temporal discharge pattern for onset SRN activity, sustained SRN activity, PNB, and SAMN were performed by computing and comparing the esiMTFs and miMTFs for each condition. Note that the MI and ESI metrics are not sensitive to changes in firing rate and strictly account for the temporal pattern of the response. The chosen metrics are therefore not affected by adaptation in the neural response, which is expected to occur for periodic stimulation in SAMN and PNB.

Shape discrimination analysis. How much information does the CNIC neuronal population contain for discriminating shape information in periodic signals? This was addressed by comparing the neuronal activity for PNB and SAMN with a modified version of the sensitivity index (d') (Green and Swets, 1966). Because the neuronal response patterns for PNB and SAMN consisted of 40-dimensional SC vectors, it was necessary to modify the conventional d' metric so that it could accommodate these multidimensional response patterns. The sensitivity index (d') was defined as follows:

$$d' = \frac{\|\bar{\Phi}_{\text{PNB}} - \bar{\Phi}_{\text{SAMN}}\|}{\sqrt{\text{Tr}[\mathbf{C}_{\text{PNB}}] + \text{Tr}[\mathbf{C}_{\text{SAMN}}]}}, \quad (6)$$

where $\bar{\Phi}_{\text{PNB}}$ and $\bar{\Phi}_{\text{SAMN}}$ are the population averaged SC for PNB and SAMN (see Fig. 9A,B) expressed as 40 sample point vectors (i.e., 10 samples/cycle), $\|\cdot\|$ is the vector norm operator, $\text{Tr}[\cdot]$ is the matrix trace operator (i.e., the sum of the diagonal terms), and \mathbf{C}_{PNB} and \mathbf{C}_{SAMN} are the population covariance matrices for PNB and SAMN, respectively. The d' metric measures the distance between the population SCs normalized by the SC spread across the neuronal population. The d' therefore quantifies the relative distance between population activity for PNB

and SAMN while taking into account the neuronal variability in the population response. Higher values of d' indicate improved discrimination.

Results

We examined how neurons in the CNIC encode periodic information concurrently with envelope shape information with sounds that encompass the range of modulation frequencies for rhythm and pitch perception. Three sounds were generated that allowed us to characterize how repetition and envelope shape contribute to the neuronal response pattern (Fig. 1A). The response pattern to PNB, SRN, and SAMN was examined in 142 units (65 SUs; 77 MUs) in the cat CNIC. The SRN envelope (Fig. 1A, bottom panel) consists of a single cycle of a SAMN envelope (Fig. 1A, center panel) and as such does not contain repetition information. The envelope shape for the SRN, however, covaries with the modulation frequency parameter such that the envelope is shorter and more succinct for higher frequencies. For SAMN, the envelope shape covaries with the repetition of the sound similar to the SRN envelope. In contrast, PNB (Fig. 1A, top panel) contains pure repetition information, and the envelope shape does not vary despite changes in modulation frequency. Figure 1, B and C, illustrates typical responses to these three sounds at 10 and 36 Hz. This unit displayed a periodic but phasic response to PNB (Fig. 1B,C, top panel). In contrast, the SAMN discharge pattern was characterized by a sustained but modulated response that accurately reflected the SAMN envelope shape (Fig. 1B,C, center panel). Similarly, the SRN response consisted of a sharp onset response followed by a sustained response that resembled the SAMN envelope (Fig. 1B,C, bottom panel). These three conditions allow us to examine how pure repetition (PNB), shape (SRN), or combined repetition–shape factors (SAMN) contribute to the neuronal representation.

Temporal response pattern and modulation tuning

Neurons in our sample could be distinguished on the basis of their temporal discharge pattern to SRN. These could be characterized as either a temporally succinct onset response, a sustained response in which the firing pattern is modulated by the SRN envelope, or an onset followed by a sustained response pattern (mixed response). Two examples with a mixed SRN response pattern are illustrated as dot rasters in Figure 2, A and B (observed in 43 of 65 single units). The transient onset response to SRN exhibits precise spike timing at all modulation frequencies and is followed by

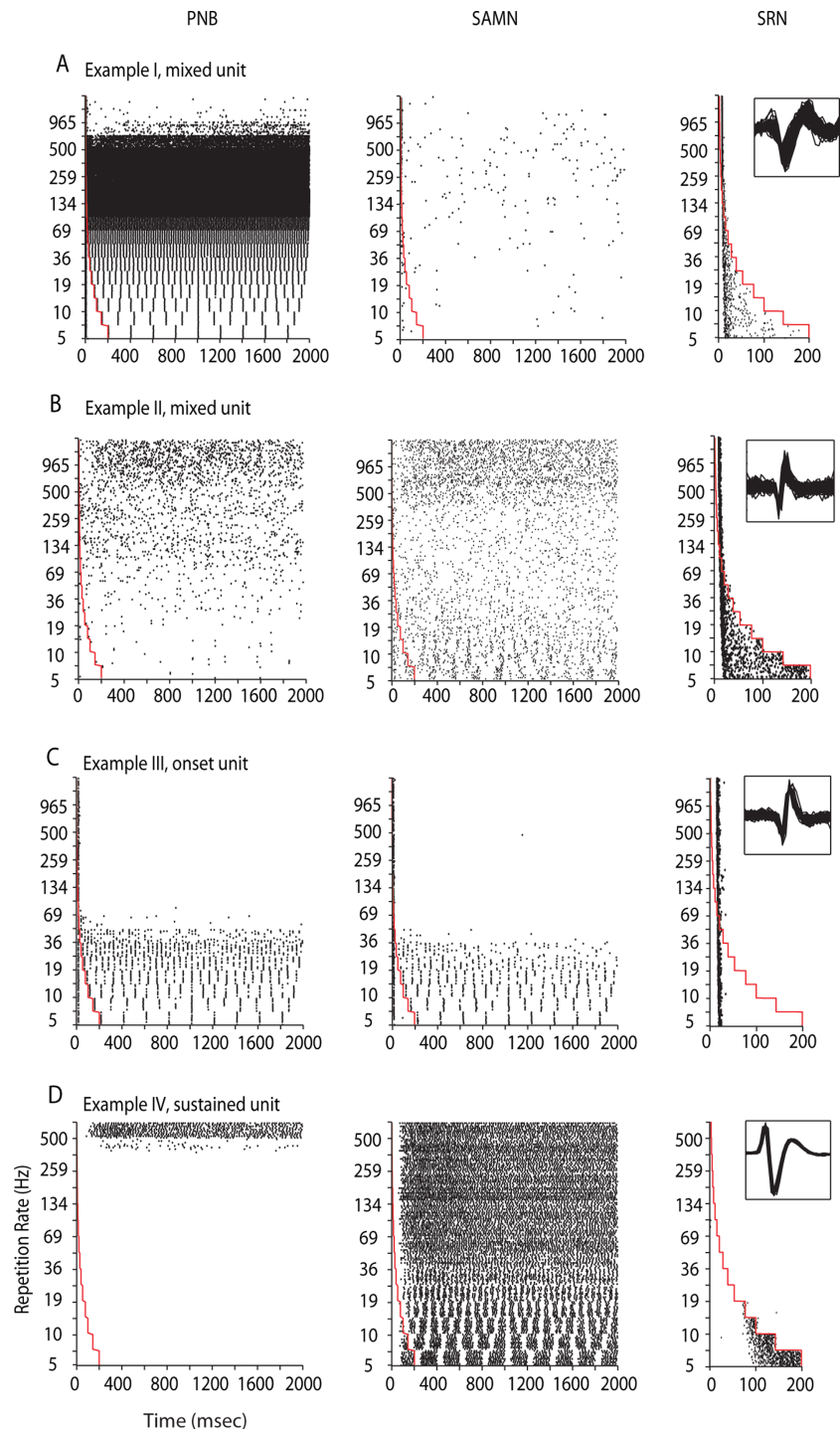


Figure 2. Spike rasters to PNB (left), SAMN (middle), and SRN (right) of four single units. Modulation frequencies were varied logarithmically from 5 to 1341 Hz for the units in A–C (18 modulation frequencies, shown along the ordinate; 10 trials per condition for PNB and SAMN; 100 trials per condition for SRN) or 5 to 500 Hz for D (15 modulation frequencies). A, B, Mixed single units exhibit an SRN response pattern with both a brief onset and longer lasting sustained component that follows the sound envelope. C, Example onset unit is characterized by a single phasic response to SRN. D, Example unit with sustained discharge pattern to SRN. PNB and SAMN rasters are shown only for the first 2 s. The red curves indicate the duration of the first cycle of response to each stimulus modulation frequency. The inset plots display an overlay of randomly chosen 100 action potential traces in a 2 ms window.

a sustained response over the stimulus duration that disappeared at higher modulation frequencies (Fig. 2A,B, right panels). The response of these neurons to PNB and SAMN are illustrated in the left and middle panels, respectively, of Figure 2, A and B. The

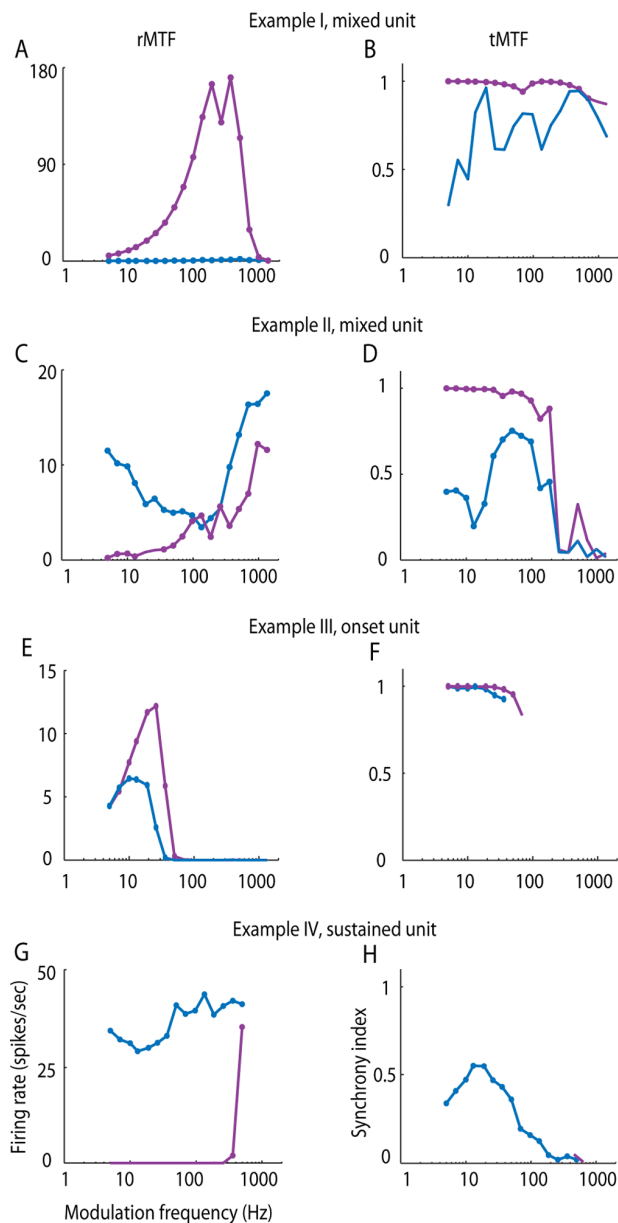


Figure 3. Comparison of modulation tuning characteristics. Example rate (left column) and synchrony (right column) MTFs obtained with SAMN (blue) and PNB (purple) for the representative neurons in Figure 2. PNB and SAMN tuning characteristics could vary dramatically; however, some units produced similar tuning patterns that were conserved for both sounds (**E, F**). Tuning properties were generally not conserved between synchrony and rate MTFs even for the same sound. Unmarked data points for the synchrony MTF (**B, D, F, H**) did not exceed the Raleigh statistic ($p < 0.001$), whereas data points with zero spikes were not plotted (**F, H**). **A, B**, MTFs of the mixed unit in Figure 2A [correlation coefficient comparisons: $r(\text{rate vs synchrony for SAMN}) = 0.54$; $r(\text{rate vs synchrony for PNB}) = 0.28$; $r(\text{PNB vs SAMN for rate}) = 0.68$; $r(\text{PNB vs SAMN for synchrony}) = -0.28$]. **C, D**, MTFs of the mixed unit in Figure 2B [correlation coefficient comparisons: $r(\text{rate vs synchrony for SAMN}) = -0.70$; $r(\text{rate vs synchrony for PNB}) = -0.89$; $r(\text{PNB vs SAMN for rate}) = 0.64$; $r(\text{PNB vs SAMN for synchrony}) = 0.70$]. **E, F**, MTFs of the onset unit in Figure 2C [correlation coefficient comparisons: $r(\text{rate vs synchrony for SAMN}) = 0.74$; $r(\text{rate vs synchrony for PNB}) = 0.50$; $r(\text{PNB vs SAMN for rate}) = 0.80$; $r(\text{PNB vs SAMN for synchrony}) = 0.93$]. **G, H**, MTFs of the sustained unit in Figure 2D [correlation coefficient comparisons: $r(\text{rate vs synchrony for SAMN}) = -0.89$; $r(\text{PNB vs SAMN for rate}) = 0.28$].

neuron in Figure 2A precisely followed the PNB stimuli (Fig. 2A, left panel) resulting in bandpass rMTF tuning and lowpass tMTF tuning (Fig. 3A, B, purple curve). However, this same unit exhibited a weak response to SAMN (Figs. 2A, middle panel; 3A, blue

curve). The neuron in Figure 2B displayed a weak albeit synchronized discharge pattern to PNB stimuli at low modulation frequencies (Fig. 2B, left panel). Although few burst stimulus events evoked action potentials, the evoked spikes were tightly aligned in time to the noise burst events resulting in high synchrony (Fig. 3D). This behavior is reflected in the tuning characteristics for this unit, which exhibited a lowpass tMTF despite an opposing highpass rMTF response pattern (Fig. 3C, D, purple curve). The SAMN tMTF of this unit by comparison exhibits a bandpass tuning profile (Fig. 3D, blue curve), whereas the rMTF exhibits reduced spiking at intermediate frequencies (~ 100 Hz) (Fig. 3C, blue curve). In this unit, the upper cutoff frequency for significant synchrony to both the PNB and SAMN stimuli were similar (Raleigh statistic, $p < 0.001$). However, at lower modulation frequencies, the PNB and SAMN responses exhibit vastly different firing rate trends indicating that the stimulus repetition alone does not account for the rMTF pattern.

Although most neurons exhibited mixed response pattern, other neurons exhibited only an onset firing pattern (11 of 65 single units). A neuron with an onset SRN response is illustrated in Figure 2C. The SRN firing pattern consisted of a brief onset component at all modulation rates (Fig. 2C, right panel). Similarly, the firing patterns to PNB and SAMN both exhibited a brief “onset”-like discharge to each modulation cycle and adapted to near zero spike rate at modulation frequencies above ~ 50 Hz (Fig. 2C, left and middle panels). For both PNB and SAMN, the unit exhibited similar tuning characteristics with a bandpass rMTF and lowpass tMTF (Fig. 3E, F). Interestingly, most onset neurons (7 of 11 units) failed to respond to the continuous SAMN, although they evoked a highly synchronized lowpass pattern to PNB. It seems that onset units are the least likely to respond to SAMN, and as such the envelope shape of the SAMN contributes little to their neuronal discharge pattern. This result is consistent with studies in bats and rats in which most of the CNIC units that failed to respond to envelope modulations were onset units (Condon et al., 1996; Shaddock Palombi et al., 2001). Despite this limitation, onset units exhibit high temporal fidelity that enables them to accurately entrain to the stimulus repetition.

Neurons with strictly a sustained SRN response pattern and no discernable onset were the least likely to occur in our sample (Fig. 2D) (7 of 65 single units). The example neuron responded robustly (allpass) to SAMN (Figs. 2D, middle panel; 3G, blue curve) although it exhibited a bandpass tuned temporal synchrony response pattern (Fig. 3H). In comparison, response could not be evoked to PNB except at high repetition frequencies (Figs. 2D, left panel; 3G, purple) at which an unsynchronized increase in spike rate was observed (Fig. 3H). A similar unsynchronized highpass discharge pattern was seen in all sustained only neurons ($n = 7$). An additional four single units displayed an off response pattern to SRN that was likely attributable to inhibition (data not shown).

In general, there was a weak correspondence between synchrony and rate MTFs (Fig. 3). The MTF filter patterns could exhibit a lowpass, bandpass, allpass, or even highpass response pattern that was not necessarily conserved for the rate or synchrony metrics (Fig. 3). In the examples shown, the rMTF and tMTF tuning patterns could differ substantially (Fig. 3, compare A, B; C, D; G, H), although some neurons did exhibit similar rate and synchrony tuning (Fig. 3, compare E, F) ($r = 0.8$; $p < 0.01$). Similarly, direct comparisons between the tuning patterns observed for PNB versus SAMN were not necessarily in agreement (Fig. 3A–D, G, H, purple curve vs blue curve).

To quantify how closely matched the tuning properties were

across different conditions (rMTF vs tMTF; PNB vs SAMN), we computed the correlation coefficient between the MTFs obtained for each of these conditions. Consistent with the examples of Figure 3, tuning properties lacked a consistent relationship across the population (Fig. 4). This was evident in the distribution of correlation coefficients between the rMTF and tMTF (Fig. 4*A,B*), which was broadly distributed (SAMN, range, -1 to 0.93 ; PNB, range, -0.98 to 0.78) and weakly correlated for both single units (mean \pm SD, SAMN, $r = 0.0 \pm 0.57$; PNB, $r = -0.01 \pm 0.50$) and multiunits (mean \pm SD, SAMN, $r = 0.41 \pm 0.43$; PNB, $r = 0.24 \pm 0.35$). Furthermore, although for some neurons the SAMN response mimicked the PNB response pattern, this was in general not true. The MTFs often differed substantially between PNB and SAMN sounds (Fig. 3, purple vs blue). The distribution of correlation coefficients between SAMN and PNB MTFs (Fig. 4*C,D*) was broadly distributed (rMTF, range, -1 to 1 ; tMTF, range, -0.86 to 1) and weakly correlated (mean \pm SD, rMTF, SU, $r = 0.19 \pm 0.49$; MU, $r = 0.20 \pm 0.47$; tMTF, SU, $r = 0.48 \pm 0.42$; MU, $r = 0.26 \pm 0.39$) indicating that the temporal or rate pattern of activity was not strongly correlated between these two types of sounds. This result is consistent with the observation that the rate best modulation frequency (BMF) obtained for PNB and SAMN were not correlated (Fig. 4*E*) (SU: Spearman's $r = 0.22 \pm 0.16$ SE; t test, $p = 0.13$, NS; MU: Spearman's $r = 0.15 \pm 0.12$ SE; t test, $p = 0.19$, NS). Because neuronal responses vary as a function of the envelope shape (as demonstrated for the SRN) (Heil and Irvine, 1997; Heil, 2001), it is possible that the observed disparity between the PNB and SAMN is accounted for by the fact that the PNB stimulus contains purely repetition information, whereas for SAMN the envelope shape is dependent on the modulation frequency and is intertwined with the periodicity information. As such, a pure rate or synchrony code as measured with rate or synchrony MTF may provide a misleading interpretation of how CNIC neurons encode stimulus repetition and shape.

The contribution of onset and sustained activity to stimulus shape and repetition

To test the hypothesis that envelope shape and repetition are concurrently encoded in the temporal discharge pattern, we used a SC analysis that allowed us to extract and compare the stimulus-evoked response for SAMN, PNB, and SRN (see Materials and Methods). The procedure for computing the SC involves circularly cross-correlating the neuronal spike trains across independent stimulus segments while averaging all correlation pairs (Fig. 5*A–D*). As can be seen, example SC functions for a typical SAMN response from a mixed neuron (Fig. 5*E*) (10, 50, 259, and 695 Hz) accurately reflect

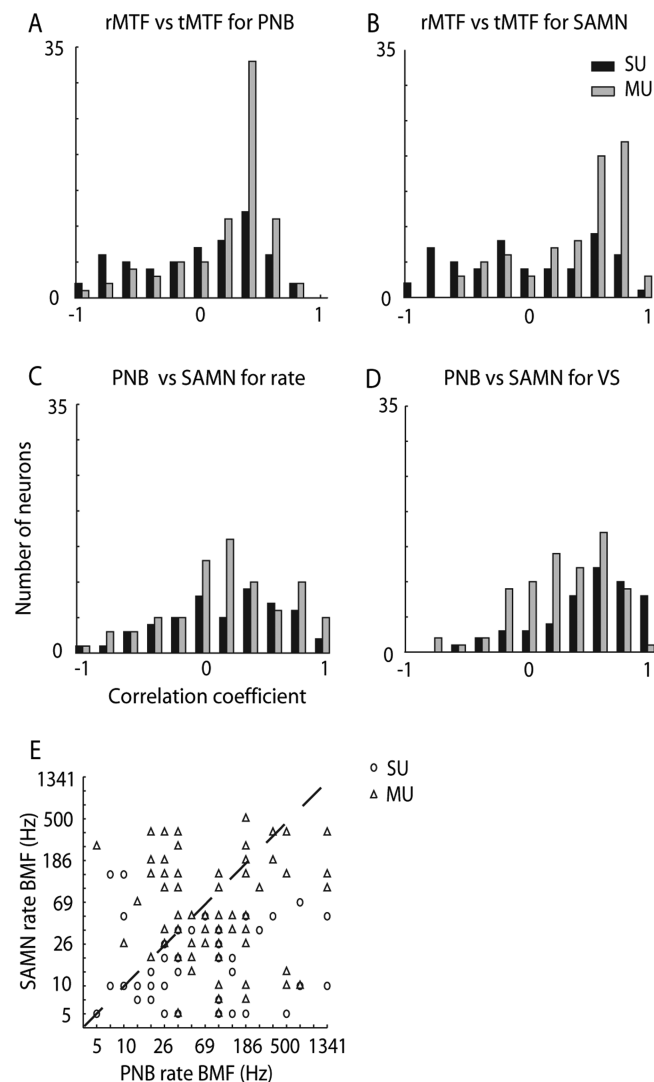


Figure 4. Lack of invariance in modulation tuning characteristics. A quantitative analysis of the MTF shape similarity was performed to compare the rate versus synchrony metrics (*A, B*) and the PNB versus SAMN sound conditions (*C, D*). Distribution of correlation coefficients between the rate and synchrony MTFs for PNB (*A*) and SAMN (*B*). Distribution of correlation coefficients comparing PNB versus SAMN for the rate (*C*) and synchrony (*D*) MTF. The black bars indicate single units, and the gray bars indicate multiunits. For all comparisons, correlation coefficients spanned the entire space suggesting that tuning was not conserved across stimuli or response metrics. *E*, Rate BMF values also indicate a weak relationship between PNB and SAMN. BMF values obtained from the rate MTF for the SAMN and PNB sounds were weakly correlated.

the sinusoidal envelope shape of the stimulus even at 259 Hz. Furthermore, the peak-to-peak response amplitude as assessed by the SC is tuned for modulation frequency. This trend was measured by computing the response modulation index (see Materials and Methods) for each condition, which we then used to plot the miMTF of each neuron (Fig. 5*F*). The bandpass tuned miMTF of this example indicates that the neuron most efficiently used its spiking activity ~ 50 – 100 Hz. For each neuron, we also estimated how well the neuronal response pattern encoded the stimulus waveform shape. The ESI (see Materials and Methods) was used to quantitatively compare the similarity of the stimulus envelope autocorrelation to the response SC at each modulation frequency (Fig. 5*G*). For this example, the ESI is ~ 1 for frequencies < 360 Hz indicating that the spike-timing pattern for this neuron accurately reflected the stimulus envelope shape for this range of frequencies. The ability of the neuron to follow the stimulus envelope drops precipitously at modulation frequencies greater than ~ 400 Hz as reflected in the esiMTF.

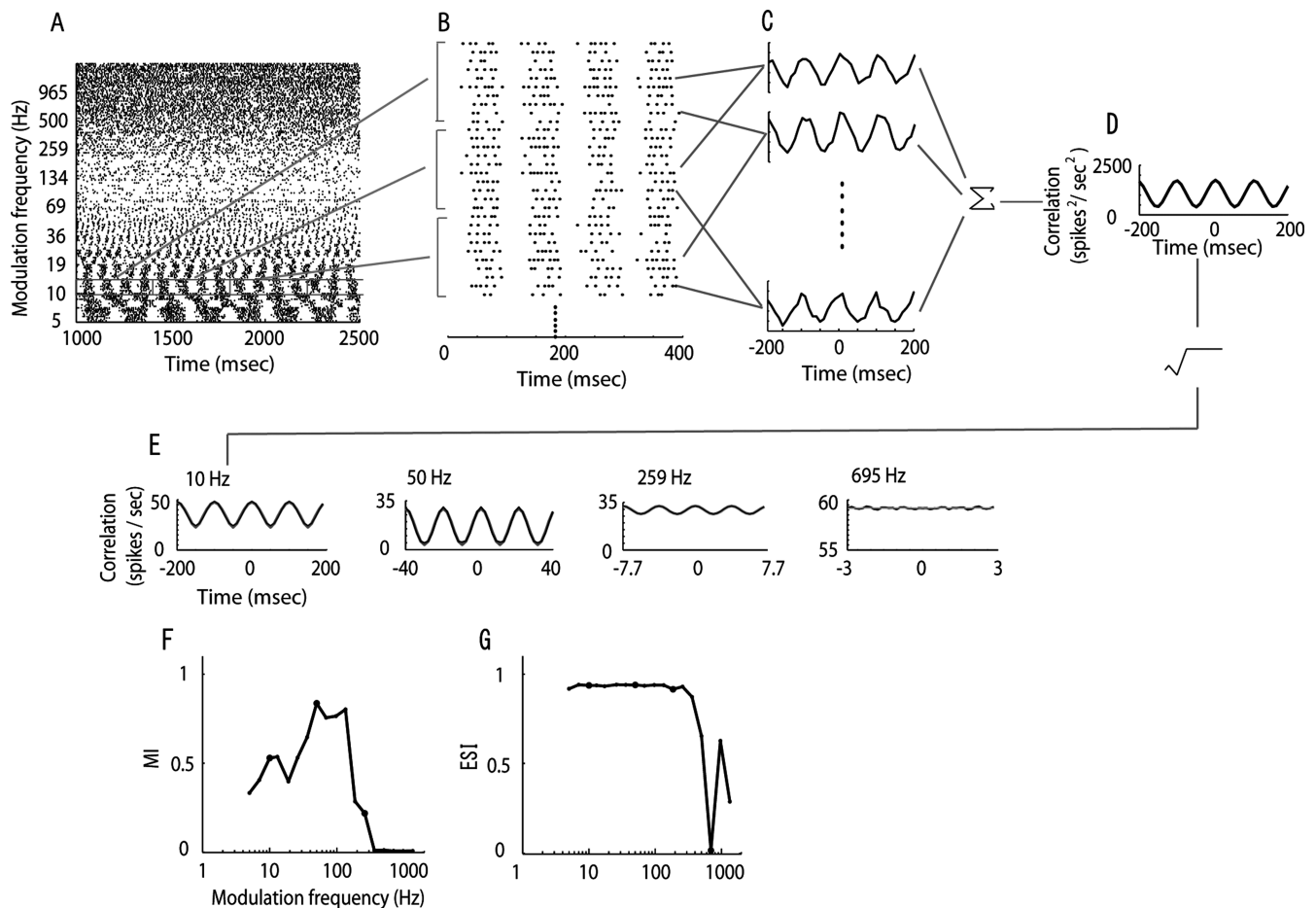


Figure 5. Quantitative analysis of enveloped phase-locked activity obtained with a modified shuffled correlation procedure (see Materials and Methods). The shuffled correlation is illustrated with the SAMN sound for a typical neuron. **A**, Spike raster of a typical response to SAMN is shown at all modulation frequencies tested. Modulation frequencies increase vertically, and 10 response trials are shown per condition. **B**, The data in **A** was first partitioned into independent four-cycle segments for each stimulus modulation frequency. **C**, **D**, Pairwise shuffled correlations were then performed across all possible segments combinations (**C**), and the average shuffled correlation was obtained by averaging across all trials (**D**). **E**, Example shuffled correlations are shown at 10, 50, 259, and 695 Hz. Note that, for this neuron, the neuronal correlation (black) and the stimulus correlation (gray) are tightly superimposed. For each of these conditions, we then measured the ESI and MI (see Materials and Methods) and represented the results as a modulation transfer function for each of these parameters (**F**, **G**). Superimposed dots on each of the MTFs represent data points for the selected conditions in **E**. The MI quantifies the normalized strength of the neuronal output, whereas the ESI measures how faithfully the neuronal discharge pattern encoded the stimulus envelope.

The SC of a mixed unit is shown for PNB and SAMN in Figure 6, *A* and *B* (same unit as Fig. 2*B*), along with the corresponding miMTF and esiMTF (Fig. 6*C*,*D*). For this unit, the SAMN shuffled correlation accurately reflected the sinusoidal envelope shape for all SAMN modulation frequencies between 10 and 500 Hz (Fig. 6*B*; *C*, blue, esiMTF), whereas for PNB the SC pattern consists of brief periodic events that accurately reflect the PNB envelope structure up to ~259 Hz (Fig. 6*A*; *C*, purple, esiMTF). The miMTF exhibited a lowpass pattern for PNB and a bandpass pattern for SAMN (Fig. 6*D*).

The examples of Figure 6 bring up the possibility that the differences in the response activity pattern for PNB and SAMN may be attributable to differences in the envelope shape of the two sounds. If this is so, the SRN response pattern should account for the differences in the responses to these two periodic stimuli. Hypothetically, for a linear time-invariant neuron the SAMN discharge pattern would consist of periodic copies of the SRN discharge. To compare the SAMN response with the observed discharge for onset and sustained components for the SRN stimulus, we therefore created a PSC by assuming that the SAMN response consisted of periodic copies of the SRN response pattern (see Materials and Methods) (Fig. 7). The periodic extension

procedure was performed independently for the onset and sustained SRN discharge components (Fig. 7*C*,*D*) and the PSC was then computed using the procedure of Figure 5 for each component.

The neuronal discharge pattern for the onset and sustained activity is shown in Figure 6*E*–*H*. The PSC analysis shows that the sustained SRN response component accurately represents the sound envelope shape at low frequencies (~10–69 Hz) (Fig. 6*E*, *G*) but deteriorated at higher rates. Note that, for this unit, sustained responses were present only for frequencies less than ~100 Hz, and thus the sustained response PSC analysis could only be performed for this range of frequencies. In comparison, the onset SRN component of the same unit resembled brief periodic events at low frequencies (Fig. 6*F*) (5 Hz) but was closely matched to a sinusoid waveform at higher modulation frequencies (Fig. 6*F*) (186–695 Hz). Thus, the temporal structure of the sustained SRN response closely resembles the observed pattern for SAMN at low frequencies, whereas the onset activity dominates and appears to provide a better representation of the envelope at high modulation frequencies. This is evident in the esiMTFs and miMTFs for the onset and sustained SRN components (Fig. 6*G*,*H*).

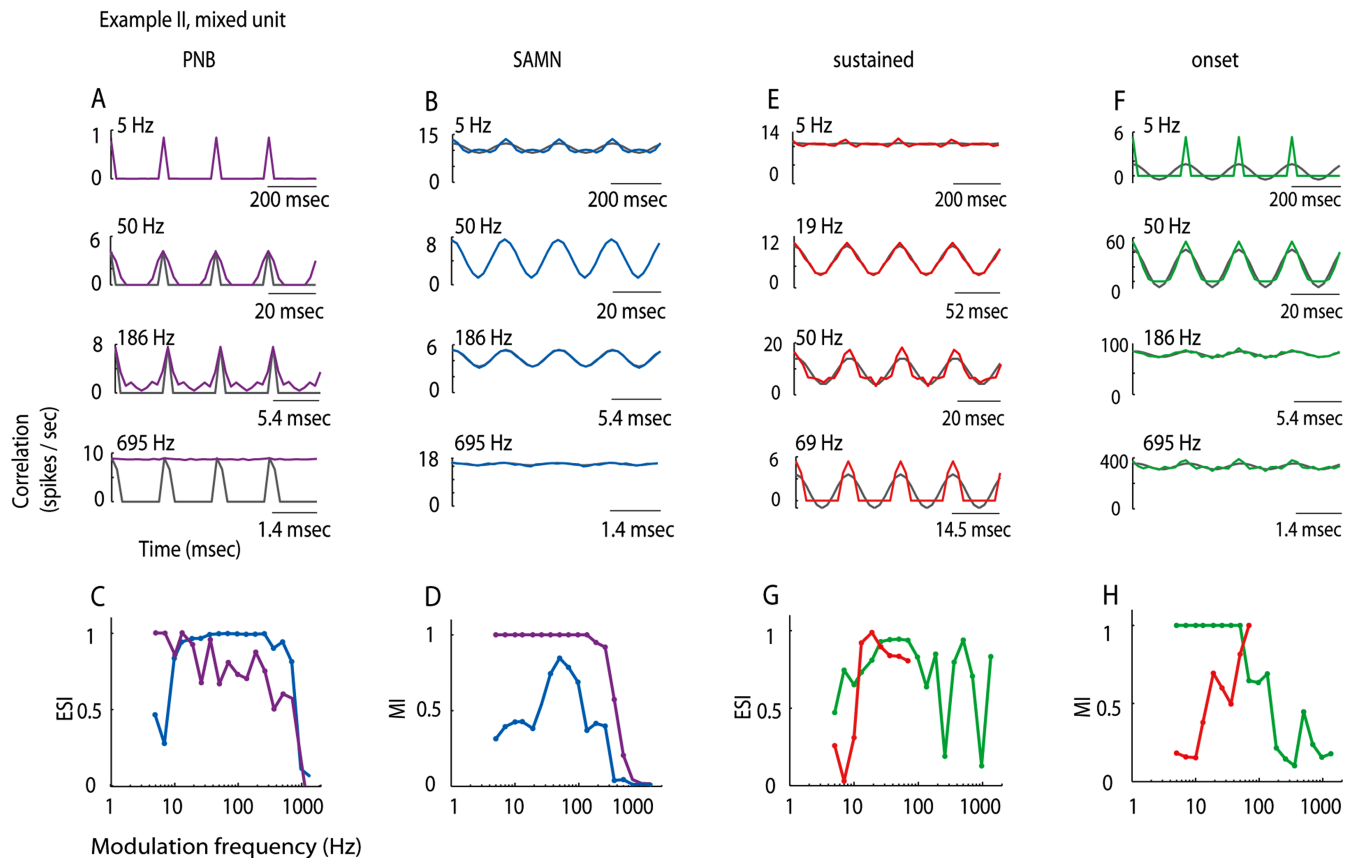


Figure 6. Shuffled correlation analysis for a mixed single unit (same unit as in Fig. 2B). **A, B**, Shuffled correlation function for PNB (purple), SAMN (blue), and their stimulus envelope autocorrelation (gray) at 5, 50, 186, and 695 Hz. **C, D**, esiMTF and miMTF for PNB (purple curve) and SAMN (blue curve). **E, F**, Periodic shuffled correlation function for sustained SRN response (red) at 5, 19, 50, and 69 Hz. The shuffled correlation is shown for onset SRN activity (green) at 5, 50, 186, and 695 Hz. The envelope autocorrelation of the stimulus is overlaid for reference (gray). **G, H**, esiMTF and miMTF for the sustained (red curve) and onset (green curve) SRN response. The dots on the curves of **C, D, G**, and **H** indicate significant results (t test, $p < 0.05$) relative to a Poisson neuron with identical firing rate. Conditions with zero spikes were excluded from the analysis and not plotted.

In contrast to the mixed unit of Figure 6, the SC of an onset neuron consisted of brief periodic events for both SAMN and PNB (Fig. 8A, B) that was not consistent with an envelope shape code (Fig. 8C, purple curve, esiMTF). This neuron responded only at low modulation frequencies (< 50 Hz) for both PNB and SAMN and it exhibited high temporal precision with a modulation index of 1 over this range of frequencies (Fig. 8D). The SRN response for this unit consisted of brief onset responses with no sustained component. This is evident in the PSC analysis for SRN (Fig. 8E–G) that exhibits a temporally succinct periodic correlation pattern that is dominant at low frequencies. Thus, it appears that this neuron is capable of encoding the exact occurrence of the stimulus event but provides a poor representation for the stimulus shape. Together, these two examples (Figs. 6, 8) demonstrate that neurons in the CNIC are capable of encoding the envelope shape and periodicity and suggest that onset and sustained responses could play separate functions.

How does onset and sustained activity in the CNIC population contribute to the analysis of the sound envelope shape and stimulus repetition? To address this question, we computed population averaged shuffled correlation for all stimuli and sound repetition conditions (Fig. 9) (see Materials and Methods). Figure 9, A and B, displays the averaged SC for PNB and SAMN from 65 single units. For PNB, the population SC consisted of a temporally precise periodic pattern (Fig. 9A, purple) that resembled the correlation for the PNB stimulus (Fig. 9A, gray) and was significant up to 259 Hz (compared with a Poisson population of

neurons with equal spike rates: bootstrap t test, $p < 0.05$). This was particularly true at rates < 134 Hz at which the SC consisted of brief phasic responses such that the ESI and MI were close to 1 (Fig. 9E, G, purple curve). Similarly, for SAMN, the population SC was closely matched to a cosine waveform up to ~ 500 Hz (Fig. 9B). This was evident in the population envelope shape index, which was significant (compared with a Poisson population of neurons with equal spike rates: bootstrap t test, $p < 0.05$) and close to 1 up to 500 Hz (Fig. 9E, blue curve). The strength of the SAMN response (MI), in comparison, was tuned with a maximum modulation index at 26 Hz and was not significant for the neuronal population beyond 259 Hz (Fig. 9G, blue curve) (bootstrap t test, $p < 0.05$). Thus, the CNIC population activity exhibits a high degree of temporal acuity and contains significant envelope information for modulation rates to ~ 250 Hz.

We next compared the average PSC from onset and sustained responses in the SRN stimulus and compared these directly with the observed population SC for PNB and SAMN. To do this, we estimated the MI and ESI directly from the population average PSC (Fig. 9C, D). The modulation and envelope shape index trends for onset and sustained response components exhibited a unique pattern that accurately predicted the observed results for the SAMN stimulus. The MI obtained from onset responses exhibited a lowpass trend in which the MI was 1.0 for frequencies below ~ 50 Hz (Fig. 9G, green curve) and which closely matched the observed pattern for onset activity from PNB (Fig. 9G, purple curve) (multiway ANOVA, $p = 0.53$, NS). In comparison, the MI

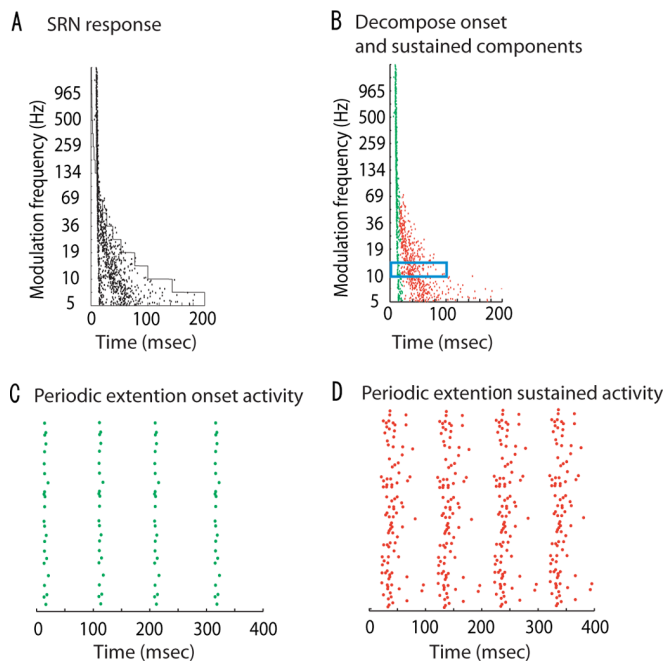


Figure 7. Predicting periodic responses to SAMN from SRN responses. **A**, The example dot rasters obtained for SRN at all modulation frequencies tested (5–1341 Hz). **B**, The dot raster is first decomposed by separating the onset (green) and sustained (red) responses (see Materials and Methods). To predict the neuronal activity for SAMN, we assume a time-invariant neuronal discharge, so that the predicted SAMN response pattern consists of periodic copies of the SRN discharge pattern. A periodic extension is performed by randomly interleaving SRN responses from each trial to generate a four-cycle dot raster. This procedure is illustrated in **C** for the 10 Hz condition shown in **B** (data inside the blue contour). **C**, **D**, Predicted dot rasters are illustrated separately for the onset component (green) and the subsequent sustained response component (red).

for the sustained activity peaked at ~ 0.5 , although it exhibited a bandpass pattern with maximum at ~ 100 Hz (Fig. 9G, red curve) that was closely matched to the bandpass pattern observed for SAMN (Fig. 9G, blue curve) (multiway ANOVA, $p = 0.10$, NS). Thus, sustained activity in the CNIC population accurately replicates the observed population pattern for SAMN, whereas onset activity fully accounts for the observed PNB trends.

The combined ESI patterns from onset and sustained activity to SRN (Fig. 9F) could explain the observed ESI trend for SAMN. Although the SAMN ESI was near 1.0 up to 500 Hz (Fig. 9E, blue curve), the combined envelope shape index for the sustained SRN response was near 1 only at low modulation frequencies (< 100 Hz) and deteriorated at higher modulation frequencies (Fig. 9F, red curve). In contrast, the combined onset activity from mixed and onset neurons to SRN exhibited a highpass pattern with maximal envelope shape index at high modulation frequencies (Fig. 9F, green curve). A distinct crossover boundary between these two response categories was clearly observed around 100 Hz. The opposing patterns for onset and sustained activity suggest that these could differentially contribute to high and low modulation frequencies and thus could together account for the broad ESI range observed for SAMN. Sustained activity appears to faithfully represent the envelope shape at low modulation frequencies (< 100 Hz), whereas onset responses represent the envelope shape at higher modulation frequencies.

The observed temporal response patterns for PNB and SAMN and the corresponding patterns for onset and sustained SRN activity suggest that precise temporal activity plays a role in the analysis of envelope shape. Furthermore, it suggests that onset

and sustained activity could underlie some of the differences in the spiking patterns for PNB and SAMN. We therefore asked whether the observed neuronal discharge patterns could be used to discriminate envelope shape information. We addressed this question by computing the psychophysical metric d' (Green and Swets, 1966) directly from the population SC patterns between SAMN and PNB for each modulation condition (see Materials and Methods). The resulting d' curve exhibits a lowpass pattern with $d' > 1$ for frequencies up to 97 Hz (Fig. 10). Thus, based on the population temporal activity pattern, differences in the PNB and SAMN envelope shape could be easily discriminated up to ~ 100 Hz.

Discussion

These results challenge long-standing theories of repetition and modulation coding in the auditory midbrain that have traditionally assumed that rate-tuned periodicity filters convey essential information about the temporal structure of the stimulus (Langer, 1992; Joris et al., 2004). For such a population-coding scheme to work, it is a prerequisite that the rate MTF of each neuron exhibit an invariant response to the stimulus periodicity regardless of the exact shape of the stimulus envelope or its spectral content. Rate and synchrony MTFs in our data could indeed be tuned; however, they were in general not consistent with this simple scheme because they were mostly different across sound conditions for the same neuron (Figs. 3, 4C,D). Furthermore, rate and synchrony MTFs differed dramatically from each other (Figs. 3, 4A,B). Accordingly, an invariant representation for stimulus periodicity on the basis of rate or synchrony alone is not consistent with these tuning properties. This behavior is accounted in our data by the fact the neurons exhibit distinct response properties to sound envelope shape and repetition. At low stimulus modulation rates sustained firing patterns contribute by producing a high-fidelity representation of the sound envelope. In contrast, onset activity extends the encoding abilities to high modulation rates at which high temporal precision is a prerequisite. These data thus suggest that the neuronal code for modulated sounds in the CNIC is more akin to temporal pattern analysis in which sustained and onset activity separately contribute to the encoding of envelope shape and repetition.

Implications for envelope coding

Our comparison of rate and synchrony MTFs suggests that a simple coding scheme based on modulation tuning alone is unlikely because tuning properties varied considerably and were generally uncorrelated across stimulus conditions (Fig. 4E). This suggests that the repetition frequency parameter is not the most fundamental stimulus attribute and that rate and synchrony provide only partial envelope information. Instead, our data support the conclusion that the precise temporal discharge pattern is significantly more important. Thus, we propose that neurons in the CNIC use their discharge pattern to provide a high-fidelity representation of the envelope waveform.

In general, the observed temporal discharge patterns could be quite varied from unit-to-unit or stimulus condition; however, several patterns appear to hold. First, pure onset neurons (7 of 11) responded strictly to PNB and produced no significant response for SAMN. In comparison, most of the neurons that exhibited a sustained or mixed SRN response responded to both PNB and SAMN (43 of 54). Several tuning properties were also consistent across stimuli. The vast majority of MTFs obtained for PNB with our SC procedure exhibited a lowpass response pattern (Fig. 9G, purple curve). In comparison, the MTF patterns ob-

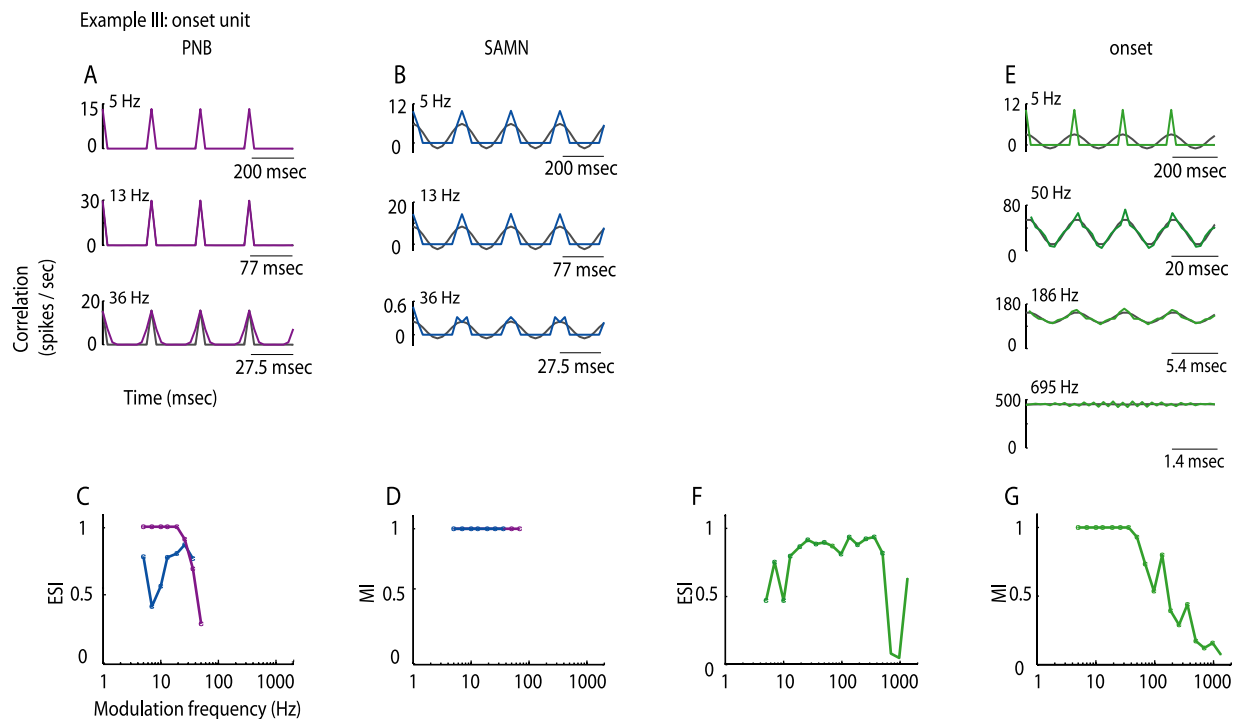


Figure 8. Shuffled correlation analysis for an onset single unit (same unit as in Fig. 2C). **A, B**, Shuffled correlation function for PNB (purple), SAMN (blue), and their stimulus envelope autocorrelation (gray) at 5, 13, and 36 Hz. **C, D**, esIMTF and miMTF for PNB (purple curve) and SAMN (blue curve). **E**, Periodic shuffled correlation function for onset (green) activity at 5, 50, 186, and 695 Hz to SRN. The envelope autocorrelation of the stimulus is superimposed for reference (gray). This unit did produce a sustained discharge, and therefore sustained activity analysis was not performed. **F, G**, esIMTF and miMTF for onset responses to SRN (green curve). The dots on the curves of **C, D, F**, and **G** indicate significant results (t test, $p < 0.05$) relative to a Poisson neuron with identical firing rates. Conditions with zero spikes were excluded from the analysis and not plotted.

served for SAMN were much more varied. One possible explanation for this disparity between PNB and SAMN is that the transient onsets in the PNB (250 μ s duration) evoked highly succinct neuronal responses, such that synchronized discharges were strictly limited by the stimulus repetition rate as a result of the time constant of the neuron. In comparison, the SAMN responses would also be influenced by the SAMN envelope shape because the rise and decay times vary as function of modulation frequency.

Direct comparison between the discharge pattern for SRN, SAMN, and PNB suggests that differences between PNB and SAMN tuning properties could arise if the PNB and SAMN sounds differentially activate onset or sustained responses. Hypothetically, it is conceivable that the brief noise burst in PNB is more likely to evoke a phasic onset response pattern, whereas the sinusoidal envelope of the SAMN could evoke sustained responses. Indeed, the observed discharge patterns for SRN are consistent with this hypothesis. The predicted modulation tuning properties obtained from the onset activity to SRN for our neuronal population closely resembled the observed lowpass trend observed for PNB [PNB (Fig. 9G, purple curve); onset SRN (Fig. 9G, green curve)]. In both instances, the output modulation index is ~ 1 for frequencies up to ~ 100 Hz and drops for higher frequencies. This is true although the SRN and PNB envelopes bear no relationship. The close association between these two trends indicates that the PNB and onset SRN discharge patterns may arise as a result of a common neuronal mechanism. Heil and colleagues (Heil and Irvine, 1997; Heil, 2001; Heil and Neubauer, 2003) have demonstrated that the timing of the first spike to transient envelopes are characterized by short latency responses and are well accounted for by a simple lowpass integrate-and-fire mechanism that could account for the observed lowpass trend. In

contrast, the population MTF for SAMN shares similar tuning properties to the predicted MTF obtained with the sustained discharge to SRN [SAMN (Fig. 9G, blue curve); sustained SRN (Fig. 9G, red curve)]. In both instances, the population MTF is tuned with maximum response modulation index of ~ 0.5 , although the population tuning peaks at a slightly higher modulation frequency for SRN (97 vs 26 Hz). This subtle disparity may be accounted for by the fact that for SAMN neuronal discharges in the CNIC adapt to repeated presentations of the sound envelope, whereas for SRN they do not. Together, these comparisons indicate that sustained and onset discharge patterns contribute differently during periodic stimulation and that the shape of the sound envelope is a key feature that regulates the temporal discharge pattern.

The proposed SC technique is mathematically similar to the shuffled autocorrelogram analysis recently used in the auditory nerve, cochlear nucleus, and inferior colliculus to study the representation of stimulus fine-structure and envelope information (Joris, 2003; Louage et al., 2004, 2005). We have extended this general approach to study the coding of envelope periodicity and shape for periodic signals. Because unfrozen noise carriers were used in this study, we were unable to characterize how fine-structure information contributes. However, our results demonstrate that the SC contains significant information that is related to the stimulus envelope for periodic signals. The close match between the stimulus and response correlograms observed for the PNB and SAMN and the various modulation frequencies tested (Figs. 5–10) provides evidence that the spiking pattern of CNIC neurons can provide an accurate representation of the envelope waveform especially for modulation frequencies below ~ 100 Hz. The conclusion that shape information is preserved only at low modulation frequencies and degrades beyond 100 Hz is sup-

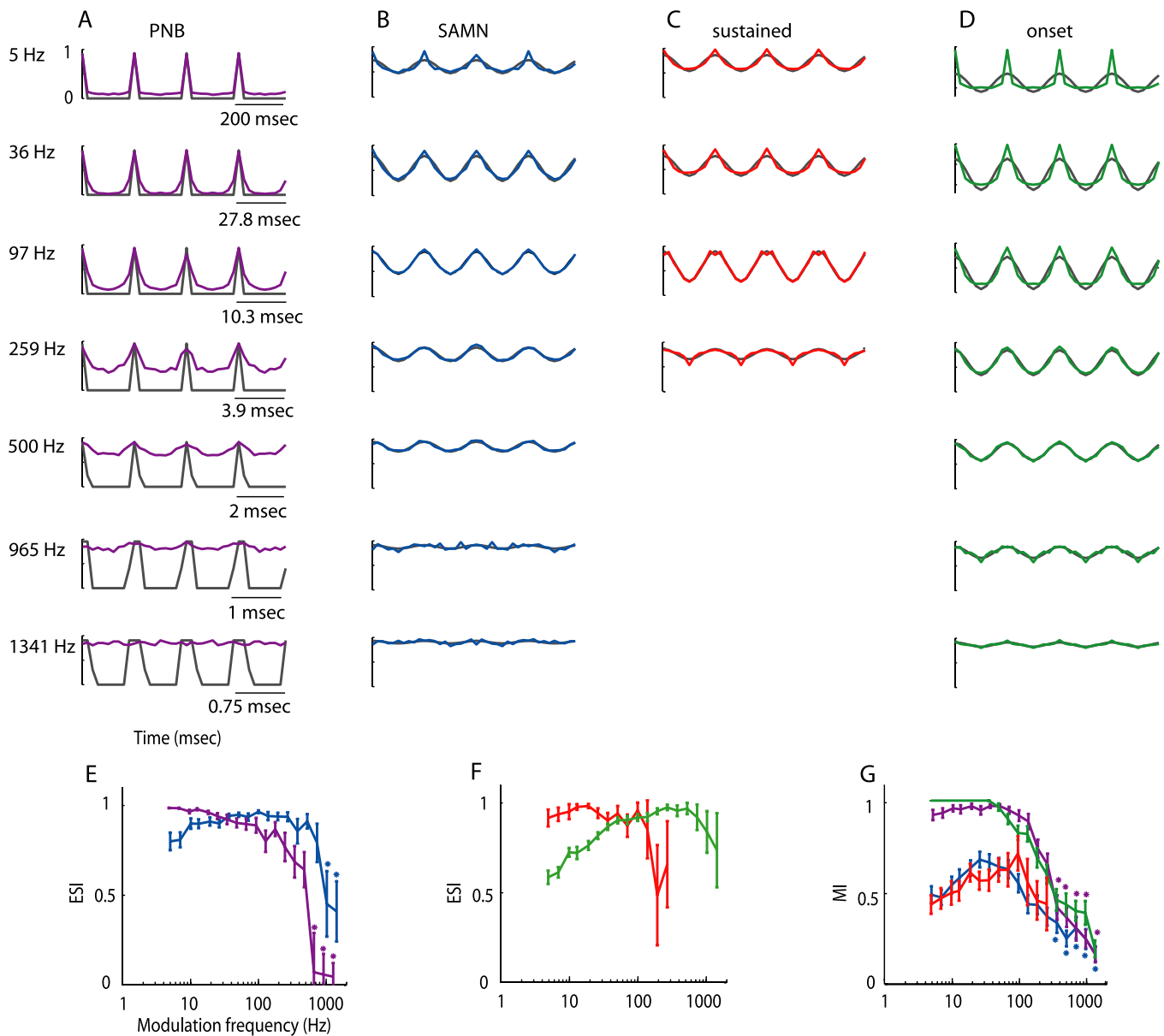


Figure 9. Comparison of the single-unit population activity pattern for PNB, SAMN, and SRN. Population averaged SC ($n = 65$ single neurons) for PNB (**A**, purple) and SAMN (**B**, blue). The PSC from sustained (**C**, red) and onset (**D**, green) SRN components. Sustained responses were only present for modulation frequencies up to 259 Hz of the SRN (**C**), because secondary spikes needed to measure sustained responses were not present in the neuronal discharges at these high modulation frequencies. The gray curve indicates the stimulus correlation. Note that the maximum value of each individual SC or PSC is normalized to 1 before averaging to avoid the effect of spike rate. **E**, esiMTF for PNB (purple) and SAMN (blue) from the population average SC. **F**, esiMTF for sustained (red) and onset (green) from the population average PSC. **G**, miMTF for PNB (purple), SAMN (blue), sustained (red), and onset (green) responses. Note that the miMTF pattern is closely matched between PNB versus onset activity and the SAMN versus sustained activity. The asterisks indicate nonsignificant ESI or MI values. Error bars indicate SEM.

ported directly by our quantitative discrimination analysis between PNB and SAMN (Fig. 10) (for additional discussion, see below, Relationship to psychophysics).

The proposed neuronal coding strategy for stimulus periodicity and envelope shape is likely initiated in brainstem circuits in which various cell classes have been identified on the basis of their temporal firing pattern (Young et al., 1988; Rhode, 1995; Stabler et al., 1996; Batra, 2006). Similar onset and sustained temporal firing patterns have been observed previously in the CNIC (Langner and Schreiner, 1988; Rees et al., 1997; Nelson and Carney, 2007). Although projecting brainstem nuclei and cell classes likely contribute to the observed response patterns (Langner and Schreiner, 1988; Rees et al., 1997; Ramachandran et al., 1999), it is also likely that intrinsic membrane and synaptic response char-

acteristics in the CNIC potentially enhance and contribute as well (Burger and Pollak, 2001; Sivaramakrishnan and Oliver, 2001; Pollak et al., 2003; Tan and Borst, 2007).

Onset and sustained discharge patterns have previously been observed in the brainstem, midbrain, and cortex of various awake species (Pollak et al., 1978; Langner et al., 2002; Batra, 2006; Ter-Mikaelian et al., 2007). Given the prevalence of these two discharge patterns, it is likely that onset and sustained activity could also serve to encode periodicity and envelope shape in the awake state. Although anesthesia is known to alter neural discharge patterns, this appears to be substantially less pronounced in the inferior colliculus compared with the auditory cortex (Ter-Mikaelian et al., 2007). Thus, it is likely that onset and sustained activity serve as a mechanism that is conserved

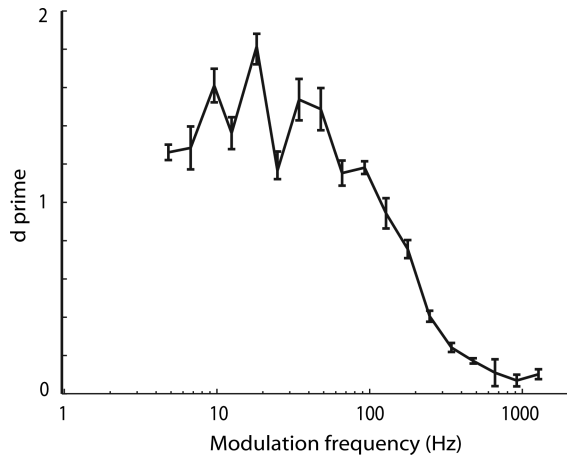


Figure 10. Shape discrimination analysis. Discrimination MTF was obtained from single units ($n = 65$) by computing the d' metric to compare the temporal activity pattern between PNB and SAMN for the CNIC population (see Materials and Methods). The d' MTF follows a lowpass trend indicating that discrimination between PNB and SAMN on the basis of the population spiking pattern is optimal for modulation frequencies below ~ 100 Hz. Error bars indicate SEM.

across structures and species for encoding temporal sound information. Our general proposition that the temporal discharge pattern is significantly more important than the response firing rate is in fact consistent with a recent study by Malone et al. (2007) in the auditory cortex of awake primates. This study demonstrated that temporal discharge pattern of cortical neurons is essential for the encoding SAM tones. In this study, most of the information regarding the stimulus identity was evident in the cycle histogram and not in the rate or synchrony tuning profile as traditionally assumed. In particular, they observed that systematically changing fundamental stimulus parameters (e.g., intensity or modulation index) significantly alters the MTF similar to the lack of invariance we observe with respect to envelope shape (Figs. 3, 4).

Relationship to psychophysics

The observed patterns of CNIC activity are consistent with various aspects of pitch and envelope shape perception. The response modulation index for the CNIC population was statistically significant for the SAMN and PNB up to ~ 500 Hz. This result mirrors human psychophysical data, which has identified the upper limit for periodicity pitch of modulated noise to lie between 500 and 800 Hz (Pollack, 1969; Burns and Viemeister, 1981). Unlike harmonic tone complexes, the type of modulated noise used in these psychophysical studies and in the present study contains no spectral information from which to extract the sound pitch. This implies that a temporal activity pattern analysis would be required to extract such information from our sounds. Although our results do not suggest that pitch is strictly represented as a temporal code, they imply that temporal information does indeed play a role.

Differences in the temporal pattern of activity for PNB and SAMN could also potentially account for the perception of the waveform shape, which contributes directly to the perceived timbre of complex sounds (Iverson and Krumhansl, 1993; Irino and Patterson, 1996). At low modulation rates, the pattern of neuronal activity for PNB is characterized by brief response epochs that mirror the sound envelope, whereas for SAMN the population response resembles a nearly perfect sinusoidal pattern (Fig. 9A,B). These differences were evident in the d' MTF, which dem-

onstrates that CNIC population activity is capable of encoding the shape information reliably up to ~ 100 Hz (Fig. 10). Interestingly, the ability for humans to discriminate periodic ramped and damped noise and tones, which are physically distinguished on the basis of their temporal shape, begins to deteriorate for frequencies beyond 100 Hz (Akeroyd and Patterson, 1995; Akeroyd and Patterson, 1997).

Summary

These data shed light on the neuronal code for temporal repetition and envelope shape information in the auditory midbrain. Collectively, the findings suggest that simple rate or synchrony codes oversimplify the manner in which CNIC neurons encoded temporal information and that the precise temporal pattern of the neuronal discharge is essential. How this information is used at subsequent levels of processing and how it contributes to basic auditory percepts of pitch, rhythms, and timbre still needs to be determined. Future studies will therefore need to consider how patterned temporal activity in the CNIC is altered in the transition to high-level cortical areas.

Appendix

Here, we derive the proof for the fast implementation of the shuffled correlation function. This implementation significantly reduces the computational load because it requires a total of $N + 1$ correlation compared with $N(N - 1)$ for the direct formula (Eq. 2). We start off with Equation 3:

$$\Phi_{\text{shuffled}}(\tau) = \sqrt{\frac{1}{N(N-1)} \left(N^2 \cdot \Phi_{\text{PSTH}}(\tau) - \sum_{k=1}^N \Phi_{kk}(\tau) \right)}, \quad (\text{A1})$$

and will use this to derive Equation 2. The correlation function for the response PSTH in Equation 4 can be expressed as follows:

$$\begin{aligned} \Phi_{\text{PSTH}}(\tau) &= \frac{1}{T} \int_0^T \text{PSTH}(t) \cdot \text{PSTH}(t + \tau) dt \\ &= \frac{1}{T} \int_0^T \frac{1}{N} \sum_{k=1}^N s_k(t) \cdot \frac{1}{N} \sum_{l=1}^N s_l(t + \tau) dt \\ &= \frac{1}{N^2} \sum_{k=1}^N \sum_{l=1}^N \frac{1}{T} \int_0^T s_k(t) \cdot s_l(t + \tau) dt = \frac{1}{N^2} \sum_{k=1}^N \sum_{l=1}^N \Phi_{kl}(\tau) \\ &= \frac{1}{N^2} \left[\sum_{k=1}^N \sum_{k \neq l} \Phi_{kl}(\tau) + \sum_{k=1}^N \Phi_{kk}(\tau) \right]. \quad (\text{A2}) \end{aligned}$$

Note that the autocorrelation of the PSTH is simply the average correlation between all possible trial pairs. Combining the above result with the inner argument of Equation A1 gives the following:

$$\begin{aligned} N^2 \cdot \Phi_{\text{PSTH}}(\tau) - \sum_{k=1}^N \Phi_{kk}(\tau) &= N^2 \cdot \frac{1}{N^2} \left[\sum_{k=1}^N \sum_{l \neq k} \Phi_{kl}(\tau) + \sum_{k=1}^N \Phi_{kk}(\tau) \right] \\ &\quad - \sum_{k=1}^N \Phi_{kk}(\tau) = \sum_{k=1}^N \sum_{l \neq k} \Phi_{kl}(\tau), \quad (\text{A3}) \end{aligned}$$

which corresponds to the sum of all the correlation functions between the off-diagonal terms. Substituting this result into Equation A1 gives the desired result as follows:

$$\Phi_{\text{shuffled}}(\tau) = \sqrt{\frac{1}{N(N-1)} \sum_{k=1}^N \sum_{l \neq k} \Phi_{kl}(\tau)}. \quad (\text{A4})$$

References

- Akeroyd MA, Patterson RD (1995) Discrimination of wideband noise modulated by a temporally asymmetric function. *J Acoust Soc Am* 98:2466–2474.
- Akeroyd MA, Patterson RD (1997) A comparison of detection and discrimination of temporal asymmetry in amplitude modulation. *J Acoust Soc Am* 101:430–439.
- Batra R (2006) Responses of neurons in the ventral nucleus of the lateral lemniscus to sinusoidally amplitude modulated tones. *J Neurophysiol* 96:2388–2398.
- Bilsen FA (1977) Pitch of noise signals: evidence for a “central spectrum.” *J Acoust Soc Am* 61:150–161.
- Bregman AS (1990) Auditory scene analysis: the perceptual organization of sound. Cambridge, MA: MIT.
- Burger RM, Pollak GD (2001) Reversible inactivation of the dorsal nucleus of the lateral lemniscus reveals its role in the processing of multiple sound sources in the inferior colliculus of bats. *J Neurosci* 21:4830–4843.
- Burns EM, Viemeister NF (1981) Played-again SAM: further observations on the pitch of amplitude-modulated noise. *J Acoust Soc Am* 70:1955–1960.
- Buunen TJ, Rhode WS (1978) Responses of fibers in the cat’s auditory nerve to the cubic difference tone. *J Acoust Soc Am* 64:772–781.
- Cariani PA, Delgutte B (1996) Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch. *J Neurophysiol* 76:1717–1734.
- Cedolin L, Delgutte B (2005) Pitch of complex tones: rate-place and interspike interval representations in the auditory nerve. *J Neurophysiol* 94:347–362.
- Chung DY, Colavita FB (1976) Periodicity pitch perception and its upper frequency limit in cats. *Perception and psychophysics* 20:433–437.
- Colavita FB (1977) Temporal pattern discrimination in the cat. *Physiol Behav* 18:513–521.
- Condon CJ, White KR, Feng AS (1996) Neurons with different temporal firing patterns in the inferior colliculus of the little brown bat differentially process sinusoidal amplitude-modulated signals. *J Comp Neurol A* 178:147–157.
- Cynx J, Shapiro M (1986) Perception of missing fundamental by a species of songbird (*Sturnus vulgaris*). *J Comp Psychol* 100:356–360.
- Dreyer A, Delgutte B (2006) Phase locking of auditory-nerve fibers to the envelopes of high-frequency sounds: implications for sound localization. *J Neurophysiol* 96:2327–2341.
- Efron B (1981) Nonparametric estimates of standard error: the jackknife, the bootstrap and other methods. *Biometrika* 68:589–599.
- Goldberg JM, Brown PB (1969) Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J Neurophysiol* 32:613–636.
- Goldstein JL (1973) An optimum processor theory for the central formation of the pitch of complex tones. *J Acoust Soc Am* 54:1496–1516.
- Green DN, Swets JA (1966) Signal detection theory and psychophysics. New York: Wiley.
- Heffner H, Whitfield IC (1976) Perception of the missing fundamental by cats. *J Acoust Soc Am* 59:915–919.
- Heil P (2001) Representation of sound onsets in the auditory system. *Audiol Neurootol* 6:167–172.
- Heil P, Irvine DR (1997) First-spike timing of auditory-nerve fibers and comparison with auditory cortex. *J Neurophysiol* 78:2438–2454.
- Heil P, Neubauer H (2003) A unifying basis of auditory thresholds based on temporal summation. *Proc Natl Acad Sci U S A* 100:6151–6156.
- Helmholtz HLF (1885) On the sensation of tone. New York: Dover.
- Irino T, Patterson RD (1996) Temporal asymmetry in the auditory system. *J Acoust Soc Am* 99:2316–2331.
- Iverson P, Krumhansl CL (1993) Isolating the dynamic attributes of musical timbre. *J Acoust Soc Am* 94:2595–2603.
- Joris PX (2003) Interaural time sensitivity dominated by cochlea-induced envelope patterns. *J Neurosci* 23:6345–6350.
- Joris PX, Yin TCT (1992) Responses to amplitude-modulated tones in the auditory nerve of the cat. *J Acoust Soc Am* 91:215–232.
- Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. *Physiol Rev* 84:541–577.
- Kim DO, Sirianni JG, Chang SO (1990) Responses of DCN-PVCN neurons and auditory nerve fibers in unanesthetized decerebrate cats to AM and pure tones: analysis with autocorrelation/power-spectrum. *Hear Res* 45:95–113.
- Krishna BS, Semple MN (2000) Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. *J Neurophysiol* 84:255–273.
- Langner G (1992) Periodicity coding in the auditory system. *Hear Res* 60:115–142.
- Langner G, Schreiner CE (1988) Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *J Neurophysiol* 60:1799–1822.
- Langner G, Albert M, Briede T (2002) Temporal and spatial coding of periodicity information in the inferior colliculus of awake chinchilla (*Chinchilla laniger*). *Hear Res* 168:110–130.
- Licklider JC (1951) A duplex theory of pitch perception. *Experientia* 7:128–134.
- Louage DH, van der Heijden M, Joris PX (2004) Temporal properties of responses to broadband noise in the auditory nerve. *J Neurophysiol* 91:2051–2065.
- Louage DH, van der Heijden M, Joris PX (2005) Enhanced temporal response properties of anteroventral cochlear nucleus neurons to broadband noise. *J Neurosci* 25:1560–1570.
- Malone BJ, Scott BH, Semple MN (2007) Dynamic amplitude coding in the auditory cortex of awake rhesus macaques. *J Neurophysiol* 98:1451–1474.
- Merzenich MM, Reid MD (1974) Representation of the cochlea within the inferior colliculus of the cat. *Brain Res* 77:397–415.
- Moore BCJ (1997) An introduction to the psychology of hearing. San Diego: Academic.
- Nelson PC, Carney LH (2007) Neural rate and timing cues for detection and discrimination of amplitude-modulated tones in the awake rabbit inferior colliculus. *J Neurophysiol* 97:522–539.
- Ohm GS (1844) Noch ein Paar Worte uber die Definition des Tones. *Ann Phys Chem* 62:1–18.
- Oliver DL, Morest DK (1984) The central nucleus of the inferior colliculus in the cat. *J Comp Neurol* 222:237–264.
- Oxenham AJ, Bernstein JG, Penagos H (2004) Correct tonotopic representation is necessary for complex pitch perception. *Proc Natl Acad Sci U S A* 101:1421–1425.
- Paquette C, Peretz I (1997) Role of familiarity in auditory discrimination of musical instrument: a laterality study. *Cortex* 33:689–696.
- Pollack I (1969) Periodicity pitch for interrupted white noise—fact or artifact? *J Acoust Soc Am* 45:237–238.
- Pollak GD, Burger RM, Klug A (2003) Dissecting the circuitry of the auditory system. *Trends Neurosci* 26:33–39.
- Pollak GK, Marsh DS, Bodenhamer R, Souther A (1978) A single-unit analysis of inferior colliculus in unanesthetized bats: response patterns and spike-count functions generated by constant-frequency and frequency-modulated sounds. *J Neurophysiol* 41:677–691.
- Ramachandran R, Davis KA, May BJ (1999) Single-unit responses in the inferior colliculus of decerebrate cats. I. Classification based on frequency response maps. *J Neurophysiol* 82:152–163.
- Rees A, Møller AR (1987) Stimulus properties influencing the responses of inferior colliculus neurons to amplitude-modulated sounds. *Hear Res* 27:129–143.
- Rees A, Palmer AR (1989) Neuronal responses to amplitude-modulated and pure-tone stimuli in the guinea pig inferior colliculus, and their modification by broadband noise. *J Acoust Soc Am* 85:1978–1994.
- Rees A, Sarbaz A, Malmierca MS, Le Beau FE (1997) Regularity of firing of neurons in the inferior colliculus. *J Neurophysiol* 77:2945–2965.
- Rhode WS (1995) Interspike intervals as a correlate of periodicity pitch in cat cochlear nucleus. *J Acoust Soc Am* 97:2414–2429.
- Rhode WS, Greenberg S (1994) Encoding of amplitude modulation in the cochlear nucleus of the cat. *J Neurophysiol* 71:1797–1825.
- Risset JC, Wessel DL (1982) Exploration of timbre by analysis and synthesis. In: *The psychology of music*, pp 26–58. New York: Academic.
- Schouten JF (1940) The residue and the mechanisms of hearing. *Proc K Ned Akad Wet* 43:991–999.

- Seebeck A (1841) Beobachtungen über einige Bedingungen der Entstehung von Tönen. *Ann Phys Chem* 53:417–436.
- Shaddock Palombi P, Backoff PM, Caspary DM (2001) Responses of young and aged rat inferior colliculus neurons to sinusoidally amplitude modulated stimuli. *Hear Res* 153:174–180.
- Shamma S, Klein D (2000) The case of the missing pitch templates: how harmonic templates emerge in the early auditory system. *J Acoust Soc Am* 107:2631–2644.
- Sivaramakrishnan S, Oliver DL (2001) Distinct K currents result in physiologically distinct cell types in the inferior colliculus of the rat. *J Neurosci* 21:2861–2877.
- Stabler SE, Palmer AR, Winter IM (1996) Temporal and mean rate discharge patterns of single units in the dorsal cochlear nucleus of the anesthetized guinea pig. *J Neurophysiol* 76:1667–1688.
- Tan ML, Borst JG (2007) Comparison of responses of neurons in the mouse inferior colliculus to current injections, tones of different durations and sinusoidal amplitude-modulated tones. *J Neurophysiol* 98:454–466.
- Terhardt E (1974) Pitch, consonance, and harmony. *J Acoust Soc Am* 55:1061–1069.
- Ter-Mikaelian M, Sanes DH, Semple MN (2007) Transformation of temporal properties between auditory midbrain and cortex in the awake Mongolian gerbil. *J Neurosci* 27:6091–6102.
- Tomlinson RW, Schwarz DW (1988) Perception of the missing fundamental in nonhuman primates. *J Acoust Soc Am* 84:560–565.
- Wightman FL (1973) The pattern-transformation model of pitch. *J Acoust Soc Am* 54:407–416.
- Young ED, Robert JM, Shofner WP (1988) Regularity and latency of units in ventral cochlear nucleus: implications for unit classification and generation of response properties. *J Neurophysiol* 60:1–29.