Midbrain-Level Neural Correlates of Behavioral Tone-in-Noise Detection: Dependence on Energy and Envelope Cues

Yingxuan Wang,¹ Kristina S. Abrams,² Laurel H. Carney,^{1,2} and Kenneth S. Henry^{1,2,3}

¹Departments of Biomedical Engineering, ²Neuroscience, and ³Otolaryngology, University of Rochester, Rochester, New York 14642

Hearing in noise is a problem often assumed to depend on encoding of energy level by channels tuned to target frequencies, but few studies have tested this hypothesis. The present study examined neural correlates of behavioral tone-in-noise (TIN) detection in budgerigars (*Melopsittacus undulatus*, either sex), a parakeet species with human-like behavioral sensitivity to many simple and complex sounds. Behavioral sensitivity to tones in band-limited noise was assessed using operant-conditioning procedures. Neural recordings were made in awake animals from midbrain-level neurons in the inferior colliculus, the first processing stage of the ascending auditory pathway with pronounced rate-based encoding of stimulus amplitude modulation. Budgerigar TIN detection thresholds were similar to human thresholds across the full range of frequencies (0.5–4 kHz) and noise levels (45–85 dB SPL) tested. Also as in humans, thresholds were minimally affected by a challenging roving-level condition with random variation in background-noise level. Many midbrain neurons showed a decreasing response rate as TIN signal-to-noise ratio (SNR) was increased by elevating the tone level, a pattern attributable to amplitude-modulation tuning in these cells and the fact that higher SNR tone-plus-noise stimuli have flatter amplitude envelopes. TIN thresholds of individual neurons were as sensitive as behavioral thresholds under most conditions, perhaps surprisingly even when the unit's characteristic frequency was tuned an octave or more away from the test frequency. A model that combined responses of two cell types enhanced TIN sensitivity in the roving-level condition. These results highlight the importance of midbrain-level envelope encoding and off-frequency neural channels for hearing in noise.

Key words: budgerigar; envelope; inferior colliculus; operant conditioning; roving level; tone in noise

Significance Statement

Detection of target sounds in noise is often assumed to depend on energy-level encoding by neural processing channels tuned to the target frequency. In contrast, we found that tone-in-noise sensitivity in budgerigars was often greatest in midbrain neurons not tuned to the test frequency, underscoring the potential importance of off-frequency channels for perception. Furthermore, the results highlight the importance of envelope processing for hearing in noise, especially under challenging conditions with random variation in background noise level over time.

Introduction

Hearing in noise is a common challenge faced in everyday life that is often thought to depend on energy-level encoding by neural processing channels tuned to target frequencies (Fletcher, 1940; Patterson, 1976). For the simplified case of tone-in-noise (TIN) detection, listeners are more likely to detect a tone when the stimulus energy level is

greater and when the TIN time-varying amplitude envelope is flatter, suggesting that both energy and envelope cues contribute to detection (Kohlrausch et al., 1997; Mao et al., 2013). Minimal threshold shifts when energy is made unreliable through equalization of the stimulus level across test trials (Richards, 1992) or use of a roving-level paradigm with random level variation (Kidd et al., 1989) further implicate envelope fluctuations as a potentially important cue for hearing in noise.

The inferior colliculus (IC) of the midbrain is a key brain region for understanding neural encoding of signals in noise because the IC is the first nucleus of the ascending pathway that encodes envelope structure through substantial changes in average response rate, known as amplitude-modulation tuning (Joris et al., 2004). Many IC neurons can be characterized both by a characteristic frequency (CF), indicating the tone frequency of maximal sensitivity, and by a best modulation frequency (BMF) in response to periodic envelope fluctuations (Langner and

Received Dec. 10, 2020; revised June 30, 2021; accepted July 8, 2021.

Author contributions: Y.W., L.H.C., and K.S.H. designed research; Y.W., K.S.A., and K.S.H. performed research; Y.W. and K.S.H. analyzed data; and Y.W., K.S.H., and L.H.C. wrote the paper.

This work was supported by Grants R01-DC017519 and R01-DC001641 from the National Institute on Deafness and Communication Disorders. Cassidy Amburgey, Caleb Connelly, Lucinda Hinojosa, Brett Tingley, and Stephanie Wong assisted with behavioral experiments. Douglas Schwarz provided software and technical support.

The authors declare no competing financial interests.

Correspondence should be addressed to Kenneth S. Henry at kenneth_henry@urmc.rochester.edu.

<https://doi.org/10.1523/JNEUROSCI.3103-20.2021>

Copyright © 2021 the authors

Figure 1. TIN stimuli. **A, B**, Example spectra of band-limited noise, with and without the addition of a tone. Stimulus SNR is indicated at the top of each panel. Noise is 0.33 octave bandwidth, log centered on the 2 kHz tone frequency, 65 dB SPL overall level. **C**, Overall energy level of the TIN stimulus increases with increasing SNR from -12 – 9 dB. White horizontal lines in each symbol show the median, filled boxes show the IQR, and vertical lines extend ± 2.7 SDs from the mean. **D, E**, Waveforms of the stimuli from **A** and **B**. Note that tone and noise waveforms were simultaneously gated for a duration of 300 ms; the central 30 ms of each stimulus is shown to better illustrate change in the quality of stimulus envelope fluctuations on addition of the tone. Thick black lines indicate the stimulus envelope. **F**, Normalized envelope slope (see below, Modeling TIN responses in individual units) decreases with increasing SNR from -12 to 9 dB; symbols as in **C**.

Schreiner, 1988; Krishna and Semple, 2000). The BMF is the amplitude-modulation frequency, determined from a modulation transfer function (MTF), that evokes the greatest response rate. Across many species, IC neurons commonly show band-enhanced modulation tuning (Kim et al., 2015, 2020) with BMFs up to several hundred Hz (Rees and Palmer, 1989; Müller-Preuss et al., 1994; Keller and Takahashi, 2000; Krishna and Semple, 2000; Woolley and Casseday, 2005; Nelson and Carney, 2007; Baumann et al., 2011; Kim et al., 2020).

The extent to which IC amplitude-modulation tuning contributes to TIN encoding is unknown. Previous studies have identified neurons with either increasing or decreasing response rates as the signal-to-noise ratio (SNR) is increased by elevating the tone level (Jiang et al., 1997; Ramachandran et al., 2000; Rocchi and Ramachandran, 2018). Although in some cases related to energy-dependent excitation or inhibition at the tone frequency (e.g., because of strong inhibition by high-energy CF tones in type-O neurons; Ramachandran et al., 1999, 2000), decreasing rate-SNR functions could also result from amplitude-modulation tuning (Mao and Carney, 2015; Fan et al., 2018). For example, IC neurons with band-enhanced modulation tuning could show decreasing response rates with increasing TIN SNR because of flattening of the stimulus envelope by the addition of higher-level tones (Fig. 1). Because previous physiological studies focused largely on single mechanisms, typically energy-based encoding, the extent to which modulation tuning contributes to TIN sensitivity remains unclear.

The present study quantified IC neural correlates of behavioral TIN sensitivity in the budgerigar (*Melopsittacus undulatus*), a small parrot species with behavioral performance similar to humans on tasks including frequency discrimination of tones and vowel formants (Dent et al., 2000; Henry et al., 2017b), amplitude-modulation detection (Dooling and Searcy, 1981; Carney et al., 2013; Henry et al., 2016), and TIN detection (Dooling and Saunders, 1975;

Saunders and Pallone, 1980). Moreover, recent behavioral studies suggest that this species uses the same energy- and envelope-based cues for TIN detection as human listeners (Henry et al., 2020; Henry and Abrams, 2021). Neurons in the budgerigar IC, also known as nucleus mesencephalicus lateralis pars dorsalis, show band-enhanced modulation tuning and other response properties similar to those found in the mammalian IC (Henry et al., 2017a).

Behavioral sensitivity to 0.5–4 kHz tones was measured in 0.33 octave band-limited noise using operant-conditioning procedures. Noise was log centered in frequency on the tone, ranged from 45–85 dB SPL, and was either fixed in level or randomly varied across test trials (together with the tone level) to assess the impact of a challenging roving-level condition on TIN sensitivity. Neural responses were recorded from the IC in awake animals using identical stimuli to differentiate energy- from envelope-based TIN encoding strategies.

Materials and Methods

Animals. Behavioral and neurophysiological studies were conducted in adult budgerigars under a protocol approved by the University Committee on Animal Resources at the University of Rochester. Animals ranged in age from 2 to 5 years and were of either sex. Behavioral experiments were conducted in four animals (two male) trained using operant-conditioning procedures. A subset of the behavioral results (for low and high noise levels) were included as part of the control group in a different study on the effects of auditory-nerve damage on behavioral TIN detection (Henry and Abrams, 2021). Neurophysiological recordings were made from the IC in four animals (two male) without anesthesia using chronically implanted microelectrodes. Different animals were used for behavioral and neurophysiological experiments.

Behavioral experiments. Behavioral experiments were conducted in trained budgerigars using previously reported procedures and equipment (Henry et al., 2017b; Henry and Abrams, 2021). Briefly, testing was performed in four single-walled acoustic isolation chambers (0.3 m³) lined with sound-absorbing foam. Animals perched under a loudspeaker (MC60, Polk Audio) inside the chamber facing three response switches. Experiments were controlled by a PC running a custom MATLAB program (MathWorks) and linked to a data acquisition card (PCI 6151 or PCIe 6251, NI), microcontroller (Arduino Leonardo), and custom hardware. Computer-generated stimuli (50 kHz sampling frequency) were digitally filtered to correct for the frequency response of the system and converted to analog by the data acquisition card before power amplification (D-75A, Crown Audio) and presentation by the loudspeaker. The calibration filter was determined based on the output of an 0.5 inch precision microphone (type 4134, Brüel & Kjær) in response to 249 log-spaced tone frequencies from 0.05 to 15.1 kHz.

Tone sensitivity in band-limited noise was evaluated at 4 octave spaced test frequencies from 0.5 to 4 kHz and three noise levels (12 conditions total). Noise levels were 55, 65, and 85 dB SPL for 0.5 kHz and 45, 65, and 85 dB SPL for the higher test frequencies. The lowest noise level was 55 dB SPL for the 500 Hz test frequency rather than 45 dB SPL to ensure sufficient stimulus audibility (i.e., at least 20 dB above the

audiometric threshold; Wong et al., 2019). Stimulus conditions were completed in different order across animals and tested repeatedly until behavioral sensitivity was stable, as discussed below. Noise was 0.33 octave in bandwidth and log centered on the tone frequency in all cases. Tones were simultaneously gated on and off with noise waveforms. All stimuli were 0.3 s in duration with 10 ms cosine-squared ($\cos^2$) onset and offset ramps.

Animals began each trial by pecking the center switch, which initiated presentation of a single stimulus. The stimulus was either a standard noise-alone waveform or a target tone-plus-noise waveform. The correct response to the standard stimulus was the right switch, and the correct response to the target was the left switch. Correct and incorrect responses were reinforced by dispensing individual millet seeds and by timeouts during which the chamber light was turned off, respectively. The number of dispensed seeds was adaptively varied during testing based on the last 50 trials to control response bias. Bias was calculated as 0.5 times the sum of the Z score of the hit rate and the Z score of the false alarm rate (Macmillan and Creelman, 1991). Test sessions with absolute bias >0.3 , computed using all trials within the session, were excluded from further analysis. Behavioral testing was conducted 6–7 d per week in morning and afternoon blocks lasting ~ 30 min each.

Animals were first trained to discriminate between the standard noise-alone stimulus and a high SNR (10–15 dB) target stimulus. When animals reached 90% correct discrimination on this task, TIN sensitivity was assessed using two-down, one-up tracking procedures during which SNR was adaptively varied within single test sessions to determine detection thresholds (Levitt, 1971). The initial SNR of the target stimulus was 10–15 dB; target SNR increased following each incorrect response to a target stimulus and decreased following two consecutive correct responses to target stimuli with the same SNR. Trials at which the direction of the track (SNR across target trials) changed from increasing to decreasing, or vice versa, were identified as reversals. The step size of the track decreased from a starting value of 3 dB to 2 dB after two reversals and 1 dB after four reversals.

Within each test session, the level of the band-limited noise remained fixed for a minimum of 15 reversals until the following two stability criteria were met: (1) the SD of the SNR of the final eight reversal points was <3 dB, and (2) the mean SNR difference between the final four reversal points and the preceding four reversals was <3 dB. Thereafter, the track continued under a roving-level condition for which the overall level of the stimulus (including the tone for tone-plus-noise trials) was randomly scaled by ± 10 dB on each trial (uniform distribution with 1 dB resolution). The track continued for a minimum of 10 additional reversals until the same stability criteria defined above for the fixed-level condition were again satisfied. Animals typically completed 4–6 tracks per day consisting of 150–200 trials each. Reversal-based thresholds were calculated for fixed-level and roving-level portions of each track as the mean SNR of the final eight reversal points within each portion.

Tracking sessions were conducted repeatedly on the same condition until (1) at least 13 thresholds were obtained, (2) the SD of the final six track thresholds was <3 dB, and (3) the mean difference between the final three track thresholds and the preceding three was <3 dB. When all these criteria were met for both fixed- and roving-level thresholds, animals moved on to the next condition. Each animal completed the conditions in random sequence at least twice. Testing on a condition was discontinued when there was no significant threshold difference from the previous testing block for the same stimulus condition. The total duration of behavioral testing ranged from 11 to 21 weeks across animals. Final thresholds were calculated for each condition and in each animal as the mean reversal-based threshold of the last 10 stable tracks.

Electrode implantation procedure. Assemblies consisting of one to two metal microelectrodes (tungsten, iridium, or platinum-iridium; 3–5 M Ω ; Microprobes for Life Science) attached to a miniature microdrive (nano-Drive; Cambridge NeuroTech) were implanted into the IC of anesthetized animals using previously described methods (Henry et al., 2016, 2017a).

Briefly, anesthesia was induced with a bolus injection of ketamine (3–5 mg/kg) and dexmedetomidine (0.08–0.1 mg/kg, s.c.), and maintained throughout the ~ 2 h implantation procedure using continuous infusion of the same anesthetics (ketamine, 6–10 mg/kg/h s.c.; dexmedetomidine, 0.16–0.27 mg/kg/h, s.c.). Breathing rate was monitored and body temperature was maintained at 39–41°C using a warming pad (HTP-1500, Adroit Medical Systems).

Animals were placed in a head holder with the nares positioned ~ 5 mm above the interaural line. A region of the dorsal cranial surface was exposed using standard surgical procedures and a craniotomy made for insertion of the microelectrodes. The craniotomy was ~ 1 mm in diameter, 3.5 mm lateral from the midline of the skull, and positioned rostro-caudally so that the trajectory of the electrodes intersected with a point ~ 0.5 mm posterior to the interaural line. Noise bursts and tones were presented as the microelectrodes were lowered into the brain to guide initial placement of the recording tips in the central nucleus of the IC near its dorsal margin (~ 8.5 mm depth).

Following successful targeting of the IC, the craniotomy was sealed (Kwik-Sil, World Precision Instruments), and the microdrive assembly adhered to the skull surface using M0.6 anchor screws and dental cement. A lightweight plastic cap was secured over the assembly with an external miniature electrical connector on the posterior surface to interface with the electrophysiological recording system, described below. After the assemblies were mounted, the position of the electrode tip or tips was adjusted so that the initial location of the recording tips was in the dorsal, low-frequency region of the IC.

Neurophysiological recordings. Recordings were made beginning 1–2 d after the implantation surgery during daily 2 h recording sessions over several weeks, with the microelectrodes extended an additional 30 μ m each day using the control screw on the microdrive to sample neural responses along the tonotopic gradient of the IC. A total of 4–7 reimplantations were performed in each animal, and an average of 16 recording sessions were conducted. In total, the recording procedures yielded 207 multiunit sessions and seven sessions with good single-unit isolation.

Animals perched during recording sessions in a wire cage that was centered on a table in a sound-isolation booth. The table top and walls and ceiling of the booth were lined with sound absorbing foam. A free-field loudspeaker (MC60, Polk Audio) was mounted at one end of the table facing the animal at a distance of 45 cm. Birds were visually monitored with a closed-circuit video camera system to ensure that they remained perched and facing the loudspeaker throughout the recording session.

Stimulus waveforms were generated using a custom MATLAB program. Computer-generated stimuli (50 kHz sampling frequency) were converted to analog signals using a data acquisition card at full scale (± 10 V; PCIe-6251, National Instruments) and attenuated to the desired level using a programmable attenuator (PA5; Tucker-Davis Technologies). A power amplifier (D-75A, Crown Audio) drove the loudspeaker. Stimulus calibration was accomplished using a digital filter that compensated for the frequency response of the system. The filter was designed based on the output of a 0.25 inch precision microphone (type 4938, Brüel & Kjær) placed at the location of the animal's head in response to 249 log-spaced tone pips ranging in frequency from 0.05 to 15.1 kHz.

Electrophysiological recordings were referenced to one of the anchor screws of the microdrive using a multichannel recording system (RHD2132 amplifier chip and C3100 USB interface board, Intan Technologies). Recordings were hardware filtered (150 Hz, high pass) and sampled on the head-mounted amplifier chip, then saved to the hard drive of the computer with an additional trigger channel that indicated the onset time of each stimulus.

Recordings were made at a sampling frequency of 30 kHz. For subsequent analyses, raw recordings were resampled at 50 kHz and then bandpass filtered in MATLAB [500-point finite impulse response (FIR), 0.75–10 kHz] to minimize the local field potential. A third-order Teager Energy Operator (Choi et al., 2006) was then applied to the filtered signal for spike detection. Spikes were detected based on a visually determined threshold applied to the transformed response waveform, once per recording session. Recordings with consistent spike shapes throughout the

Figure 2. Representative single-unit neurophysiological recording from the budgerigar IC. **A**, Waveforms of the raw recording (black) and after transformation by a Teager energy operator (TEO; red; Choi et al., 2006). The blue dotted line indicates the threshold for spike detection from the Teager-transformed waveform. **B**, Mean waveform of 44 spikes detected over 2 s (black); individual waveforms are shown in gray. **C**, The inter-spike-interval (ISI) distribution of the spikes in **B**; intervals 4 ms are not shown.

session and <1% of interspike intervals <1 ms were defined as single-unit responses (Fig. 2); otherwise, responses were considered multiunit. The average response rate was calculated over the time interval beginning 50 ms after stimulus onset, to exclude the contribution of the onset response.

Frequency response maps and MTFs. The frequency response maps (RM) and MTFs were measured at the beginning of each recording session to characterize the recording site. An RM was measured in response to pure tones of varying frequency (0.25–8 kHz, 12 steps per octave) and level (15–75 dB SPL, 10 dB step size). A set of silent stimuli was included for spontaneous-rate measurement. Tones were 100 ms in duration with 10 ms $\cos^2$ onset and offset ramps. Stimuli were presented in random sequence (all frequency and level combinations) for three repetitions with a 350 ms silent period between successive stimuli.

Average response rates were normalized by subtracting the spontaneous rate, interpolated on a 100×100 frequency-by-level grid, and smoothed using a 3×3 moving-average Tukey window. The resulting response rate matrices were used to calculate pure-tone tuning curves plotting the threshold stimulus level necessary to evoke a criterion discharge rate as a function of stimulus frequency. The criterion response rate was typically set at 20% of the highest tone-evoked response rate across all stimuli. In rare cases where the maximum rate at 20 dB SPL exceeded 20% of the maximum response rate (i.e., in particularly sensitive units), the tuning curve criterion was redefined as the maximum rate at 20 dB SPL. The CF of each unit was defined as the frequency of the minimum of the pure-tone tuning curve.

The MTF was obtained in response to sinusoidal amplitude modulated tones with carrier frequency equal to the estimated CF, 100% modulation depth, and variable modulation frequency as in prior studies of IC modulation tuning (Langner and Schreiner, 1988; Krishna and Semple, 2000; Nelson and Carney, 2007; Henry et al., 2016, 2017a). Stimuli were 0.8 s in duration with 50 ms $\cos^2$ onset and offset ramps. Because of the mismatch of on-line estimation of unit CF and the subsequent off-line analysis, 159/207 multiunits and 7/7 single units had the MTF measured with the carrier frequency within 0.16 octave of CF. Modulation frequencies ranged from four to either 1024 Hz or 0.75 times the carrier frequency, whichever value was lower, with three steps per octave. MTF stimuli were presented at 65 dB SPL with a 350 ms silent period between successive stimuli. Modulation frequencies were presented in random sequence for four repetitions.

MTFs were smoothed by fitting a spline curve ($p = 0.99$) to the response rate as a function of modulation frequency. The BMF was determined as the geometric mean of frequencies crossing 0.99 of the maximum rate in the smoothed MTF. The percentage of enhancement (i.e., strength) of the MTF was quantified as the rate difference between BMF and the unmodulated stimulus condition, normalized by the sum of the two values.

Neural TIN detection thresholds. Responses to TIN stimuli were then obtained with tones matched to unit CF and to the four test frequencies used in the behavioral experiments (0.5, 1, 2, and 4 kHz). Responses to

different stimulus frequencies were recorded in separate testing blocks. Stimuli were generated by adding a 0.3 s tone to a 0.33 octave band-limited noise waveform of the same duration. Noise waveforms were generated independently for each stimulus presentation using a 5000-point FIR filter and were always log centered on the tone frequency. Noise level varied from 35 to 75 dB SPL in 10 dB steps, and SNR varied from –12 to 9 dB in 3 dB steps. A noise-alone stimulus ($-\infty$ SNR) was also included for each noise level. TIN stimuli were presented with 10 ms $\cos^2$ onset and offset ramps in random sequence for a total of 20 repetitions. The silent interval between stimuli was 250 ms.

Neural TIN detection thresholds were estimated by receiver-operating characteristic (ROC) analysis (Egan, 1975) of the functions plotting response rate versus stimulus SNR. For each SNR, the classification performance of the neuron was defined as the percentage of separation (i.e., area under the ROC curve) between the rate distribution observed for the noise-alone stimulus and the rate distribution observed for tones plus noise. The performance SNR function was interpolated, and the threshold was calculated as the lowest SNR above which classification performance consistently exceeded 70.7%.

Thresholds based on responses pooled across multiple recording sites were estimated using a population-pattern decoder (Jazayeri and Movshon, 2006; Day and Delgutte, 2013). Pooling was conducted both across all units and for units with CFs within ± 0.16 octave of the tone frequency. For each stimulus SNR θ , pairwise discrimination performance of the neural population was calculated between TIN stimuli and noise-alone stimuli by first drawing 1000 population responses for each of the two stimulus alternatives at random. For each population draw, the decoder calculates the log likelihood of the two alternatives as in Jazayeri and Movshon (2006) as follows:

$$\log L(\theta) = \sum_{i=1}^N n_i \log f_i(\theta) - \sum_{i=1}^N f_i(\theta) - \sum_{i=1}^N \log(n_i!),$$

where N is the total number of units, $f_i(\theta)$ is the average discharge rate for unit i as a function of θ , and n_i is an average rate for one trial that was randomly drawn from the responses of the unit i . The first term is an optimally weighted sum of response rates across the population, whereas the second term is the sum of response rates. The last term can be ignored because it is independent of θ . For each random population draw, the selected response rate n_i was excluded from the dataset in calculating $f_i(\theta)$ to avoid overfitting. Discrimination performance was calculated for each SNR as the proportion of population draws for which the log likelihood of the correct stimulus condition was greater than the alternative. The population neural threshold was defined as the minimum SNR above which the decoder model was consistently $\geq 70.7\%$ correct in discriminating noise-alone from TIN stimuli.

Modeling TIN responses in individual units. Multiple regression models were used to determine the extent to which the response rate in individual neurons could be predicted based on energy and envelope cues. Values of both cues were calculated on a stimulus-by-stimulus basis in a manner similar to previous studies of cues for TIN detection (Fletcher, 1940; Richards, 1992; Davidson et al., 2009; Mao et al., 2013; Henry et al., 2020). The energy cue was calculated as the root mean square amplitude of the stimulus waveform in dB SPL. The envelope cue was calculated by first computing the Hilbert envelope of the stimulus and normalizing this function to have a mean value of one. The envelope cue was then calculated as mean absolute value of the time-varying envelope slope. This is the same envelope-slope metric used in several earlier psychophysical studies of TIN detection (Davidson et al., 2006; Mao et al., 2013) but with the preceding critical-band filter omitted.

The first analysis modeled TIN response rate, Y , with only energy terms as follows:

$$Y = \beta_E(\max(x_E, thr_E) - thr_E) + Y_0,$$

where x_E is the energy cue, β_E is the energy coefficient relating response rate to x_E , thr_E is the energy threshold above which rate is energy dependent, and Y_0 is the spontaneous rate. The second multiple regression model included envelope terms in addition to the energy terms in the above model as follows:

$$Y = \beta_E(\max(x_E, thr_E) - thr_E) + Y_0 + \beta_{env}x_{env} + \beta_{E*env}x_{env}(\max(x_E, thr_E) - thr_E),$$

where β_{env} and β_{E*env} are the coefficients of the envelope term and the energy by envelope interaction, respectively. For both models, the values of the coefficients were estimated using a nonlinear least-squares solver (MATLAB function `lsqcurvefit`). Y_0 was constrained to be positive and thr_E was constrained to be positive and <85 dB SPL. The goodness of fit was quantified by the adjusted R^2 value, representing the proportion of variance in the data explained by the model while controlling for the number of parameters. R^2_{adj} was calculated as $1 - \frac{(1 - R^2)(n - 1)}{n - k - 1}$, where n is the number of observations and k is the number of parameters included in the model.

Statistical analyses. Behavioral TIN thresholds and threshold differences between the fixed and roving-level condition were analyzed using linear mixed-effects models in R (version 3.6.2; Bates et al., 2015). Models incorporated a random effect of animal identity to account for repeated measures within subjects and fixed effects of frequency, level, and the frequency by level interaction. Degrees of freedom for t tests were calculated based on the Satterthwaite approximation, and tests of simple slopes were used to explore significant interactions. Neural TIN thresholds were analyzed using a similar approach but with a random effect of unit identity to account for repeated measures and fixed effects of noise level (low, moderate, and high) and tone frequency (four categories; 1 octave frequency bands centered at 0.5, 1, 2, and 4 kHz). Other analyses included χ^2 tests performed in R, and Pearson correlations conducted in MATLAB.

Results

Budgerigars show similar behavioral TIN sensitivity to that of humans under fixed- and roving-level conditions

Budgerigars were trained to discriminate a tone-plus-noise stimulus from a noise-only standard stimulus using operant-conditioning procedures. Behavioral TIN sensitivity was assessed using adaptive tracking sessions during which the SNR of the TIN stimulus was varied across trials (Fig. 3). TIN detection thresholds were calculated as the mean SNR of the final eight reversal points in the two-down one-up track, which corresponds to ~70.7% correct detection performance (Levitt, 1971). Noise level was held constant during the initial fixed-level part of each tracking session. Thereafter, during roving-level testing, the overall level of the stimulus was randomly varied over a 20 dB range across trials, thereby increasing the complexity of the task. Note that the level shift was applied after combining the tone

Figure 3. Representative behavioral results from two-down one-up adaptive tracking sessions. Thick lines show the mean stimulus SNR of 10 repeated tracking sessions as a function of target trial number. Thin lines show the results of individual sessions. Noise level was fixed during the first part of each track (black), which was followed by a more challenging roving-level test period (blue) for which the overall stimulus level (noise or tone plus noise) was randomly varied over a 20 dB range across trials while preserving stimulus SNR. Individual tracking sessions continued until reversal-based stability criteria were met (see above, Behavioral experiments). Test frequencies are labeled at the top of each column. Noise levels are indicated at the top left of each row (low: 45 [1–4 kHz] or 55 dB SPL [500 Hz]; mid: 65 dB SPL; high: 85 dB SPL). Results are from animal B35.

and noise, thus preserving the SNR of the stimulus on tone-plus-noise trials.

The SNR of the stimulus decreased rapidly over the first 20–30 trials of testing before stabilizing near the animal's fixed-level TIN detection threshold (Fig. 3, gray lines). Fixed-level thresholds generally ranged from –5 to 0 dB SNR across animals and showed minimal variation across test frequencies in moderate- and high-level noise (i.e., 65 and 85 dB SPL, respectively; Fig. 4). In contrast, for the low noise level (45 or 55 dB SPL), most animals showed consistently lower (more sensitive) thresholds with increasing frequency by 1–2 dB per octave. A repeated-measures mixed-model analysis of fixed-level TIN thresholds showed significant effects of frequency ($F_{(3,33)} = 7.24$, $p = 0.0007$) and the frequency-by-noise level interaction ($F_{(6,33)} = 2.57$, $p = 0.037$). The effect of noise level was not significant ($F_{(2,33)} = 0.74$, $p = 0.48$).

Differences in threshold during the second roving-level part of test sessions, known as the rove effect (Fig. 3, blue lines), were generally near zero and almost invariably less than +2 dB. These small rove effects suggest a minimal behavioral impact of this test paradigm, despite the fact that it makes single-channel energy cues less reliable for performing the TIN detection task. A mixed-model analysis showed no significant variation of rove

Figure 4. Behavioral TIN detection thresholds of budgerigars. **A**, TIN detection thresholds in fixed-level noise. **B**, Thresholds shifts under the roving-level condition. Mean noise level is indicated at the top of each column [low: 45 (1–4 kHz) or 55 dB SPL (500 Hz); midlevel: 65 dB SPL; high: 85 dB SPL]. Symbols show thresholds of individual animals; gray bands show 2 SEs above and below the across-subject mean (thick black lines). Thresholds are relatively similar across test frequencies and levels and minimally affected by the roving-level condition. Results are similar to those of normal-hearing human subjects tested previously with the same stimuli (red dotted lines; Leong et al., 2020).

effect with frequency or noise level (frequency, $F_{(3,33)} = 0.38$, $p = 0.77$; noise level, $F_{(2,33)} = 2.07$, $p = 0.14$; frequency $\times$ noise level, $F_{(6,33)} = 1.13$, $p = 0.37$). The average rove effect (0.69 ± 0.53 dB; mean $\pm$ SE) was not significantly different from zero ($t_{(28,2)} = 1.31$, $p = 0.20$).

In summary, tone-detection thresholds in budgerigars decreased slightly with increasing test frequency in low-level noise while showing less variability for moderate- and high-level noise. TIN sensitivity was relatively unaffected by the roving-level test condition, suggesting that this species may use cues other than (or in addition to) the single-channel energy cue, perhaps envelope related, to detect tones in noise.

Finally, budgerigar behavioral thresholds were compared with those reported previously in normal-hearing human subjects, who were tested using stimuli and tracking procedures identical to the present study but with a two-interval discrimination task rather than the single-interval task used in budgerigars (Leong et al., 2020). Both average TIN detection thresholds and the impact of the roving-level paradigm were remarkably similar between budgerigars and humans (Fig. 4, red dotted lines show human results).

Frequency and modulation tuning in the budgerigar IC

Neural recordings were made from a total of 207 multiunit clusters and seven single units in the IC of four awake and unrestrained budgerigars to gain insight into the neural mechanisms underlying behavioral TIN sensitivity. Recordings characterized the basic frequency and modulation tuning properties of neurons as well as TIN responses to CF-matched and behavioral test stimuli. The basic tuning properties of the IC units were similar to those reported previously in this species (Henry et al., 2016,

2017a). Pure-tone frequency response maps typically showed V-shaped tuning curves (Fig. 5A–D), with CFs ranging from 0.4 to 5.8 kHz [median, 2.01 kHz; interquartile range (IQR), 0.95–3.33 kHz; Fig. 5I] and excitatory rate thresholds at sound levels of 20 dB SPL or lower. Inhibitory rate responses were also sometimes observed at frequencies above or below the unit's CF (e.g., note strong below-CF inhibition in Fig. 5D).

Modulation tuning is a dominant response property of IC neurons in birds and mammals and was characterized in budgerigars using MTFs measured with a CF-matched tone carrier (frequency within ± 0.16 octave of CF) in a subset of 159 multiunits and the seven single units. Among these units, all MTFs showed some degree of band-enhanced modulation tuning associated with a greater response rate for a range of modulation frequencies, centered around the BMF, compared with the unmodulated-tone response (Fig. 5E–H). The BMFs of the MTF varied between 54 Hz and 1024 Hz (median, 399 Hz; IQR, 274–482 Hz; 1024 Hz was the highest modulation frequency tested), with no apparent relationship between CF and BMF found for units with CFs > 1 kHz (Fig. 5I; $r = 0.02$, $p = 0.8$, $n = 124$; Pearson's correlation

between log-transformed variables; note, slightly greater BMF variation in the highest CF units). In contrast, for units with CFs < 1 kHz, BMF increased significantly with increasing CF ($r = 0.55$, $p = 0.0002$, $n = 42$).

The strength of modulation tuning was quantified as the normalized difference between the response rates at BMF and to an unmodulated CF tone (i.e., the difference divided by the mean of the two rates). Symbol shading in Figure 5I denotes this response property, with darker shades of gray indicating stronger modulation tuning. Modulation-tuning strength ranged from 0.3 to 2 across units (median, 1.02; IQR, 0.74–1.2) and was consistently highest in units with intermediate CFs (1–4 kHz). No obvious differences in modulation-tuning properties were noted between multiunits and the sample of single units included in the study (Fig. 5I; black circles, multiunits; red squares, single units).

Dependence of TIN responses on energy and envelope cues

CF-matched responses

IC sensitivity to TIN stimuli was initially assessed using tone frequencies matched to the CF of each recorded neuron to maximize overlap between the stimulus spectrum and the frequency region of greatest neural sensitivity as in most prior neurophysiological studies. TIN stimuli were generated by combining a pure tone and a third-octave band-limited noise waveform with the tone frequency log-centered in the noise band and equal to the estimated CF of the IC unit. TIN responses were measured using CF-matched stimuli (tone frequency within ± 0.16 octave of CF) in 157 multiunits and seven single units. Note that the remainder of the 207 multiunits studied were tested with off-CF stimuli only (see below, Off-CF responses). Tones and band-limited

Figure 5. Typical frequency and modulation tuning characteristics of budgerigar IC units. **A–D**, Representative frequency response maps showing response rate as a function of tone frequency and level. Red shading shows the excitatory response area for which the tone response exceeded the spontaneous rate; blue shading shows inhibition. Red solid lines indicate the excitatory threshold tuning curve. CF is indicated within each panel. **E–H**, MTFs showing response rate to sinusoidally amplitude modulated tones with the carrier frequencies (MU1, 0.5 kHz; SU1, 2.0 kHz; MU2, 3.3 kHz; MU3, 6.0 kHz) similar to the estimated CF. Circles indicate the response rate to the unmodulated tone. BMF is indicated within each panel. **I**, BMF increases with increasing CF for units with CFs <1 kHz and appears unassociated with CF in units of higher CF. Multiunits are outlined in black, single units are outlined in red; shading depth is proportional to the strength of modulation tuning. Note that 1024 Hz was the highest modulation frequency tested. **J**, Histogram showing the distribution of IC unit CFs.

noise were presented simultaneously for a stimulus duration of 300 ms. The level of the noise ranged from 35 to 85 dB SPL in 10 dB steps. At each noise level, the tone level was varied to generate SNRs ranging from –12–9 dB in 3 dB steps, and a noise-alone

stimulus ($-\infty$ modulation depth) was included to facilitate calculation of the TIN detection threshold (see below).

For CF-matched stimuli, the proportion of units showing significant variation in the response rate with changing SNRs (i.e., TIN-sensitive units) ranged from 41 to 62% across noise levels. Among TIN-sensitive units, most showed a decreasing response rate with increasing SNR (i.e., decreasing rate-SNR functions; Fig. 6A), a perhaps surprising result considering that the stimulus energy level increases with increasing SNR for these stimuli. The percentage of TIN-sensitive units showing decreasing rate-SNR functions increased from 65% for the noise level of 35 dB SPL to 91% for the noise level of 75 dB SPL. Note that decreasing rate-SNR functions, although negatively correlated to the overall energy level of the stimulus, are expected in neurons with band-enhanced modulation tuning: higher SNR stimuli have smaller normalized envelope fluctuations (Fig. 1), which should evoke less activity from modulation-sensitive cells. In contrast, across noise-alone stimuli at different levels these units showed increasing rates for higher stimulus energy levels, and hence, a positive correlation of the response rate to energy level (Fig. 6A–B). In summary, IC units with decreasing rate-SNR functions displayed response properties consistent with both envelope and energy-based encoding of CF-matched TIN stimuli.

Two regression models were fit to the TIN response rates of individual units (combining responses across noise levels and SNRs in a single analysis) to quantify the amount of variance explained by stimulus energy and envelope cues. The energy cue was calculated as the root mean square amplitude of the stimulus waveform in dB SPL. The envelope cue was calculated by first normalizing the amplitude envelope of the stimulus to a mean of one. The envelope cue was then calculated as the mean of the absolute value of the first derivative of the normalized envelope as in prior studies (i.e., normalized envelope slope; Richards, 1992; Davidson et al., 2009). Note that higher envelope slope values indicate a stimulus with deeper and/or faster envelope fluctuations, whereas lower values indicate a flatter envelope, more similar to that of a pure tone. Variation of both energy and envelope cues across stimuli of the same SNR was because of the random nature of the band-limited noise waveform. The first model consisted of an intercept and two parameters: an energy threshold, below which the response rate was energy independent, and an energy term, indicating the slope of the relationship between energy and response rate for suprathreshold energy levels. This simple energy model provided a relatively poor fit to rate responses in neurons with decreasing rate-SNR functions (Fig. 6A, red dashed lines; $R^2_{adj E} = 0.37$ for example unit MU2; Fig. 6B; note that all model R^2 values throughout this article are adjusted based on the number of model parameters) because the model failed to capture response variation with SNR at each noise level. An alternative energy-plus-envelope model, which incorporated an envelope term and an energy-by-envelope interaction, captured response variations within noise levels and hence improved the model fit (Fig. 6A, purple dashed lines; $R^2_{adj E+env} = 0.77$ for example unit MU2). The weight of envelope-related terms in the combined model was quantified as $[R^2_{adj E+env} - R^2_{adj E}] / R^2_{adj E+env}$ (0.52 for example unit MU2), which corresponds to the proportion of predictable variance accounted for by including these terms. In summary, responses of units with decreasing rate-SNR functions for CF-matched TIN stimuli were best explained by a model combining energy and envelope cues.

In contrast to the decreasing rate-SNR functions found in most IC units, the remainder of TIN-sensitive units (9–35%,

Figure 6. Representative IC neural responses to CF-matched TIN stimuli. **A**, IC mean response rate (black solid lines; error bars indicate SD) as a function of stimulus SNR at five noise levels (35–75 dB SPL, top). Response rates of unit MU2 (CF = 3.33 kHz) decrease with increasing SNR at each noise level. Inset, The frequency response map (Figure 5C) with a black arrow at the tone frequency (2.9 kHz). Vertical black dotted lines indicate neural SNR thresholds above which TIN stimuli are discriminable from noise alone (denoted as $-\infty$). Model fits to the data are shown in red dash-dotted lines [energy-only (E) model, $R^2_{adj} E = 0.37$] and purple dotted lines [energy-plus-envelope ($E+env$) model, $R^2_{adj} E+env = 0.77$, $envelopeweight = 0.52$]. **B**, Response rates from **A** (all noise levels and SNRs) plotted as a function of stimulus energy level. Responses to the same noise level are drawn in the same color. **C**, TIN response patterns, as in **A**, of unit MU1 (CF = 0.54 kHz) showing increasing response rate with increasing SNR. Tone frequency was 0.5 kHz. Model fits: $R^2_{adj} E = 0.88$; $R^2_{adj} E+env = 0.88$; envelope weight = 0.003. **D**, Response rates from **C** plotted as a function of stimulus energy level for each stimulus waveform.

depending on noise level) showed an increasing response rate with increasing SNR at each noise level (i.e., increasing rate-SNR functions; Fig. 6C). The response rate in these units was positively correlated with the overall energy level of the stimulus, both within and across noise levels. Consequently, for units showing this trend, the energy model explained a large proportion of the variance in response rates across noise levels and SNRs (Fig. 6C, red dashed lines; $R^2_{adj} E = 0.88$ for example unit MU1; Fig. 6D), and the addition of envelope-related terms to the model failed to substantially improve the goodness of fit (Fig. 6C, blue dashed lines; $R^2_{adj} E+env = 0.88$ for example unit MU1; envelope weight = 0.003).

In summary, IC responses to CF-matched TIN stimuli were most often correlated with both the energy level and envelope structure of the stimulus, resulting in decreasing rate-SNR functions within noise levels. Less commonly, and particularly for low-CF units and at low sound levels, as expanded on in subsequent sections, IC TIN responses had increasing rate-SNR functions that were readily explainable by a simple energy model.

Off-CF responses

IC neural responses to TIN stimuli were also recorded for test frequencies up to several octaves away from the estimated CF to test the possible utility of off-frequency neural channels for TIN detection. Off-frequency neurons are rarely considered in neurophysiological studies of TIN detection but could theoretically contribute to behavioral sensitivity based on a substantial spread of excitation across neural frequency channels at moderate-to-high stimulus levels. Note that the tone frequency remained log-centered in third-octave band-limited noise for off-CF stimuli and ranged from 0.5 to 4 kHz in octave steps to match the behavioral experiment. Neural responses to off-CF TIN stimuli (test frequencies more than ± 0.16 octave from CF) showed the same increasing and decreasing rate-SNR functions (Fig. 7) described

above; although surprisingly, both response trends were frequently observed in the same unit depending on the test frequency. As illustrated in Figure 7 for a representative unit, and discussed later, increasing rate-SNR functions were more common for test frequencies below CF (Fig. 7A), whereas higher test frequencies (e.g., at CF in Fig. 7C) tended to evoke decreasing rate-SNR functions.

The same multiple-regression models used above to explain CF-matched TIN responses were also applied to off-CF responses to further explore the possible dependence of IC TIN responses on stimulus energy and envelope cues. For units showing increasing rate-SNR functions in response to off-CF TIN stimuli, the energy model once again provided a good fit to the responses ($R^2_{adj} E = 0.92$ for example unit SU1; Fig. 7A, red dashed lines). Note the threshold term in these units accounted for invariance of the response rate below a minimum stimulus energy level (~ 65 dB SPL for example unit SU1 in Fig. 7B; median across the neural population, 46 dB SPL; IQR, 37–62 dB SPL) and that higher thresholds could generally be attributed to the stimulus spectrum falling outside the sensitive region of a unit's frequency response map (Fig. 7A, inset). Adding envelope terms to the model failed to appreciably increase the goodness of fit in units with increasing rate-SNR functions in response to off-CF TIN stimuli.

For units with decreasing rate-SNR functions for off-CF TIN stimuli, the energy model showed a poor fit to the responses, as expected (Fig. 7C, red dashed lines; $R^2_{adj} E = 0.24$ for example unit SU1; Fig. 7D), and adding envelope-related terms (envelope and the energy-by-envelope interaction) substantially improved the goodness of fit (Fig. 7C, blue dotted line; $R^2_{adj} E+env = 0.80$ for example unit SU1; envelope weight = 0.70). These results show that similar to the CF-matched responses, rate responses to off-CF TIN stimuli vary with both energy and envelope cues present in the stimulus, raising the possibility that off-CF neural channels also contribute to behavioral TIN detection.

Figure 7. Representative IC responses to off-CF TIN stimuli. **A, C, E**, IC TIN response patterns as in Figure 6, but for off-CF stimuli. **A**, Unit SU1: CF = 2.04 kHz, test frequency = 1 kHz; model fits: $R^2_{adj, E} = 0.92$, $R^2_{adj, E+env} = 0.93$, envelope weight = 0.014. **C**, Unit SU1: CF = 2.04 kHz, test frequency = 2.4 kHz; model fits: $R^2_{adj, E} = 0.24$, $R^2_{adj, E+env} = 0.80$, envelope weight = 0.70. **E**, Unit MU3: CF = 5.83 kHz, test frequency = 2 kHz; model fits: $R^2_{adj, E} = 0.63$, $R^2_{adj, E+env} = 0.63$, envelope weight = 0. **B, D, F**, Response rates from **A, C, E**, respectively, as a function of stimulus energy level, as in Figure 6.

Trends across the neural population

The multiple-regression models applied above in representative units displaying increasing (Figs. 6C, 7A) and decreasing (Figs. 6A, 7C) rate-SNR functions were applied to all units to further explore dependence of IC rate on energy and envelope cues across the IC population. Adjusted R^2 values were calculated for the energy-only model and the combined energy-plus-envelope model in each unit to assess goodness of fit. Moreover, for units in which R^2_{adj} of the combined model was at least 0.5, the relative weighting of the envelope terms in the combined model was calculated as $(R^2_{adj, E+env} - R^2_{adj, E})/R^2_{adj, E+env}$, a quantity corresponding to the proportion of predictable variance accounted for by adding in the envelope-related terms. Adjusted R^2 values of the energy-only and combined models are shown in the left and center columns of Figure 8, respectively, where red-filled circles show results from units with increasing rate-SNR functions at most noise levels, and blue-filled circles show results from units with decreasing rate-SNR functions at most noise levels. Results are not shown for TIN-insensitive units. The right column shows the weight of the envelope terms in the combined model. Units were divided into three groups with low ($n = 106$), medium ($n = 76$), and high CFs ($n = 18$) to test for possible variation in energy and envelope coding across neural channels tuned to different frequency ranges.

In units with CFs in the low and medium ranges (i.e., CFs <2 kHz and from 2 to 4 kHz, respectively; Fig. 8A–F), the increasing rate-SNR curves were more common when the tone frequency was lower than CF, whereas higher tone frequencies tended to evoke decreasing rate-SNR curves. The energy-only

model provided a good fit to the increasing rate-SNR curves in these units, with R^2_{adj} values up to 0.97 (Fig. 8A,D) and relatively low weighting of the envelope term in the combined model (Fig. 8C,F) because of minimal improvement of R^2_{adj} on adding these terms. The R^2_{adj} value of the energy model dropped sharply when the tone frequency became equal to or higher than CF, whereas rate-SNR functions showed a concomitant shift to the decreasing rate-SNR curves (tone frequency/CF > 0; Fig. 8A,D). For tone frequencies at or above CF, the addition of the envelope terms greatly improved the model fit (see R^2_{adj} values of the combined model in Fig. 8B,E), resulting in a large increase in the weight of envelope terms (Fig. 8C,F). These results show that for most IC units, TIN responses to test frequencies below CF were positively correlated with stimulus energy and well explained by a simple energy model. In contrast, for test frequencies at or above CF, response rates became increasingly dependent on envelope cues.

In contrast to units with CFs <4 kHz, responses of higher-CF units appeared more likely to show decreasing rate-SNR functions for test frequencies below CF (Fig. 8G–J). Whereas the decreasing rate-SNR functions in lower-CF units (CF < 4 kHz) was readily explained by dependence of response rate on envelope cues, as outlined above, rate responses of higher-CF units were often well explained by the energy model (Fig. 8G) and showed little or no increase in R^2_{adj} with the addition of envelope terms to the model (Fig. 8H). Consequently, the weight of the envelope term ($n = 8$ units with decreasing rate-SNR functions) was relatively low (Fig. 8I). Indeed, envelope weight was lower in these high-CF units than in low-CF ($t_{(18)} = -8.20$, $p < 0.0001$) and mid-CF ranges ($t_{(16)} = -10.64$, $p < 0.0001$), and slightly lower in

Figure 8. Model fits to IC TIN responses. **A**, E -model fits (adjusted R^2 values) to TIN responses of units with CFs below 2 kHz. Model fits are shown as a function of normalized test frequency (octave scale relative to CF). The dashed vertical line indicates a test frequency equal to CF. Results from units with different rates of SNR functions are drawn with different symbols (red filled circles, increasing at most noise levels; blue filled circles, decreasing at most noise levels; black open circles, no dominant pattern across noise levels). **B**, $E+env$ -model fits to TIN responses of units with CFs < 2 kHz, plotted as a function of normalized test frequency as in **A**. Adjusted R^2 values above the horizontal dotted line exceed 0.5. **C**, Envelope weights in units with CFs below 2 kHz, plotted as a function of normalized test frequency. Envelope weight is the proportion of variance that was predicted by the $E+env$ -model (only shown in units for which the adjusted R^2 value of the $E+env$ -model exceeded 0.5). Symbol meanings are as in **A**. **D**, **E**, **F**, E -model fit, $E+env$ -model fit, and envelope weight, respectively, as in **A**, **B**, **C**, for units with CFs from 2 to 4 kHz. **G**, **H**, **I**, E -model fit, $E+env$ -model fit, and envelope weight, respectively, as in **A**, **B**, **C**, for units with CFs > 4 kHz. Results are from 106 IC units in **A** and **B**, 104 in **C**, 76 in **D** and **E**, 74 in **F**, and 18 in **G**, **H**, and **I**.

mid-CF units than low-CF units ($t_{(300)} = -2.44$, $p = 0.015$). Closer inspection of responses for a representative high-CF unit, MU3 (Fig. 7E), showed that the response rate was negatively correlated with energy both within and across noise levels, potentially because of strong inhibitory sidebands at below-CF tone frequencies in the frequency response map shown for MU3 (Fig. 5D). Slightly weaker modulation tuning in these high-CF units (Fig. 5H,I) might also explain why response rate was largely energy dependent.

Off-frequency channels account for behavioral TIN thresholds under fixed-level conditions

On-frequency neural thresholds

Thresholds for behavioral TIN detection are often thought to depend on the thresholds of neural channels for which the CF of the processing channel is matched to the frequency of the target tone. To test this assumption, neural thresholds for CF-matched TIN detection were calculated in each IC unit based on ROC analysis of rate responses to stimuli of varying SNR. Threshold was defined for each noise level as the lowest SNR above which separation between the rate distribution for the noise-only stimulus and the rate distribution for TIN stimuli exceeded 70.7% (distributions were based on 20 response repetitions for each stimulus SNR; note that 70.7% is the correct performance level of an unbiased observer performing a two-down one-up tracking session; Levitt, 1971). Most units showed a threshold for CF-matched TIN detection within the range of SNRs tested (−12 to +9 dB) for at least one noise level (131/157 multiunits and 5/7 single units; Fig. 9). The proportion of units without thresholds in the tested range of SNRs (i.e., TIN insensitive) varied across CF ranges (octave-wide ranges log-centered at 0.5, 1, 2, and 4 kHz) and with noise level (Fig. 9, histograms). TIN-insensitive

responses were significantly more common in the highest-CF range ($\chi^2 = 65.54$, $df = 3$, $p < 0.001$) and for lower noise levels ($\chi^2 = 23.64$, $df = 4$, $p < 0.001$).

CF-matched, TIN-detection thresholds were similar across noise levels (Fig. 9), with no main effect of noise level revealed by statistical analysis ($F_{(4,436)} = 0.83$, $p = 0.50$; mixed-effects model). In contrast, thresholds varied across CF ranges ($F_{(3,142)} = 6.54$, $p < 0.001$). Neural thresholds in the 1 and 2 kHz CF ranges (CFs from 0.7 to 2.8 kHz) were lowest, with the most sensitive units having thresholds slightly lower (i.e., more sensitive) than the behavioral thresholds of trained animals (Fig. 9, black circles; error bars show the across-subject SD). In contrast, neural thresholds in the 0.5 and 4 kHz CF ranges were higher and rarely, if ever, as sensitive as those observed behaviorally. The insensitivity of neural thresholds in the 0.5 kHz CF range could be because of the relatively small number of units found with low CFs (i.e., sparse sampling; Fig. 5J). However, among the 34 units found with CFs within ± 0.25 octave of 4 kHz, 21–28 units (depending on the noise levels) did not have a CF-matched TIN detection threshold (Fig. 9, histograms), and others had thresholds considerably less sensitive than those observed behaviorally. These results suggest that although CF-matched TIN responses may be adequate to account for behavioral TIN detection at moderate test frequencies (0.7–2.8 kHz), these responses appear insufficient to explain behavioral sensitivity to low- or high-frequency TIN stimuli in this species.

CF-matched TIN detection thresholds of units displaying increasing and decreasing rate-SNR functions in response to TIN stimuli are shown with red upward-facing and blue downward-facing triangles, respectively, in Figure 9. Although the increasing pattern was generally less common, it was observed more frequently in low-CF units and at low noise levels. Moreover, TIN detection thresholds were slightly higher in units showing increasing rate-SNR functions compared with units with decreasing rate-SNR functions (2.4 ± 0.6 ; least-squares mean difference $\pm$ SE; $t_{(304)} = -4.13$, $p < 0.001$; *post hoc* comparison of least-squares means; though note difficulty controlling for the potential effect of CF).

Off-frequency neural thresholds

To test whether off-frequency channels might exhibit greater TIN sensitivity, particularly at test frequencies for which on-frequency channels appeared insufficient (i.e., at 0.5 and 4 kHz), neural thresholds were also evaluated at test frequencies up to several octaves away from CF using the same ROC-based approach applied above for CF-matched stimuli. Thresholds are shown as a function of CF in Figure 10, where each column is one of four frequencies tested in most units (i.e., 0.5, 1, 2, and 4 kHz; sample sizes of 83, 109, 147, and 99 units, respectively) and each row is a different noise level. Thresholds plotted outside the purple vertical bands in each panel were considered off CF,

Figure 9. CF-matched neural thresholds for fixed-level TIN detection. Thresholds are plotted as a function of test frequency. Noise level in dB SPL is indicated at the top left of each panel. Thresholds of units showing increasing and decreasing rate-SNR functions are drawn with red upward-pointing and blue downward-pointing triangles, respectively. Mean behavioral TIN detection thresholds are drawn with black circles; error bars indicate the mean SD across animals. Histograms show the total number of units (black lines) and the number of units without a TIN detection threshold (gray filled area); bin width is 0.33 octave; 164 units were tested with CF-matched TIN stimuli.

because the unit's CF was more than ± 0.16 octave from the test frequency. Surprisingly, off-CF thresholds for TIN detection could be equally if not more sensitive than thresholds for CF-matched stimuli, especially for high noise levels where a large proportion of neurons responded to all stimuli regardless of CF. Notably, at test frequencies of 0.5 and 4 kHz, where CF-matched neural thresholds appeared potentially inadequate to explain behavioral thresholds (horizontal lines), a large number of off-CF responses had sensitive thresholds within the behavioral range. This result suggests that off-CF neural channels may make an important contribution to behavioral TIN detection.

Pearson correlations were used to evaluate the extent to which neural TIN sensitivity might depend on energy-based versus envelope-based coding of these stimuli. Results are shown in Figure 11 for units with both increasing (red symbols) and decreasing

(blue symbols) rate-SNR functions. For units with increasing rate-SNR functions, TIN thresholds were lower when the adjusted R^2 value of the energy model was high (Fig. 11A; $r = -0.31$, $p < 0.001$) and uncorrelated with the log-transformed envelope weight in the model combining energy and envelope cues (Fig. 11B; $r = -0.10$, $p = 0.18$). On the other hand, for units displaying decreasing rate-SNR functions (i.e., in putatively more envelope-sensitive units), TIN thresholds were unassociated with the adjusted R^2 value of the energy model (Fig. 11A; $r = 0.04$, $p = 0.38$) and decreased markedly with increasing log-transformed envelope weight in the combined model (Fig. 11B; $r = -0.21$, $p < 0.001$). These results show that stronger dependence of response rate on envelope cues, and to a lesser extent on energy, was associated with lower neural threshold for TIN detection.

Finally, we tested for relationships of TIN detection thresholds and envelope weight with the strength of amplitude-modulation tuning, as measured from the traditional MTF using a CF tone carrier (Fig. 11C,D). Only units for which the test frequency was within ± 0.16 octave of CF were included in these analyses. For neurons with decreasing rate-SNR functions, TIN thresholds tended to decrease with increasing modulation tuning strength as might be expected ($r = -0.29$, $p = 0.002$), whereas envelope weight was uncorrelated with modulation tuning strength ($r = 0.14$, $p = 0.19$). Neurons with increasing rate-SNR functions showed no association of TIN threshold ($r = 0.34$, $p = 0.064$) or envelope weight ($r = 0.22$, $p = 0.30$) with modulation tuning strength. In summary, these results suggest that the strength of amplitude-modulation tuning based on the MTF has limited capacity to predict aspects of TIN responses.

Sensitivity of the pooled neural population

The TIN detection thresholds of individual IC units discussed above suggest that off-frequency neural channels are needed to explain behavioral TIN sensitivity under some test conditions but do not rule out the possibility that on-frequency neural responses might be sufficient after pooling responses across units. To test this hypothesis, we calculated population-level neural thresholds through optimal pooling of information across individual IC units using a maximum likelihood-based decoder analysis (Jazayeri and Movshon, 2006; Day and Delgutte, 2013, 2015). Population thresholds were computed for both CF-matched responses only (i.e., responses for which CFs were within ± 0.16 octave of the tone frequency; Fig. 12, purple lines), and all responses (black lines). Thresholds were evaluated for each test frequency and level as a function of sample size as the number of units also influences decoder performance. For each sample size investigated, units were selected randomly 20 times to determine the median and interquartile range of the model's performance.

TIN thresholds based on the CF-matched neural population were as sensitive as the behavioral thresholds of trained animals at test frequencies of 1 and 2 kHz but were 5–10 dB higher than the 4 kHz behavioral threshold despite a seemingly adequate sample size of units ($n = 17$). Thus, CF-matched neural responses appeared insufficient to explain behavioral TIN sensitivity at 4 kHz, even when the information from the individual units with similar CFs was combined. In contrast, introducing off-CF responses into the pooling procedure resulted in IC population thresholds that were typically low enough to explain behavioral performance and notably lower (more sensitive) under some conditions than CF-matched thresholds calculated with the same number of units (e.g., at 65 and 75 dB SPL for 2 and 4 kHz test frequencies). These results further underscore the likely

Figure 10. Fixed-level TIN detection thresholds across CFs at behavioral test frequencies. Test frequencies are indicated at the top of each column; noise level is at the right of each row. Thresholds of units with decreasing and increasing rate-SNR functions are marked with blue downward-pointing and red upward-pointing triangles, respectively. Mean behavioral thresholds are indicated with horizontal dash-dot lines; dotted lines are 1 SD above and below the mean. Histograms show the total number of units (black lines) and the number of TIN-insensitive units (gray-filled area); bin width is 0.33 octave. The total number of IC units tested was 83, 109, 147, and 99 at test frequencies of 0.5, 1, 2, and 4 kHz, respectively.

importance of off-CF neural channels for perception of TIN stimuli, especially for relatively low and high test frequencies.

Roving-level neural thresholds

Random stimulus-level variation for the roving-level condition decreases the reliability of single-channel energy cues compared

with fixed-level listening and was therefore expected to result in higher neural thresholds as IC responses generally varied with stimulus energy in addition to the envelope cue. Neural thresholds for roving-level TIN detection were estimated by combining responses across 55, 65, and 75 dB SPL noise levels (three levels vs 21 possible levels spanning the same 20 dB range in behavioral

Figure 11. Variation in neural fixed-level TIN thresholds with model fits based on energy and envelope cues. **A**, Threshold versus E -model fit in units with increasing and decreasing rate SNR functions. **B**, Threshold versus envelope weight in units with decreasing rate-SNR functions. **C**, Envelope weight versus modulation-tuning strength in units with increasing and decreasing rate-SNR functions. **D**, Threshold versus modulation-tuning strength in units with increasing and decreasing rate-SNR functions. Unit thresholds shown are the lowest value observed across all tested noise levels. Results are from 200 units in **A**, 196 units in **B**, and 98 units for which $R^2_{adj E+env}$ exceeded 0.5 in **C** and **D**.

experiments) and applying the same ROC analysis described above for calculation of fixed-level neural thresholds. Consistent with predictions, fewer units had measurable thresholds for the roving-level condition (Fig. 13) than for fixed-level noise at 65 dB SPL (Fig. 13A). Furthermore, among units with a measurable roving-level threshold, nearly all had CFs less than the tone frequency and showed decreasing rate-SNR functions. These results further underscore the likely importance of off-frequency neural channels and envelope-based encoding for TIN sensitivity, especially under challenging listening conditions with random variation in background noise level.

Because all IC units showed some dependence of TIN responses on energy, neural thresholds were expected to increase under the roving-level conditions. Rove effects in IC neurons were generally positive (Fig. 13B), consistent with this expectation, and were typically greater than the small rove effects observed in the behavioral experiments (note that many neurons had no measurable roving-level threshold and therefore do not appear in Fig. 13). Nonetheless, at all test frequencies except 500 Hz, a small proportion of IC neurons had roving-level thresholds that were approximately as sensitive as those of behaviorally trained animals (Fig. 13A). These results suggest that the response properties of the most sensitive IC neurons could be sufficient to account for roving-level TIN sensitivity in budgerigars.

To evaluate the potential benefit of combining responses across neurons for roving-level TIN detection, given the relatively small proportion of neurons with sensitive roving-level thresholds, we first used the maximum likelihood-based decoder analysis to calculate IC neural population thresholds based on all units as in the previous analysis of fixed-level results. Population thresholds were calculated using 20 stimulus repetitions per

SNR, randomly selected across the three noise levels, for direct comparison to the fixed-level results. Roving-level population thresholds at test frequencies of 0.5, 1, 2, and 4 kHz were 2.81 ± 0.46 , 1.71 ± 0.12 , -0.88 ± 0.06 , -1.88 ± 0.61 dB, respectively (means $\pm$ SD across three analyses; Fig. 12, red stars); that is, generally a few dB higher than the fixed-level thresholds.

As an alternative approach to test the utility of combining responses for roving-level TIN detection, we first simulated fixed- and roving-level TIN responses in two model neurons, one for which the response rate depended on both energy and envelope cues (based on the model fit for unit MU2; Fig. 6A; decreasing rate-SNR function) and a second for which TIN encoding was primarily energy dependent (based on MU1; Fig. 6C; increasing rate-SNR function). For the roving-level condition, the noise level varied over a 20 dB range across trials based on a random uniform distribution with 1 dB resolution, as in the behavioral experiments. Single-trial responses were predicted based on the multiple regression models described above, which explained 87% of the variance in response rates of MU1 and 78% of the variance in response rates of MU2.

The simulated TIN thresholds of the energy-dependent model unit, which showed an increasing rate-SNR function, increased by 17.7 dB between the fixed- and roving-level noise conditions (Fig. 14A,B). In contrast, the predicted threshold of the energy-and-envelope-dependent unit, with the decreasing rate-SNR function, increased by 6.2 dB in roving-noise level (Fig. 14D,E). These results show that incorporating envelope coding can reduce the effect of the roving stimulus level on TIN detection thresholds. However, note that because of the partial correlation of response rate to energy in MU2 and other envelope-sensitive units, the predicted neural rove effect was still larger than the average behavioral rove effect of ~ 0.7 dB (Fig. 4).

Finally, we tested whether a model neuron that received an excitatory input from a typical IC unit with a decreasing rate-SNR function (i.e., MU2; energy-plus-envelope based) and an inhibitory input from a typical IC unit with an increasing rate-SNR function (i.e., MU1; energy-based) could better account for roving-level behavioral TIN detection (Fig. 14E). This model structure was based on the premise that inhibition by the energy-dependent input might partly counteract energy dependence of the response rate in the energy-plus-envelope based input, resulting in a model response related primarily to envelope structure (the cue unaffected by roving stimulus level). For each condition (fixed and roving level), response rates of the two inputs were normalized to have an SD of one for the noise-alone stimulus. The decision variable of the model neuron was calculated as $R_{E+env} - I \cdot R_E$, where R_{E+env} and R_E were the standardized response rates of the energy-plus-envelope (excitatory) and energy-dependent (inhibitory) inputs, respectively, and I was the normalized strength of the inhibitory input ($I = 1$ indicates equal inhibitory and excitatory strengths; note that both inputs were standardized by dividing by the SD observed for the noise-alone condition). The smallest rove effect of 0.69 dB was found when I was 0.88.

These model simulations show that a simple model neuron receiving roughly equal amplitude excitatory input from an envelope-sensitive IC unit and inhibitory input from an energy-sensitive unit substantially improves TIN sensitivity under the roving-level condition. Indeed, the rove effect observed in this model unit was remarkably similar to the average rove effect observed in behavioral experiments (~ 0.7 dB in both cases). Note that this model could take the form of one based on

Figure 12. Neural fixed-level population thresholds estimated with a maximum-likelihood-based decoder analysis. Test frequencies are indicated at the top of each column; noise level is at the left of each row. Population thresholds are plotted for on-frequency units with CFs within ± 0.16 octaves of the test frequency (purple) and all CFs (black) as a function of the number of units included in the analysis. Thick dark lines and shaded regions indicated the median and interquartile range of the population threshold across different randomly selected groups of neurons. Mean behavioral thresholds are indicated with horizontal dash-dot lines. Dotted lines are 1 SD above and below the mean. Mean roving-level population thresholds are indicated with red stars.

response differences across processing channels, where the inhibitory input arises from a CF channel above the tone frequency (commonly showing energy-dependent, increasing responses) and the excitatory input arises from a CF channel equal to or less than the tone frequency (typically an envelope-based, decreasing response).

Discussion

This study compared behavioral and midbrain-level neural sensitivity to TIN stimuli in a single species, the budgerigar. Budgerigar behavioral thresholds for TIN detection were similar to those of humans across the full range of frequencies and noise levels tested. Also as in humans, budgerigars showed minimal threshold shifts for TIN detection under a roving-level condition with random variation in noise level. IC neural responses to CF-matched stimuli (tone frequency within ± 0.16 octave of CF)

measured in awake budgerigars were sensitive enough to explain behavioral TIN thresholds for frequencies from 0.7 to 2.8 kHz. In contrast, off-CF neural responses were required to account for behavioral performance at other frequencies. IC TIN responses could usually be predicted by a model combining energy and envelope cues, and in other cases by a simple energy model. A model neuron receiving input from both neuron types was able to achieve rove resistance of TIN detection thresholds similar to the level found behaviorally.

Similar behavioral TIN sensitivity between budgerigars and humans suggests that these species may use the same cues to perform the task. A previous study in budgerigars quantified the pattern of hit and false alarm rates for 500 Hz TIN detection across an ensemble of reproducible noise waveforms, known as the detection pattern (Henry et al., 2020). Budgerigar detection patterns were significantly correlated to those of human subjects

Figure 13. Roving-level TIN sensitivity across CFs at behavioral test frequencies. **A**, Roving-level thresholds of IC neurons calculated from responses pooled across noise levels of 55, 65, and 75 dB SPL. Test frequencies are indicated at the top of each column; thresholds of units with decreasing and increasing rate-SNR functions are marked with blue downward-pointing and red upward-pointing triangles, respectively. Mean roving-level behavioral thresholds are indicated with horizontal dash-dot lines; dotted lines are 1 SD above and below the mean. **B**, Neural rove effects showing the threshold difference between the fixed- (65 dB noise level) and roving-level condition. Positive values indicate a higher roving-level threshold. Dash-dot and dotted lines indicate the mean and SD of the behavioral rove effect. **C**, Histograms show the total number of units (black lines) and the number of TIN-insensitive units under the roving-level condition (gray bars); bin width is 0.33 octave. The total number of IC units tested was 83, 109, 147, and 99 at test frequencies of 0.5, 1, 2, and 4 kHz, respectively.

tested with the same noise waveforms (Evilsizer et al., 2002; Davidson et al., 2006; Mao et al., 2013) and could be predicted along with trial-by-trial responses by a simple psychophysical model combining energy and envelope cues (Henry et al., 2020). This same energy-plus-envelope model was found in a subsequent budgerigar study to predict substantial variance in behavioral TIN responses for test frequencies up to 4 kHz (Henry and Abrams, 2021). An alternative model based on energy differences across frequency channels could also predict 500 Hz behavioral results, whereas models including temporal fine structure generally failed to explain responses to noise-alone trials (Henry et al., 2020). Together, these studies suggest that budgerigars and humans both rely on energy and envelope-based cues for TIN detection.

IC responses to TIN stimuli usually showed decreasing rate-SNR functions and could be predicted by a model combining energy and envelope cues. Decreasing rate-SNR functions were most common when the tone frequency was equal to or greater than CF. In other cases, especially when neurons were tested with tone frequencies lower than CF, we observed increasing rate-SNR functions consistent with predictions of the energy model. These same basic rate-SNR curves have also been found in the mammalian IC (Jiang et al., 1997; Ramachandran et al., 2000; L. Fan, KS Henry, and LH Carney, unpublished observations; Rocchi and Ramachandran, 2018), for which they have frequently been interpreted in the context of frequency/level-dependent excitation and inhibition, according to the frequency-response map (Ramachandran et al., 1999). In contrast, a rabbit study (L. Fan, KS Henry, and LH Carney, unpublished observations) and the present results highlight the potential importance of envelope cues and midbrain-level amplitude-modulation tuning for TIN detection. Many IC

neurons in birds and mammals show increasing response rate for stimuli with greater depth of envelope fluctuations over a limited range of amplitude-modulation frequencies, a well-known response type known as band-enhanced modulation tuning (Langner and Schreiner, 1988; Kim et al., 2015, 2020). Because the envelope of TIN stimuli becomes flatter with increasing SNR (Fig. 1), it follows that these modulation-tuned neurons should show a decreasing rate with an increasing SNR, as was observed here and in many rabbit IC neurons by L. Fan, KS Henry, and LH Carney. (unpublished observations). Our finding that a model combining energy and envelope cues predicted most decreasing rate-SNR functions further supports the interpretation that modulation tuning for envelope fluctuations is a key factor shaping TIN neural responses. On the other hand, a small number of neurons was identified, mostly with relatively high CFs and weak modulation tuning, for which the decreasing rate-SNR curve was attributable to inhibition by stimulus energy at the test frequency rather than by an envelope-based processing mechanism. In summary, although most IC TIN responses were correlated to energy and envelope cues and showed decreasing rate-SNR functions, other responses were correlated to energy alone. These energy-dependent neurons typically showed increasing rate-SNR functions but in rare cases showed decreasing rate-SNR curves.

For stimuli presented within ± 0.16 octave of units' CFs, TIN detection thresholds of individual IC units were sensitive enough to explain behavioral performance for test frequencies from 0.7 to 2.8 kHz in fixed-level noise. In contrast, a large proportion of higher-CF units ($CF > 2.8$ kHz) showed no threshold for CF-matched TIN detection, at least within the range of SNRs tested (up to +9 dB). Furthermore, thresholds of the few higher-CF TIN-sensitive units were substantially less sensitive than

Figure 14. Neural processing of roving-level TIN stimuli. **A**, Predicted mean rate-SNR function of an *E*-dependent model IC unit under fixed- (black) and roving-level (blue) test conditions. Error bars indicate the SD. Dotted vertical lines denote the threshold SNR above which neural TIN detection exceeds 70.7%. **B**, Neurometric functions based on rate-SNR responses from **A**, plotting the percentage of correctly identified stimuli across stimulus SNRs. **C**, Predicted rate-SNR functions, for fixed- and roving-level conditions as in **A**, of an *E+env*-dependent IC unit. **D**, Neurometric functions based on rate-SNR responses from **C**. **E**, Predicted fixed- and roving-level TIN thresholds of a model neuron receiving excitatory input from the *E+env*-dependent unit (in **C**) and inhibitory input from the *E*-dependent unit (in **A**). Open and filled circles indicate thresholds of the *E*-dependent and *E+env*-dependent inputs, respectively.

behavioral thresholds. Optimal pooling of response rates (Jazayeri and Movshon, 2006; Day and Delgutte, 2013) across CF-matched responses was investigated as a possible mechanism to improve performance in the high-CF range, but the population neural threshold remained insufficient to account for observed behavioral thresholds despite a seemingly adequate sample of units. Off-CF neural responses were also explored by presenting TIN stimuli at frequencies up to several octaves from CF. Perhaps surprisingly, a large number of off-CF IC units were found that had thresholds equal to or lower than those observed behaviorally across the full range of test frequencies. Thus, while hearing in noise is often assumed to depend on neural channels tuned in frequency to the target signal, these results highlight the

potential importance of off-CF neural channels for masked detection.

Whereas responses of individual IC units that were sensitive to both energy and envelope cues could generally account for behavioral TIN thresholds under both fixed- and roving-level conditions, the proportion of TIN-sensitive units was considerably smaller for the roving-level condition. Furthermore, in model simulations, the expected roving-level threshold shift in MU2 was 6.2 dB, considerably higher than the 0.7 dB roving-level threshold shift observed behaviorally. To explore whether a combination of neurons that were sensitive to energy and energy-plus-envelope might provide better resistance to the effect of roving level, we tested thresholds of an upstream model neuron receiving excitatory input from a typical IC unit with a decreasing rate-SNR function (energy- and envelope dependent) and inhibitory input from a typical unit with an increasing rate-SNR function (energy dependent). The rove effect of the model unit was lowest (0.69 dB) and consistent with behavioral results when the strength of the inhibitory input was 0.88 times that of the excitatory input. Further studies are needed to determine whether neurons of this type exist at higher auditory processing levels.

It remains unknown why envelope dependence of IC responses was typically weaker for test frequencies lower than the CFs of units. One possible explanation is that neural responses in this frequency range were dominated by excitation or inhibition, whereas balanced excitation and same-frequency inhibition is thought to produce modulation tuning in midbrain neurons, with inhibitory input lagging excitation (Nelson and Carney, 2004, 2007). Indeed, several experimental studies have shown that pharmacological blockage of GABAergic inhibition alters rate-based modulation transfer functions of IC units (Burger and Pollak, 1998; Caspary et al., 2002, 2008; Zhang and Kelly, 2003); thus, weaker envelope dependence of neural responses is expected at test frequencies for which responses are dominated primarily by excitation or inhibition. Further study is needed to test for differences in modulation tuning across test frequencies because MTFs were assessed using CF-matched stimuli only. Finally, the question of how IC TIN responses compare between birds and mammals requires further attention. Although responses to off-CF stimuli have not been studied systematically in mammals, Rocchi and Ramachandran (2018) noted an increasing rate-SNR function in a macaque IC unit tested with below-CF TIN stimuli, a result consistent with our finding in budgerigars. Moreover, Jiang et al. (1997) observed decreasing rate-SNR functions in a small proportion of units with CFs higher or lower than the 500 Hz tone frequency used in their study. However, note that both of these studies use wideband noise. L. Fan, KS Henry, and LH Carney (unpublished observations) used the same 0.33 octave noise bandwidth used in budgerigars but did not consider off-CF responses. In general, the existence of broadly similar TIN responses between birds and mammals agrees with previous studies highlighting conserved auditory neural-processing mechanisms between these groups, from auditory-nerve fibers to the level of cortical microcircuits (Sachs et al., 1974; Manley et al., 1985; Salvi et al., 1992; Woolley and Portfors, 2013; Calabrese and Woolley, 2015).

In conclusion, behavioral and midbrain-level sensitivity to TIN stimuli was investigated in budgerigars. Behavioral TIN sensitivity was similar to that of humans (Leong et al.,

2020) and minimally affected by a roving-level paradigm for which single-channel energy cues are unreliable, consistent with previous budgerigar studies (Henry et al., 2020; Henry and Abrams, 2021). Neural recordings from IC single- and multiunits in awake budgerigars highlighted the importance of envelope encoding and off-frequency neural channels not tuned to the target frequency for TIN processing. Furthermore, neural modeling results showed that a combination of energy- and envelope-dependent neurons could enhance TIN sensitivity under challenging roving-level conditions with random variation in noise level.

References

- Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* 67:1–48.
- Baumann S, Griffiths TD, Sun L, Petkov CI, Thiele A, Rees A (2011) Orthogonal representation of sound dimensions in the primate midbrain. *Nat Neurosci* 14:423–425.
- Burger RM, Pollak GD (1998) Analysis of the role of inhibition in shaping responses to sinusoidally amplitude-modulated signals in the inferior colliculus. *J Neurophysiol* 80:1686–1701.
- Calabrese A, Woolley SMN (2015) Coding principles of the canonical cortical microcircuit in the avian brain. *Proc Natl Acad Sci U S A* 112:3517–3522.
- Carney LH, Ketterer AD, Abrams KS, Schwarz DM, Idrobo F (2013) Detection thresholds for amplitude modulations of tones in budgerigar, rabbit, and human. *Adv Exp Med Biol* 787:391–398.
- Caspary DM, Palombi PS, Hughes LF (2002) GABAergic inputs shape responses to amplitude modulated stimuli in the inferior colliculus. *Hear Res* 168:163–173.
- Caspary DM, Ling L, Turner JG, Hughes LF (2008) Inhibitory neurotransmission, plasticity and aging in the mammalian central auditory system. *J Exp Biol* 211:1781–1791.
- Choi JH, Jung HK, Kim T (2006) A new action potential detector using the MTEO and its effects on spike sorting systems at low signal-to-noise ratios. *IEEE Trans Biomed Eng* 53:738–746.
- Davidson SA, Gilkey RH, Colburn HS, Carney LH (2006) Binaural detection with narrowband and wideband reproducible noise maskers. III. Monaural and diotic detection and model results. *J Acoust Soc Am* 119:2258–2275.
- Davidson SA, Gilkey RH, Colburn HS, Carney LH (2009) An evaluation of models for diotic and dichotic detection in reproducible noises. *J Acoust Soc Am* 126:1906–1925.
- Day ML, Delgutte B (2013) Decoding sound source location and separation using neural population activity patterns. *J Neurosci* 33:15837–15847.
- Day ML, Delgutte B (2015) Neural population encoding and decoding of sound source location across sound level in the rabbit inferior colliculus. *J Neurophysiol* 115:193–207.
- Dent ML, Dooling RJ, Pierce AS (2000) Frequency discrimination in budgerigars (*Melopsittacus undulatus*): effects of tone duration and tonal context. *J Acoust Soc Am* 107:2657–2664.
- Dooling RJ, Saunders JC (1975) Hearing in the parakeet (*Melopsittacus undulatus*): absolute thresholds, critical ratios, frequency difference limens, and vocalizations. *J Comp Physiol Psychol* 88:1–20.
- Dooling RJ, Searcy MH (1981) Amplitude modulation thresholds for the parakeet (*Melopsittacus undulatus*). *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 143:383–388.
- Egan JP (1975) Signal detection theory and ROC analysis. New York, New York: Academic Press.
- Evilsizer ME, Gilkey RH, Mason CR, Colburn HS, Carney LH (2002) Binaural detection with narrowband and wideband reproducible noise maskers: i. Results for human. *J Acoust Soc Am* 111:336–345.
- Fan L, Henry KS, Carney LH (2018) Challenging one model with many stimuli: simulating responses in the inferior colliculus. *Acta Acust united Acust* 104:895–899.
- Fletcher H (1940) Auditory patterns. *Rev Mod Phys* 12:47–65.
- Henry KS, Abrams KS (2021) Normal tone-in-noise sensitivity in trained budgerigars despite substantial auditory-nerve injury: no evidence of hidden hearing loss. *J Neurosci* 41:118–129.
- Henry KS, Neilans EG, Abrams KS, Idrobo F, Carney LH (2016) Neural correlates of behavioral amplitude modulation sensitivity in the budgerigar midbrain. *J Neurophysiol* 115:1905–1916.
- Henry KS, Abrams KS, Forst J, Mender MJ, Neilans EG, Idrobo F, Carney LH (2017a) Midbrain synchrony to envelope structure supports behavioral sensitivity to single-formant vowel-like sounds in noise. *J Assoc Res Otolaryngol* 18:165–181.
- Henry KS, Amburgey KN, Abrams KS, Idrobo F, Carney LH (2017b) Formant-frequency discrimination of synthesized vowels in budgerigars (*Melopsittacus undulatus*) and humans. *J Acoust Soc Am* 142:2073–2083.
- Henry KS, Amburgey KN, Abrams KS, Carney LH (2020) Identifying cues for tone-in-noise detection using decision variable correlation in the budgerigar (*Melopsittacus undulatus*). *J Acoust Soc Am* 147:984–997.
- Jazayeri M, Movshon JA (2006) Optimal representation of sensory information by neural populations. *Nat Neurosci* 9:690–696.
- Jiang D, McAlpine D, Palmer AR (1997) Responses of neurons in the inferior colliculus to binaural masking level difference stimuli measured by rate-versus-level functions. *J Neurophysiol* 77:3085–3106.
- Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. *Physiol Rev* 84:541–577.
- Keller CH, Takahashi TT (2000) Representation of temporal features of complex sounds by the discharge patterns of neurons in the owl's inferior colliculus. *J Neurophysiol* 84:2638–2650.
- Kidd G, Mason CR, Brantley MA, Owen GA (1989) Roving-level tone-in-noise detection. *J Acoust Soc Am* 86:1310–1317.
- Kim DO, Zahorik P, Carney LH, Bishop BB, Kuwada S (2015) Auditory distance coding in rabbit midbrain neurons and human perception: monaural amplitude modulation depth as a cue. *J Neurosci* 35:5360–5372.
- Kim DO, Carney LH, Kuwada S (2020) Amplitude modulation transfer functions reveal opposing populations within both the inferior colliculus and medial geniculate body. *J Neurophysiol* 124:1198–1215.
- Kohlrausch A, Fassel R, Van Der Heijden M, Kortekaas R, Van De Par S, Oxenham AJ, Püschel D (1997) Detection of tones in low-noise noise: further evidence for the role of envelope fluctuations. *Acustica* 83:659–669.
- 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, Schreiner CE (1988) Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *J Neurophysiol* 60:1799–1822.
- Leong U-C, Schwarz DM, Henry KS, Carney LH (2020) Sensorineural hearing loss diminishes use of temporal envelope cues: evidence from roving-level tone-in-noise detection. *Ear Hear* 41:1009–1019.
- Levitt H (1971) Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 49:467–477.
- Macmillan NA, Creelman CD (1991) Detection theory: a user's guide. Cambridge, England: Cambridge UP.
- Manley GA, Gleich O, Leppelsack HJ, Oeckinghaus H (1985) Activity patterns of cochlear ganglion neurones in the starling. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 157:161–181.
- Mao J, Carney LH (2015) Tone-in-noise detection using envelope cues: comparison of signal-processing-based and physiological models. *J Assoc Res Otolaryngol* 16:121–133.
- Mao J, Vosoughi A, Carney LH (2013) Predictions of diotic tone-in-noise detection based on a nonlinear optimal combination of energy, envelope, and fine-structure cues. *J Acoust Soc Am* 134:396–406.
- Müller-Preuss P, Flachskamm C, Bieser A (1994) Neural coding of amplitude modulation within the auditory midbrain of squirrel monkeys. *Hear Res* 80:197–208.
- Nelson PC, Carney LH (2004) A phenomenological model of peripheral and central neural responses to amplitude-modulated tones. *J Acoust Soc Am* 116:2173–2186.
- 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.

- Patterson RD (1976) Auditory filter shapes derived with noise stimuli. *J Acoust Soc Am* 59:640–654.
- 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.
- Ramachandran R, Davis KA, May BJ (2000) Rate representation of tones in noise in the inferior colliculus of decerebrate cats. *J Assoc Res Otolaryngol* 1:144–160.
- 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.
- Richards VM (1992) The detectability of a tone added to narrow bands of equal-energy noise. *J Acoust Soc Am* 91:3424–3435.
- Rocchi F, Ramachandran R (2018) Neuronal adaptation to sound statistics in the inferior colliculus of behaving macaques does not reduce the effectiveness of the masking noise. *J Neurophysiol* 120:2819–2833.
- Sachs MB, Young ED, Lewis RH (1974) Discharge patterns of single fibers in the pigeon auditory nerve. *Brain Res* 70:431–447.
- Salvi R, Saunders SS, Powers NL, Boettcher FA (1992) Discharge patterns of cochlear ganglion neurons in the chicken. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 170:227–241.
- Saunders JC, Pallone RL (1980) Frequency Selectivity in the parakeet studied by isointensity masking contours. *J Exp Biol* 87:331–342.
- Wong SJ, Abrams KS, Amburgey KN, Wang Y, Henry KS (2019) Effects of selective auditory-nerve damage on the behavioral audiogram and temporal integration in the budgerigar. *Hear Res* 374:24–34.
- Woolley SMN, Casseday JH (2005) Processing of modulated sounds in the zebra finch auditory midbrain: responses to noise, frequency sweeps, and sinusoidal amplitude modulations. *J Neurophysiol* 94:1143–1157.
- Woolley SMN, Portfors CV (2013) Conserved mechanisms of vocalization coding in mammalian and songbird auditory midbrain. *Hear Res* 305:45–56.
- Zhang H, Kelly JB (2003) Glutamatergic and GABAergic regulation of neural responses in inferior colliculus to amplitude-modulated sounds. *J Neurophysiol* 90:477–490.