Identifying cues for tone-in-noise detection using decision variable correlation in the budgerigar (*Melopsittacus undulatus*)

Kenneth S. Henry,^{1,a)} Kassidy N. Amburgey,¹ Kristina S. Abrams,² and Laurel H. Carney³

¹Department of Otolaryngology, University of Rochester, Rochester, New York 14642, USA

²Department of Neuroscience, University of Rochester, Rochester, New York 14642, USA

³Department of Biomedical Engineering, University of Rochester, Rochester, New York 14642, USA

ABSTRACT:

Previous studies evaluated cues for masked tone detection using reproducible noise waveforms. Human results founded on this approach suggest that tone detection is based on combined energy and envelope (ENV) cues, but detection cues in nonhuman species are less clear. Decision variable correlation (DVC) was used to evaluate tone-in-noise detection cues in the budgerigar, an avian species with human-like behavioral sensitivity to many complex sounds. DVC quantifies a model's ability to predict trial-by-trial variance in behavioral responses. Budgerigars were behaviorally conditioned to detect 500-Hz tones in wideband (WB; 100–3000 Hz) and narrowband (NB; 452–552 Hz) noise. Behavioral responses were obtained using a single-interval, two-alternative discrimination task and two-down, one-up adaptive tracking procedures. Tone-detection thresholds in WB noise were higher than human thresholds, putatively due to broader peripheral frequency tuning, whereas NB thresholds were within ~1 dB of human results. Budgerigar average hit and false-alarm rates across noise waveforms were consistent, highly correlated across subjects, and correlated to human results. Trial-by-trial behavioral results in NB noise were best explained by a model combining energy and ENV cues. In contrast, WB results were better predicted by ENV-based or multiple-channel energy detector models. These results suggest that budgerigars and humans use similar cues for tone-in-noise detection. © 2020 Acoustical Society of America. <https://doi.org/10.1121/10.0000621>

(Received 6 August 2019; revised 2 January 2020; accepted 7 January 2020; published online 10 February 2020)

[Editor: Sarah Verhulst]

Pages: 984–997

I. INTRODUCTION

Perception of speech and other signals in competing noise is a common problem faced in everyday life. Previous human studies have investigated the cues that support perception in noise using a variety of approaches, including tones presented in reproducible noise waveforms (Davidson *et al.*, 2006; Evilsizer *et al.*, 2002; Gilkey *et al.*, 1985; Mao *et al.*, 2013). These studies used a single-interval behavioral task to examine tone detection across an ensemble of reproducible noise waveforms. Hits occurred when the tone was detected for trials containing a tone plus noise. False alarms (FAs) occurred when the subject misclassified a noise trial as tone plus noise. Hit and FA rates for an ensemble of reproducible noise waveforms yield a detection pattern.

Detection patterns across reproducible noise waveforms are repeatable within test subjects and highly correlated across young listeners with normal hearing, suggesting that different individuals use the same stimulus cues to detect tones in noise (Davidson *et al.*, 2006; Evilsizer *et al.*, 2002). The detection pattern for listeners with normal hearing is consistent with predictions of a model that combines primarily single-channel energy (overall level) and envelope (ENV) cues after critical-band filtering of the stimulus (Mao

et al., 2013; Mao and Carney, 2015; note that these same cues are implicated in studies of speech intelligibility; Bibberger and Ewert, 2017; Jørgensen *et al.*, 2013; Rhebergen *et al.*, 2006). Single-channel energy and ENV cues explain a similar proportion of variance in the normal-hearing detection pattern, even when the energy cue is made unreliable, and hence unrelated to the presence or absence of the tone, through use of a “roving-level” paradigm that randomly varies the stimulus level across trials (Mao *et al.*, 2015). Temporal fine structure (TFS) cues explain less variance in the detection pattern. Altered tone-in-noise detection patterns in individuals with sensorineural hearing loss suggest a decline in use of ENV and TFS cues to perform the task—and increased dependence on energy (Mao *et al.*, 2015)—that may be associated with well-known speech-perception difficulties experienced by this population (Mao *et al.*, 2015; see also Ewert *et al.*, 2018; Paraouty *et al.*, 2016). Finally, as an alternative to the single-channel models discussed above, a multiple-channel energy-detector model based on energy differences across frequency channels can also explain substantial variance in human tone-in-noise detection (Davidson *et al.*, 2006, 2009; Gilkey and Robinson, 1986).

Tone-in-noise detection in nonhuman animal models has most frequently been studied using random noise in the context of critical bands or critical ratios (Niemiec *et al.*, 1992; Nienhuys and Clark, 1979; Okanoya and Dooling,

^{a)}Electronic mail: kenneth_henry@urmc.rochester.edu

1987; Pickles, 1975; Watson, 1963; Yost and Shofner, 2009). One series of studies in rabbits employed reproducible noise waveforms to determine detection patterns (Early *et al.*, 2001; Gai *et al.*, 2007; Zheng *et al.*, 2002). The detection pattern in rabbits was uncorrelated with human results and model predictions based on ENV cues, suggesting that this species uses different stimulus cues from humans to detect low-frequency tones in noise. In chinchillas, behavioral results in trained animals suggest a wideband (WB) tone-in-noise detection strategy rather than processing within a single critical band, as has been proposed to occur in humans (Yost and Shofner, 2009). Given that results are from a small number of studies, it remains unknown whether other nonhuman species also use different cues from humans to perform tone-in-noise detection, particularly those with behavioral sensitivity to complex sounds that closely approximates human detection performance.

The budgerigar (*Melopsittacus undulatus*) is a small parrot species with human-like behavioral sensitivity to many simple and complex sounds and the capacity to mimic speech (Dooling *et al.*, 2000). Budgerigars have sensitive hearing from 0.25 to 6 kHz (Dooling and Saunders, 1975) and show behavioral performance similar to humans on a range of psychoacoustical tasks, including frequency discrimination of tones and vowel formants (Dent *et al.*, 2000; Henry *et al.*, 2017a; Henry *et al.*, 2017b), gap detection (Okanoya and Dooling, 1990), and amplitude-modulation detection (Carney *et al.*, 2013; Dooling and Searcy, 1981; Henry *et al.*, 2016). Budgerigar tone-detection thresholds in WB noise are within ± 1 –2 dB of human performance over the frequency range from 2 to 4 kHz, for which psychophysical tuning curves and critical bands are similar between the two species (Farabaugh *et al.*, 1998; Okanoya and Dooling, 1987; Saunders *et al.*, 1979; Saunders and Pallone, 1980). It remains unknown which cues budgerigars use to detect tones in noise.

The present study examined behavioral tone-in-noise detection patterns in the budgerigar using the same 500-Hz tone frequency and reproducible WB and narrowband (NB) noise waveforms from previous human and rabbit studies (Davidson *et al.*, 2006; Evilsizer *et al.*, 2002; Gai *et al.*, 2007). Cues for tone-in-noise detection were evaluated using decision variable correlation (DVC), a recent extension of signal detection theory that quantifies the ability of a model to predict behavioral responses of a test subject on a trial-by-trial basis (Sebastian *et al.*, 2017; Sebastian and Geisler, 2018). DVC has not yet been used to evaluate cues for tone-in-noise detection in any species. DVCs were used to assess the ability of single-channel energy, ENV, and TFS models and a multiple-channel energy-detector model (Gilkey and Robinson, 1986) to explain tone-in-noise detection in trained budgerigars.

II. METHODS

A. Subjects

Behavioral experiments were performed in four budgerigars (*M. undulatus*; two male, two female). Animals ranged from 3 to 4 years of age and weighed between 35 and 50 g.

All experimental procedures were approved by the University Committee on Animal Resources at the University of Rochester.

B. Experimental apparatus

The experimental apparatus has been described previously (Henry *et al.*, 2017b). Briefly, experiments were conducted in four, single-walled acoustic isolation chambers lined with sound-absorbing foam. Test chambers had an internal volume of 0.3 m³ and contained an overhead light and observation camera. Animals perched in a wire mesh cage with access to three horizontally arranged response switches. Individual hulled millet seeds were dispensed by a feeder (ENV-203 Mini; Med Associates, St. Albans, VT) and delivered to a trough inside the cage for reinforcement of behavioral responses. The feeder, lights, and response switches were controlled by a data acquisition board (PCI-6151; National Instruments Corporation, Austin, TX), microcontroller (Arduino Leonardo; Turin, Italy), custom-made hardware, and custom MATLAB programs (The MathWorks, Natick, MA). Stimuli were presented by a loudspeaker (Polk Audio MC60; Baltimore, MD) positioned 20 cm above the animal's head. Calibration was performed with a digital pre-emphasis filter calculated based on the output of a 1/2 in. precision microphone (Brüel and Kjær type 4134, Marlborough, MA) in response to 249 log-spaced frequencies from 0.05 to 15.1 kHz. The pre-emphasis filter corrected the frequency response of the loudspeaker system prior to stimulus presentation.

C. Behavioral test procedures

Behavioral experiments were performed using operant-conditioning procedures and a single-interval, two-alternative, non-forced choice paradigm. A trial was initiated by pecking the center switch, resulting in the presentation of a single stimulus. Following stimulus presentation, birds had 3 s to respond by pressing either the left switch, which was the correct response to a tone-plus-noise trial (a hit), or the right switch, which was the correct response to noise alone (a correct rejection). Correct responses (hits and correct rejections) were reinforced with one or two millet seeds, depending on bias toward the right or left switch (see below). Incorrect responses (misses and FAs) resulted in no seed reward, and the house light was turned off for a 5-s timeout. If a bird pecked any switch during the timeout period, the timeout was restarted for 5 s to discourage this behavior. In rare cases when the bird did not report a response following the stimulus ($<0.22\%$ of trials), a shorter 2-s timeout was given. Five tone-plus-noise and five noise-alone stimuli were presented in a random sequence for every group of ten trials. All of the 25 frozen-noise stimuli were presented in a random sequence over every group of 25 trials.

Animals were initially trained to discriminate between noise-alone and tone-plus-noise stimuli presented at a high signal-to-noise ratio [SNR; i.e., E_s/N_0 of 26.8 dB; E_s/N_0 is the tone level [dB sound pressure level (SPL)] minus the

noise spectrum level (dB SPL/Hz) plus 10 times log ten of the stimulus duration (0.3 s for the current experiments)]. Once birds reached 90% correct discrimination, two-down one-up adaptive tracking procedures (Levitt, 1971) were used to test behavioral responses near the animal's threshold ($\sim 70.7\%$ correct responses). For every two successive, correct responses to tone-plus-noise trials, the SNR was decreased by 1 dB. Incorrect responses to a tone-plus-noise trial resulted in a 1-dB increase in SNR. Responses to noise-alone trials did not affect the track. Changes in track direction from up to down or from down to up were considered a reversal. Tracking sessions continued for 300 trials or until 1.5–2 g of millet were consumed. Behavioral sessions were run 7 days per week in morning and afternoon test blocks lasting approximately 30 min each.

Bias was calculated during all sessions (tracking and non-tracking) as -0.5 times the sum of the Z-score of the hit rate and Z-score of the false-alarm rate (Macmillan and Creelman, 2005). Bias was calculated after the first 50 trials and recalculated every 20 trials thereafter based on the most recent 50 trials. The proportion of correct responses in which a bird received two seeds (for responses on the biased-against side) was adaptively varied following each calculation to control the bias. Any tracking session with cumulative absolute bias (calculated based on all trials within the session) of 0.3 or greater was not used in later analysis.

D. Stimulus conditions

Stimuli were WB (100–3000 Hz) and NB (452–552 Hz) noise waveforms from previous studies ($n = 25$ for each bandwidth), presented with or without a 500-Hz, simultaneously gated tone (Davidson *et al.*, 2006; Evilsizer *et al.*, 2002; Mao *et al.*, 2015). The WB noise waveforms were Gaussian noises that were generated with different seeds, and then band limited by transforming them into the frequency domain and setting all magnitudes outside the passband to zero before transforming back into the time domain. The NB noises were generated from the WB noises by further limiting the passband, using the same frequency-domain strategy. Thus, the spectrum of each NB noise matched that of the corresponding WB noise within the band surrounding 500 Hz.

Stimuli were 300 ms in duration with 10-ms cosine-squared onset and offset ramps. Four conditions were tested for each noise bandwidth: (1) fixed noise spectrum level of 40 dB SPL/Hz, (2) fixed noise spectrum level of 60 dB SPL/Hz, (3) roving noise spectrum level from 30 to 50 dB SPL/Hz, and (4) roving noise spectrum level from 50 to 70 dB SPL/Hz. For the two roving-level conditions, the noise spectrum level on each trial was drawn randomly from a uniform distribution with a 20-dB range. The roving-level paradigm was used because it makes the single-channel energy cue unreliable for detecting the tone and reduces the correlation between single-channel energy and ENV cues for tone-in-noise detection. Stimuli were generated in MATLAB with a sampling frequency of 50 kHz.

E. Tone-in-noise detection patterns

The eight stimulus conditions were tested in different orders across animals to yield detection patterns. Sessions were conducted repeatedly on the same condition for several weeks until thresholds and the detection pattern were stable. Thresholds were estimated for each test session as the mean SNR (E_S/N_0) of reversal points during the last half of the track. Detection patterns plotting hit and FA rate across different frozen noise waveforms were calculated based on near-threshold behavioral responses obtained when the SNR of the most recent tone-plus-noise stimulus was within ± 2 dB of the mean threshold. Note that detection patterns were based on behavioral responses pooled across many test sessions conducted over several weeks. The total duration of behavioral testing required to complete all eight stimulus conditions ranged from 24 to 34 weeks across animals (~ 3 –4 weeks of testing per condition). For each stimulus condition and in each animal, the stability of the detection pattern over the testing period was assessed using Chi-squared analyses of data pooled across sessions and by calculating the Pearson correlation between results pooled from the first and second halves of behavioral test results. The average number of trials per noise waveform for each animal and stimulus condition was 123, including both tone-plus-noise and noise-alone trials.

F. DVC analysis

DVCs were used to evaluate the ability of several models to predict trial-by-trial variance in behavioral responses to tone-plus-noise and noise-alone stimuli. With this analysis framework described previously in detail (Sebastian *et al.*, 2017; Sebastian and Geisler, 2018), responses are compared between a model and observer tested with the same near-threshold stimulus sequences (including identical noise waveforms). Importantly, performance varies from trial to trial due to the random nature of the specific noise waveform generated. DVCs quantify, separately for each stimulus category (noise alone vs tone plus noise), the extent to which responses of the model and the observer are correlated. High DVCs provide evidence that the decision variable used to calculate the model response also influences the test subject's behavioral responses. Procedures from Sebastian and Geisler (2018) were followed for maximum-likelihood estimation of DVCs.

Decision variables were calculated for single-channel energy, ENV and TFS models, and a multiple-channel energy-detector model. The first stage of the single-channel models was a fourth-order gamma-tone filter with a center frequency of 500 Hz and equivalent rectangular bandwidth of 225 Hz (bandwidth was 400 Hz at a gain 10-dB below the filter's peak). This bandwidth value falls within the range of psychophysical tuning-curve bandwidths reported previously in this species (Saunders *et al.*, 1979; Saunders and Pallone, 1980) and was found in preliminary analyses to predict slightly more variance in trial-by-trial performance than higher or lower values, and thus used as the critical

bandwidth. The decision variable of the energy model (Fletcher, 1940) was calculated as the energy level (dB SPL) of the stimulus after passing through the critical-band filter. The ENV decision variable (Richards, 1992) was calculated by first normalizing the critical-band filtered stimulus by its root mean square (RMS) value and computing the Hilbert ENV. Next, the time-varying ENV slope was computed (using the “diff” function in MATLAB) and full-wave rectified, yielding only positive slope values. The ENV decision variable was calculated as negative one times the mean value of the rectified ENV slope function (Davidson *et al.*, 2009). The TFS decision variable was calculated by first half-wave rectifying the critical-band filtered stimulus and low-pass filtering the result to capture the roll-off in auditory-nerve phase locking with increasing frequency (Gleich and Klump, 1995; Sachs and Abbas, 1974). The low-pass filter was constructed by cascading seven first-order filters, each with a 3-dB cutoff frequency of 4000 Hz, to yield a roll-off in phase locking of 3 dB at 1325 Hz and ~ 100 dB/decade from 4 to 6 kHz (Heinz *et al.*, 2001). TFS coding was then quantified in the conventional way as vector strength (Goldberg and Brown, 1969) by dividing the amplitude of the 500-Hz component in Fourier transform of the response by the response mean. Finally, the multiple-channel energy-detector model (Gilkey and Robinson, 1986) was implemented as a bank of fourth-order gamma-tone filters each with an equivalent rectangular bandwidth of 75 Hz. Filter center frequencies ranged from 425 to 575 Hz for the NB condition and from 275 to 725 Hz for the WB condition with a step size of 75 Hz. The decision variable of the multiple-channel energy-detector model was the weighted linear sum of the RMS energy level at the output of the filters. Filter weights were determined separately for each noise bandwidth and in each subject using linear logistic regression analyses (function “fitglm” in MATLAB). The dependent variable was an array of zeros (tone reported absent) and ones (tone reported present), indicating subject responses to all stimulus trials (including both tone-plus-noise and noise-alone). Independent variables were the RMS energy level at the output of each filter (NB, $N = 3$ filters; WB, $N = 7$ filters). Regression coefficients produced by the analysis were used as filter weights for calculation of the multiple-channel energy detector decision variable.

DVCs were estimated based on the same near-threshold behavioral results used to calculate roving-level detection patterns (i.e., based on the several thousand trials for which the most recent tone-plus-noise stimulus was within ± 2 dB of the mean threshold). Because animals frequently responded before the end of the 0.3-s stimulus interval, stimulus waveforms were reduced in duration to the reaction time of the subject on a trial-by-trial basis prior to calculating the decision variables. Only behavioral results from the four roving-level stimulus conditions were analyzed because this paradigm disrupts collinearity of energy with ENV and TFS decision variables and therefore limits the potential for spurious DVC results (e.g., if the energy-model decision variable is correlated with behavioral responses for fixed-

level test conditions, it is unclear whether this reflects actual use of the energy cue by test subjects or use of an alternative temporal cue to which the energy cue is correlated). DVCs were estimated for the models outlined above and for weighted linear combinations of the three single-channel model decision variables. The individual decision variables were first normalized to have equal variance across trials by dividing each one by the standard deviation of the decision variable observed across all trials. Linear combination was then accomplished by adding individual decision variables together with weights (i.e., scaled by a factor) ranging from zero to one. The combined weight of the three cues was always one. Model implementation, linear combination of decision variables, and estimation of DVCs were accomplished in MATLAB.

G. Statistical analysis

Statistical analyses included Chi-square analyses, t -tests, and linear regression analyses conducted in MATLAB, and linear mixed-effects ANOVAs (Bates *et al.*, 2015) conducted in R (version 3.4.1). Subject random intercepts were used in mixed-effects analyses to account for repeated observations within test subjects. For threshold analyses, the fixed effects were noise level (40 vs 60 dB SPL/Hz spectrum level) and listening paradigm (fixed vs roving level). For DVC analyses, fixed effects were the model decision variable (e.g., energy vs ENV cue) and trial type (noise-alone vs tone-plus-noise). Interactions were included and dropped when not significant ($p > 0.05$) in order of decreasing p value. Degrees of freedom for F tests and pairwise comparisons of least-squares means were calculated based on the Satterthwaite approximation. Visual inspection of model results showed that residuals were approximately normally distributed. The distribution of the residuals was not significantly different from a normal distribution for all models ($W = 0.97014$ – 0.98416 , $p = 0.58$ – 0.84 ; Shapiro-Wilk normality tests).

III. RESULTS

A. Detection patterns across reproducible noise waveforms

Budgerigar detection patterns show hit rates and FA rates across an ensemble of 25 reproducible NB and WB noise waveforms with spectrum level of 40 dB SPL/Hz (Fig. 1). Patterns in each animal are based on an average of 5240 behavioral trials (range 3370–7725) pooled across several weeks of testing. Detection patterns were highly consistent over this test period in all animals, as indicated by high within-subject Pearson correlations comparing mean results between the first and second halves of behavioral testing (NB hit rate, $\rho = 0.88 \pm 0.04$; NB FA rate, $\rho = 0.84 \pm 0.08$; WB hit rate, $\rho = 0.89 \pm 0.03$; WB FA rate, $\rho = 0.87 \pm 0.07$; means $\pm$ SD; $p < 0.0001$ for all correlations). Similarity of behavioral performance across different subjects was evaluated by comparing individual detection patterns between all possible subject pairs using Pearson correlations. Detection

FIG. 1. (Color online) Budgerigar 500-Hz tone-in-noise detection patterns showing variation in behavioral response probability across an ensemble of 25 reproducible WB (100–3000 Hz) and NB (452–552 Hz) noise waveforms. (a) Hit rate for NB noise waveforms; (b) hit rate for WB noise waveforms; (c) false-alarm rate for NB noise waveforms; (d) FA rate for WB noise waveforms. Detection patterns are similar between the four different test subjects (see legend). Note that the symbols for each bird are connected by lines for the purpose of illustration. Standard error (SE) bars drawn in the right of each panel [$p \pm \sqrt{p(1-p)/N}$, where p is the response probability, and N is the number of trials per noise waveform] illustrate the level of within-subject variability. The standard error is shown for p of 0.75 for hit rate (a),(b) and for p of 0.25 for FA rate (c),(d). Note that for equal N as in the present experiments, standard error is greatest for p of 0.5 (15% greater than for p of 0.25 or 0.75) and decreases for p values approaching zero or 1.

patterns were significantly correlated between all subject pairs for both bandwidth conditions (NB hit rate, $\rho = 0.82 \pm 0.07$; NB FA rate, $\rho = 0.74 \pm 0.10$; WB hit rate, $\rho = 0.74 \pm 0.09$; WB FA rate, $\rho = 0.65 \pm 0.13$; means $\pm$ SD; $p \leq 0.004$ for all comparisons) and approached the strength of within-subject correlations described above. These results suggest that different test subjects used the same cues to detect a 500-Hz tone in noise.

Tone detection in 40 dB SPL/Hz noise was compared between budgerigars, humans ($n = 8$; from Davidson *et al.*, 2006, and Evilsizer *et al.*, 2002), and rabbits ($n = 4$; from Gai *et al.*, 2007) through pairwise comparisons of average detection patterns between species. Average detection patterns for the NB condition were significantly correlated between budgerigars and humans (hit rate, $\rho_{23} = 0.44$, $p = 0.028$; FA rate, $\rho_{23} = 0.74$, $p < 0.0001$). Note that rabbits had unstable detection patterns for NB stimuli (Gai *et al.*, 2007). Average detection patterns for the WB condition were significantly correlated between budgerigars and humans (hit rate, $\rho_{23} = 0.59$, $p = 0.002$; FA rate, $\rho_{23} = 0.51$, $p = 0.009$) but not between budgerigars and rabbits (hit rate, $\rho_{23} = 0.29$, $p = 0.16$; FA rate, $\rho_{23} = 0.27$, $p = 0.19$) or between rabbits and humans (as was reported previously by Gai *et al.*, 2007; hit rate, $\rho_{23} = 0.23$, $p = 0.26$; FA rate, $\rho_{23} = -0.08$, $p = 0.69$). These results suggest that budgerigars and humans use similar cues to detect a 500-Hz tone in noise. In contrast, rabbits may use different cues to perform the task.

B. Effects of noise bandwidth on tone-detection thresholds

Budgerigar tone-detection thresholds were lower (more sensitive) in NB noise than in WB noise (Table I; $t_3 = -23.65$, $p = 0.0002$; paired t -test of the fixed-level condition with spectrum level of 40 dB SPL/Hz). The threshold in WB noise was within ± 1 dB of previously reported behavioral thresholds in this species [Dooling and Saunders, 1975; Okanoya and Dooling, 1987; Saunders *et al.*, 1979; previously reported critical ratios (difference in dB between the threshold signal level and noise spectrum level) were converted to E_s/N_0 by adding $10 \log_{10}(0.3 \text{ s})$]. Compared to human thresholds, which do not vary between the two noise bandwidths for this task (Davidson *et al.*, 2006; Evilsizer

TABLE I. Budgerigar tone-in-noise detection thresholds (E_s/N_0).

Bandwidth	Spectrum level (Hz ⁻¹)	Rove	B47	B34	B35	B54	Mean $\pm$ SD
WB	40 dB SPL	0 dB	18.1	17.7	18.2	17.0	17.8 $\pm$ 0.5
		20 dB	18.8	16.9	17.3	16.0	17.3 $\pm$ 1.2
	60 dB SPL	0 dB	18.0	14.8	16.3	14.8	16.0 $\pm$ 1.5
		20 dB	17.1	16.1	16.4	16.0	16.4 $\pm$ 0.5
NB	40 dB SPL	0 dB	13.5	12.4	13.8	12.4	13.0 $\pm$ 0.7
		20 dB	16.4	16.9	14.6	14.6	15.6 $\pm$ 1.2
	60 dB SPL	0 dB	12.2	9.8	11.6	13.0	11.7 $\pm$ 1.4
		20 dB	13.1	12.3	13.5	13.8	13.2 $\pm$ 0.7

et al., 2002), budgerigar thresholds were nearly as sensitive as those of humans for the NB condition [mean difference $\pm$ standard error (SE), 1.2 ± 0.5 dB; $t_{10} = 2.23$, $p = 0.050$; two-sample t -test] but were significantly higher for the WB condition (mean difference $\pm$ SE, 6.7 ± 0.5 dB; $t_{10} = 9.79$, $p < 0.0001$). The observed threshold elevation for the WB condition shows that in budgerigars, noise frequency components outside the bandwidth of the NB stimulus (452–552 Hz) impacted 500-Hz tone detection in WB noise more than in humans. This finding likely reflects broader peripheral tuning bandwidth at 500 Hz, observed previously in this species, as compared to humans (Saunders *et al.*, 1978; Saunders *et al.*, 1979; Saunders and Pallone, 1980).

C. Effects of roving stimulus level on tone-in-noise detection patterns and thresholds

Tone-in-noise detection in the budgerigar was also assessed under roving-level conditions, for which the stimulus level was randomly varied over ± 10 dB across trials, making single-channel energy cues unreliable (Fig. 2). The average spectrum level of the noise was either 40 dB SPL/Hz (as described above in Secs. III A and III B) or 60 dB SPL/Hz. For each noise bandwidth, average detection patterns were highly correlated between fixed- and roving-level conditions and between the two average spectrum levels (Table II). By

TABLE II. Pearson correlations of average budgerigar detection patterns across stimulus conditions. Roving-level range was 20 dB.

	40 dB SPL/Hz; roving	60 dB SPL/Hz; fixed	60 dB SPL/Hz; roving
WB hit rate			
40 dB SPL/Hz; fixed	0.94	0.82	0.87
40 dB SPL/Hz; roving		0.90	0.95
60 dB SPL/Hz; fixed			0.95
WB FA rate			
40 dB SPL/Hz; fixed	0.95	0.78	0.83
40 dB SPL/Hz; roving		0.83	0.87
60 dB SPL/Hz; fixed			0.90
NB hit rate			
40 dB SPL/Hz; fixed	0.94	0.94	0.89
40 dB SPL/Hz; roving		0.88	0.93
60 dB SPL/Hz; fixed			0.91
NB FA rate			
40 dB SPL/Hz; fixed	0.86	0.84	0.75
40 dB SPL/Hz; roving		0.91	0.88
60 dB SPL/Hz; fixed			0.93

FIG. 2. (Color online) Predicted effect of roving stimulus level based on tone-in-noise detection by the single-channel energy model. (a) The decision variable range of the energy model as a function of SNR for tones in fixed-level noise (blue). Note separation at higher SNRs from the decision variable range of noise-alone trials (red). (b) Decision variable ranges of tone-plus-noise and noise-alone trials, as in (a), but for the roving-level condition. Stimulus level varied ± 10 dB (random uniform distribution) from trial to trial. (c) Tone-in-noise detection performance of the energy model as a function of stimulus SNR for fixed- and roving-level conditions. The threshold for 70.7% correct model performance increases ~ 11 dB for the roving-level condition.

contrast, note that the expected correlation of detection patterns between a single-channel energy model for fixed-level tone-in-noise detection and an ENV model for roving-level detection ranges from 0.5 to 0.7 across test conditions (Mao *et al.*, 2013). The observed, higher cross-condition correlations in budgerigars are consistent with no change in the cues for tone-in-noise detection across fixed- and roving-level conditions or across the different sound levels.

Tone-in-noise detection thresholds of the single-channel energy model increased by ~ 11 dB between fixed- and roving-level conditions based on simulations (Fig. 2). Model thresholds were estimated by first computing the decision variable distribution in response tones in NB noise as a function of SNR [500 trials per SNR; Figs. 2(a) and 2(b); WB results were similar]. The percent discrimination performance of the model between tone-plus-noise vs noise-alone stimuli was calculated at each SNR based on the area under the receiver-operator characteristic curve with threshold defined as the SNR above which discrimination performance exceeded 70.7% [Fig. 2(c); 70.7% is the expected mean correct performance from a two-down, one-up adaptive staircase procedure; Levitt, 1971]. In contrast to expectations based on the single-channel energy model, detection thresholds of budgerigars increased by only ~ 2 dB for the roving-level compared to fixed-level thresholds in NB noise, and were unchanged with roving stimulus level in WB noise (Table I). These results suggest that budgerigars can use cues in addition to (or different from) stimulus energy to detect a 500-Hz tone in noise. A mixed-model analysis of NB thresholds showed main effects of both average spectrum level ($F_{1,13} = 13.51$, $p = 0.0028$) and roving level ($F_{1,13} = 10.70$, $p = 0.0061$). The interaction between average spectrum level and roving status was not significant ($F_{1,12} = 0.35$, $p = 0.57$). NB thresholds increased by 1.9 ± 0.6 dB under roving-level

conditions, consistent with limited influence of stimulus energy, and were slightly lower [-2.1 ± 0.6 dB; least-squares (LS) mean difference $\pm$ SE] for the 60 dB SPL/Hz spectrum level. The same analysis of WB thresholds showed a main effect of spectrum level (40 vs 60 dB SPL/Hz; $F_{1,10} = 18.67$, $p = 0.0015$) but no difference between roving- and fixed-level conditions ($F_{1,10} = 0.11$, $p = 0.74$) and no interaction between average spectrum level and roving status ($F_{1,9} = 3.14$, $p = 0.11$). WB thresholds were again slightly lower for the 60 dB SPL/Hz spectrum level (-1.4 ± 0.3 dB; LS mean difference $\pm$ SE).

The findings that detection patterns and thresholds are relatively unaffected by the roving-level paradigm are noteworthy because they suggest no major change in the cues for tone-in-noise detection between fixed- and roving-level conditions. Thus, budgerigars do not appear to transition from use of single-channel energy cues under fixed-level test conditions to rove-resistant cues (e.g., ENV- or TFS-based) when the energy cue becomes unreliable (i.e., with roving stimulus level).

D. Single-channel cues for tone detection in NB noise

DVCs (Sebastian *et al.*, 2017; Sebastian and Geisler, 2018) were used to evaluate the ability of several models to predict behavioral tone detection in noise on a trial-by-trial basis. Only responses from roving-level trials were analyzed because this stimulus paradigm minimizes collinearity between energy and temporal cues (i.e., TFS and ENV; see Table III; again, note that the correlation between energy and ENV decision variables ranges from 0.5 to 0.7 for fixed-level tone-in-noise detection; Mao *et al.*, 2013). Results were similar between 40 and 60-dB SPL/Hz noise spectrum levels and therefore combined, yielding 5638–8014 trials per animal (25 847 trials total). The decision variable of the energy model (Fletcher, 1940) was calculated as the log transformed RMS output of a single critical-band filter centered on the tone

frequency. For NB noise, budgerigar behavioral responses were correlated to the energy-model decision variable for both tone-plus-noise and noise-alone trials [Figs. 3(a), top; Figs. 4(a) and 4(c)], indicating that tone-in-noise detection was influenced by single-channel energy cues. However, note that due to the roving-level paradigm, the sensitivity of the energy model [$d' = 0.437 \pm 0.037$; mean $\pm$ SD across animals; Fig. 3(a), bottom; Fig. 4(g)] was too low to explain the behavioral sensitivity of test subjects [$d' = 1.237 \pm 0.118$; mean $\pm$ SD; Fig. 4(g)]. Consequently, budgerigars must have used other cues, in addition to energy, to detect low-frequency tones in NB noise.

ENV fluctuations and TFS at the output of the critical-band filter provide alternative, temporal cues for tone-in-noise detection that are substantially more informative than energy cues under roving-level conditions [e.g., ENV model $d' = 2.230 \pm 0.189$; TFS $d' = 4.44 \pm 0.580$; mean $\pm$ SD; Figs. 3(b) and 3(c), bottom; Fig. 4(g)]. Decision variables for ENV and TFS models were calculated based on the mean absolute slope of ENV fluctuations normalized by energy and vector strength at the output of a simple auditory-nerve model, respectively (see Sec. II). Behavioral responses were moderately correlated to decision variables of both the ENV model and TFS model for tone-plus-noise trials [Figs. 3(b) and 3(c), top, blue; Fig. 4(a)]. Which temporal cue, in fact, influenced behavior was unclear, however, due to substantial collinearity of ENV and TFS cues across stimuli (Table III). Behavioral responses to noise-alone trials for which TFS and ENV cues are more independent (Table III) were moderately correlated to the decision variable of the ENV model but not to the TFS model [Figs. 3(b) and 3(c), top, red; Fig. 4(c)]. This result suggests that tone-in-noise detection in budgerigars may be more influenced by ENV fluctuations than by TFS.

A mixed-model analysis of variance (ANOVA) confirmed significant variation in DVC across single-cue models and trial types [model cue, $F_{2,18} = 61.38$, $p < 0.0001$; trial type, $F_{1,18} = 3.87$, $p = 0.065$; model cue $\times$ trial type, $F_{2,18} = 5.115$, $p = 0.017$; Figs. 4(a) and 4(c)]. DVCs were higher for the energy model than the ENV model ($t_{18} = 7.03$, $p < 0.0001$) or TFS model ($t_{18} = 10.96$, $p < 0.0001$) and higher for the ENV model than the TFS model ($t_{18} = 3.93$, $p = 0.001$; pairwise comparisons of LS means). Between trial types, DVC was higher for tone-plus-noise trials than noise-alone trials for the TFS model ($t_{18} = 3.11$, $p = 0.006$) but not for the other models (energy, $t_{18} = 1.62$, $p = 0.12$; ENV; $t_{18} = 1.33$, $p = 0.2$).

The observed behavioral correlations to energy and temporal cues considered in isolation raise the question of whether a single-channel model combining multiple cues might better explain trial-by-trial variation in behavioral performance. This possibility was addressed through linear combination of weighted energy, ENV, and TFS cues into a single decision variable. The best-fitting (i.e., most highly correlated to behavior) combined model showed zero weighting of TFS and energy cues that were weighted more highly than ENV cues in all test subjects [Figs. 3(d), 4(e), 5,

TABLE III. Pearson correlations between model decision variables. Model abbreviations are as in Figs. 3 and 4.

	ENV	TFS	MD
WB tone-plus-noise trials			
Energy	0.1323	0.0813	0.1391
ENV		0.5686	0.8981
TFS			0.5304
WB noise-alone trials			
Energy	0.1030	−0.0635	0.1047
ENV		−0.1581	0.8781
TFS			−0.1830
NB tone-plus-noise trials			
Energy	0.0651	0.1224	0.1261
ENV		0.6412	0.8111
TFS			0.7864
NB noise-alone trials			
Energy	0.0543	−0.0301	0.0241
ENV		0.1232	0.7141
TFS			0.2250

FIG. 3. (Color online) Comparison of budgerigar behavioral responses in roving-level NB noise to decision variables of single-channel energy, ENV, TFS, and combined models (a)–(d), and a multiple-channel model (e). Top panels show the proportion of trials for which subjects responded that a tone was present as a function of the model decision-variable value. Results are plotted separately for tone-plus-noise (blue) and noise-alone trials (red). DVCs indicating the correlation of trial-by-trial behavioral responses to the model decision variable are reported in the lower right of each panel. Results are based on 25 847 trials pooled across four test subjects, for which the SNR of the most recent tone-plus-noise stimulus was within ± 2 dB of the average reversal-based threshold. Linear weights of cues in the combined model are indicated at the top of (d). The unit of the decision variable is dB for energy, vector strength for TFS, and arbitrary for ENV, combined models, and the multiple-channel model. Lower panels compare normalized distributions of the model decision variable between tone-plus-noise and noise-alone trials. Model tone-in-noise sensitivity (d') based on the distributions is reported at the top of each panel. The d' value of the test subjects was 1.21.

left]. While the increase in DVC of this model over the single-cue energy model was small (0.05 ± 0.016 ; LS mean difference $\pm$ SE; $t_{10} = 3.54$, $p = 0.005$; pairwise comparison of LS means from a mixed-model ANOVA comparing energy and combined models), note that in contrast to the energy model, the sensitivity of the energy-plus-ENV model was sufficient [$d' = 1.298 \pm 0.224$; Fig. 3(d), bottom] to explain the behavioral sensitivity of test subjects [Fig. 4(g)]. These results show that budgerigars more likely use a combination of energy and ENV cues rather than energy alone to detect tones in NB noise (a multiple-channel detection strategy is considered below).

E. Single-channel cues for tone detection in WB noise

DVC analyses of budgerigar behavioral results in WB noise were based on 4880–8120 roving-level trials per animal (23 501 total trials). WB results were similar to those described above in NB noise, except that behavioral response correlations to the single-channel energy model were lower than those described above for NB noise [Figs. 6(a), 4(b), and 4(d)]. Thus, budgerigars apparently were less reliant on single-channel energy cues for tone detection in WB noise as compared to NB noise. As was the case for NB noise, behavioral responses were once again correlated to the ENV model decision variable for both trial types and correlated to the TFS model decision variable primarily for tone-plus-noise trials [Figs. 6(b), 6(c), 4(b), and 4(d); note collinearity of TFS and ENV decision variables for WB tone-plus-noise trials in Table III]. A mixed-model

ANOVA showed significant variation in DVC across single-cue models and trial types [model cue, $F_{2,15} = 18.78$, $p < 0.0001$; trial type, $F_{1,15} = 3.25$, $p = 0.091$; model cue $\times$ trial type, $F_{2,15} = 3.92$, $p = 0.043$; Figs. 4(b) and 4(c)]. DVC was higher for the ENV model than the energy model ($t_{15} = 5.09$, $p = 0.0001$) or TFS model ($t_{15} = 5.50$, $p = 0.0001$) and similar between energy and TFS models ($t_{15} = 0.41$, $p = 0.69$). Between trial types, DVC was higher for tone-plus-noise trials than for noise-alone trials for the TFS model ($t_{15} = 3.07$, $p = 0.008$) but not for the other single-cue models (energy, $t_{15} = 0.88$, $p = 0.39$; ENV; $t_{15} = 0.93$, $p = 0.37$).

Cue weighting of the best-fitting combined-cue, single-channel model in WB noise [Figs. 6(d) and 4(f)] was zero for TFS in most test subjects, similar to the NB condition, suggesting minimal use of this cue, but weighting was generally greater for ENV than for energy [Figs. 4(f) and 5, right]. This pattern contrasts with the finding in NB noise that energy cues carried greater weight than ENV cues (although note that a combination of energy with ENV cues was required to explain the observed sensitivity of test subjects on the NB condition), and suggests that budgerigars are less influenced by single-channel energy cues when the masking noise is broadband. A mixed-model ANOVA of DVCs from the ENV model and best-fitting combined model showed no significant variation between model types ($F_{1,10} = 0.76$, $p = 0.40$). All models in WB noise showed sensitivity (d') sufficient to account for the observed behavioral sensitivity of the test subjects except for the

FIG. 4. (Color online) (a)–(d) Comparison of DVCs across energy (E), ENV, TFS, combined (c), and multiple-channel energy-detector (MD) models. DVCs are shown for (a),(b) tone-plus-noise trials and (c),(d) noise-alone trials. (e),(f) Linear weighting of energy, ENV, and TFS cues in the best-fitting combined model. (g),(h) Sensitivity (d') of the test subjects (OBS) and the various models. Results are shown for (a),(c),(e),(g) NB and (b),(d),(f),(h) WB test conditions. In all panels, horizontal and vertical bars indicate cross-subject means and standard deviations, respectively.

single-channel energy model [Fig. 4(h)]. In summary, budgerigars appear to use primarily single-channel ENV cues, possibly combined with energy, to detect 500-Hz tones in WB noise.

F. Tone-in-noise detection based on the multiple-channel energy detector

Differences in energy across frequency channels provide an alternative cue for detecting tones in noise compared to the single-channel decision variables considered above. The decision variable of a multiple-channel energy detector

FIG. 5. (Color online) Mean DVCs of the combined model for different linear combinations of energy, ENV, and TFS decision variables. Cue weight decreases from one at the labeled vertex to zero along the adjacent edge. “x” symbols indicate the decision variable combination yielding the highest average DVC across both trial types (indicated top). White shading indicates a DVC of zero and the DVC increment between contours is 0.05. The best-fitting model combines energy and ENV cues for the NB test condition and is based primarily on ENV in WB noise.

model was calculated as the linear combination of RMS energy at the output of multiple filters with center frequencies near the 500-Hz tone frequency (NB, 425–575 Hz, three filters; WB, 275–725 Hz, seven filters). Equivalent rectangular bandwidths and center-frequency spacing of the filters were both 75 Hz, as in previous studies (Davidson *et al.*, 2006, 2009). Filter weights were fit individually to each test subject’s behavioral results using linear logistic regression. Alternative multiple-channel models based on broader filters (200-Hz equivalent rectangular bandwidth; consistent with estimates of peripheral tuning bandwidth in budgerigars) did not improve model DVCs to WB behavioral results.

For the NB condition, filter weights were positive for the 500-Hz center frequency and negative for the adjacent center frequencies, similar to previous results in human listeners (Fig. 7). Budgerigar behavioral responses were moderately correlated to the decision variable of the multiple-channel energy-detector model for both trial types in all test subjects [Figs. 3(e), 4(a), and 4(c)]. The DVC of the multiple-channel model to behavioral performance in NB noise was similar to that of the single-channel ENV model and considerably lower than that of the best-fitting single-channel model based on combined energy and ENV cues. Thus, budgerigars appear more likely to use a combination of single-channel energy and ENV cues for 500-Hz tone detection in NB noise than a multiple-channel strategy.

Filter weights for the WB condition were generally positive for center frequencies from 425 to 575 Hz and negative for higher and lower center frequencies with some variability across test subjects (Fig. 7). Compared to previous human results, the peak in filter weight surrounding the 500-Hz test frequency was considerably wider—another result likely associated with species differences in peripheral tuning bandwidth. Behavioral responses in WB noise were well correlated to the decision variable of the multiple-channel model for both tone-plus-noise and noise-alone trials [Figs. 4(b), 4(d), and 6(e)]. The DVC of the multiple-channel model in WB noise was slightly greater than that of the

FIG. 6. (Color online) Comparison of budgerigar behavioral responses in roving-level WB noise to decision variables of the same models in Fig. 3. Top and bottom panels follow the same convention as in Fig. 3. Results are based on 23 501 trials pooled across test subjects. The d' value of the subjects was 1.23.

best-fitting single-channel models (multiple-channel vs ENV model, 0.067 ± 0.022 ; $t_{17} = 3.05$, $p = 0.007$; multiple-channel vs combined model, 0.051 ± 0.022 ; $t_{17} = 2.31$, $p = 0.034$; pairwise differences in mixed-model LS means). However, note that decision variables were highly correlated between the multiple-channel and ENV models (Table III), and the multiple-channel model was fit based on 4–6 more parameters than the single-channel models. Therefore, single- and multiple-channel strategies may be equally likely to explain 500-Hz tone detection by budgerigars in WB noise.

G. Variance explained in behavioral performance

The split-half DVC quantifies the correlation of behavioral responses between two groups of identical stimulus trials (here, between even- and odd-numbered presentations of the same stimulus waveform), providing a measure of within-subject consistency of behavioral performance. The square root of the split-half DVC defines the maximum possible decision correlation of any model (Sebastian and Geisler, 2018) and was used to calculate the proportion of predictable variance accounted for by the best-fitting model (i.e., as the square of the model correlation coefficient divided by the split-half DVC). Split-half DVCs in NB noise were 0.724 ± 0.090 for tone-plus-noise trials and 0.688 ± 0.054 for noise-alone trials (means $\pm$ SD). The best-fitting model (combining energy and ENV cues) accounted for $78.6 \pm 9.8\%$ of the predictable variance for NB tone-plus-noise trials and $75.2 \pm 5.5\%$ of the predictable variance for NB noise-alone trials.

The DVC of the best-fitting model in WB noise (i.e., the multiple-channel energy-detector model) was somewhat lower than that of the best-fitting NB model (tone-plus-noise trials, $t_3 = -11.86$, $p = 0.0013$; noise-alone trials, $t_3 = -5.25$,

$p = 0.013$; paired t -tests; Fig. 4). This difference was partly because the maximum predictable variance was lower for the WB condition, as indicated by reduced split-half DVCs of the average subject with itself (tone-plus-noise trials, 0.549 ± 0.102 ; noise-alone trials, 0.460 ± 0.073 ; means $\pm$ SD). Furthermore, the best-fitting model for WB noise accounted for a smaller proportion of the predictable variance (tone-plus-noise trials, $60.7\% \pm 13.2\%$; noise-alone trials, $65.5\% \pm 28.7\%$; compared to 75%–80% for both trial types in NB noise). In summary, behavioral correlation to model decision variables was lower for WB than for NB noise due to both greater variability of behavioral performance across identical stimulus trials and a less accurate fit of the best-fitting model for the WB noise condition.

IV. DISCUSSION

Tone-in-noise detection was studied in budgerigars trained through operant conditioning to detect a 500-Hz tone in NB and WB noise waveforms from an ensemble of reproducible maskers. Detection patterns, consisting of average hit and FA rates across noise waveforms, were repeatable, highly correlated between the four test subjects, and correlated to previous results in humans. DVC analyses of roving-level results showed that in NB noise, trial-by-trial behavioral responses were most strongly correlated to predictions of a single-channel model combining energy and ENV cues. In contrast, behavioral responses in WB noise were best predicted by (1) a model based primarily on ENV cues with lower weighting of energy or (2) a multiple-channel energy detector.

Detection patterns across ensembles of reproducible noise waveforms were correlated between budgerigars and humans for both WB and NB test conditions, suggesting

overlap between these species in the cues used for tone-in-noise detection. Previous studies show that average detection patterns of human subjects are consistent with model predictions based on single-channel energy and ENV cues (Davidson *et al.*, 2006, 2009) and are best explained by a nonlinear model that combines primarily energy and ENV cues at the output of a single critical-band filter (Mao *et al.*, 2013, 2015). For WB noise, an alternative multiple-channel energy detector model also explains substantial variance in human detection patterns (Davidson *et al.*, 2006, 2009). As discussed below, these same cues are implicated to support tone-in-noise detection in budgerigars, based on DVC analyses of trial-by-trial behavioral responses.

The results of the DVC analyses suggest that in NB noise, budgerigars used a combination of energy and ENV cues to detect a 500-Hz tone in competing noise. A single-channel model combining energy and ENV cues more accurately predicted behavioral responses across trials than alternative models based on energy, ENV, or TFS cues alone. While the increase in DVC of the combined model over the energy model was small, the sensitivity (d') of the combined model was sufficient to account for behavioral sensitivity, whereas that of the energy model was not. Thus, a detection strategy based solely on single-channel energy appears impossible. Weighting of energy in the best-fitting combined model was roughly twice that of the ENV cue. DVC analyses further suggest that budgerigar tone-in-noise detection was influenced less by single-channel energy cues when the masker noise was WB. Compared to results in NB noise, roving-level WB behavioral responses were less correlated to predictions of the single-channel energy model, and the best-fitting combined model had lower weighting of the energy cue. Alternatively, budgerigars may have used energy differences across frequency channels to perform the WB task because the multiple-channel energy-detector model predicted slightly more variance in trial-by-trial behavioral responses than the best-fitting single-channel model. Which strategy was, in fact, employed by budgerigars could not be determined due to collinearity of decision variables between the ENV model and the multiple-channel energy-detector model.

Another method to assess perceptual weighting of single-channel energy cues is by comparing detection thresholds between fixed- and roving-level test conditions (Kidd *et al.*, 1989; Richards, 1992). Greater threshold elevation in a roving-level condition suggests stronger dependence of tone-in-noise detection on the energy cue. Budgerigar thresholds were unchanged with roving level in WB noise and increased by ~ 2 dB with roving level in NB noise, consistent with less influence of energy cues when the masker noise was WB (note agreement with DVC results discussed above). Normal-hearing human listeners also show greater threshold elevation with the roving level for NB noise (Kidd *et al.*, 1989; Mao *et al.*, 2015), suggesting similar influence of the single-channel energy cue on tone-in-noise detection in both species. Larger effects of the roving-level paradigm on thresholds in listeners with

sensorineural hearing loss demonstrate increased reliance on single-channel energy cues in this population (Mao *et al.*, 2015).

Greater weighting of energy than ENV cues for the NB test condition is perhaps surprising because for the roving-level stimuli used here, greater weighting of the ENV cue or TFS would yield greater tone sensitivity. Greater than expected weighting of energy may reflect constraints on neural encoding of energy and ENV. Whereas auditory-nerve fibers encode energy and ENV somewhat independently (energy, average rate and spread of excitation; ENV, temporal response patterns; Joris and Yin, 1992), average discharge rates of inferior colliculus (midbrain) neurons in the budgerigar and other species vary with both stimulus energy and ENV structure rather than either cue alone (Henry *et al.*, 2016; Henry *et al.*, 2017a; Joris *et al.*, 2004; Krishna and Semple, 2000; Langner and Schreiner, 1988; Nelson and Carney, 2007). A decision variable based on ENV processing in the inferior colliculus (Mao and Carney, 2015) might better predict trial-by-trial behavioral responses than these stimulus-based variables.

Budgerigar thresholds for 500-Hz tone detection were approximately as sensitive as human thresholds in NB noise but ~ 7 dB less sensitive for the WB condition. This was because budgerigar thresholds increased in WB noise, whereas human thresholds do not (Davidson *et al.*, 2006; Evilsizer *et al.*, 2002), as also observed in prior studies. Previous band-widening experiments show that for a signal frequency of 500 Hz, the “critical bandwidth” below which tone detection thresholds increase with increasing noise bandwidth and above which thresholds are relatively constant is ~ 75 Hz in humans compared to ~ 400 Hz in the budgerigar (Saunders *et al.*, 1978; Saunders *et al.*, 1979). This difference in critical bandwidth likely reflects a species difference in peripheral frequency selectivity, which at this test frequency appears to be sharper in humans and other primates (Joris *et al.*, 2011) than in the budgerigar based on both neurophysiological (Henry *et al.*, 2016; Henry *et al.*, 2017a) and psychophysical methods (Saunders *et al.*, 1978; Saunders *et al.*, 1979; Saunders and Pallone, 1980). For detection in WB noise, broader peripheral tuning in the budgerigar allows a wider effective bandwidth of the noise to interfere with cues for tone detection (energy, ENV, or others) than occurs in humans (compare also filter weighting of the multiple-channel model in Fig. 7). In contrast, for NB noise, the effective bandwidth of the interfering noise is limited primarily by the stimulus in both species rather than peripheral frequency selectivity, and thresholds are roughly equal. Tone-in-noise masking results are more similar between budgerigars and humans at higher frequencies from 2 to 4 kHz (Dooling and Saunders, 1975; Saunders *et al.*, 1978; Saunders *et al.*, 1979), for which peripheral frequency selectivity appears more similar between these species.

Rabbits tested with the same ensemble of WB reproducible noise waveforms as used in the present study also showed repeatable detection patterns, at least, in some animals, which could be explained by energy cues at

FIG. 7. (Color online) Filter weighting of the multiple-channel energy-detector model for NB (left) and WB (right) test conditions. Top panels show results in individual budgerigars ($n=4$). Bottom panels show mean weights of a model fit to pooled budgerigar results (black) compared to mean human results obtained with the same stimuli (from Davidson *et al.*, 2006).

frequencies primarily above that of the tone (Gai *et al.*, 2007). A modified multiple-channel energy detector model could also explain significant variance in rabbit responses. Rabbit detection patterns are uncorrelated with ENV-model predictions, human results, and budgerigar results as shown here, suggesting different cues are used for tone-in-noise detection in this species. Chinchillas may also employ an alternative, “WB” detection strategy for tones in noise that is more influenced by high-frequency neural channels than in humans (Yost and Shofner, 2009). Chinchillas trained to detect a 1000-Hz signal in noise showed similar response rates to 500-Hz and 2000-Hz probe stimuli, even though probe responses were not reinforced. These results were interpreted as a possible limitation of the ability to make behavioral responses based on the output of a single frequency channel.

An accurate model for tone-in-noise detection should predict variance in behavioral responses for both tone-plus-noise and noise-alone trials. However, budgerigar behavioral responses were uncorrelated to TFS-model predictions for noise-alone trials, and the best-fitting combined model

showed low weighting of TFS cues for both noise bandwidths, suggesting that this cue has limited influence on tone-in-noise detection in budgerigars. While this result is surprising in light of robust auditory-nerve phase locking at 500 Hz in birds and mammals (Johnson, 1980; Köppl, 1997; Sachs *et al.*, 1974), previous human studies also suggest limited dependence of 500-Hz detection on TFS processing. Indeed, Davidson *et al.* (2009) and Mao *et al.* (2013) found that a phase-opponency model of TFS processing predicted relatively little variance in average tone-in-noise detection patterns compared to alternative models based on critical-band energy and ENV cues.

The best-fitting model for budgerigar tone-in-noise detection explained less of the predictable variance of trial-by-trial behavioral responses in WB noise than in NB noise. The reason for this is unknown, but the pattern could reflect the greater variety of cues available for the WB test condition. Whereas the detection in NB noise is limited to a small number of frequency channels by the bandwidth of the stimulus, no such constraint exists for WB noise, and test subjects are free to use any cue available, including (and not limited to) differences in energy, ENV statistics, or TFS across multiple channels. The exact cues might change over time as the test subject attempts to maximize the success rate for performing the task, presenting difficulty for establishing a model to describe the average results. This same phenomenon, introduced here as a possible explanation for differences in model fit, would explain why behavioral results were generally less consistent for WB noise compared to NB noise (as quantified with the split-half DVC). A subject might respond differently to the same WB stimulus on different occasions due to changes in the detection strategy.

V. CONCLUSIONS

Budgerigar detection patterns for 500-Hz tone detection in reproducible NB and WB noise waveforms were repeatable, highly correlated between test subjects, and correlated to human results obtained previously with the same stimuli. DVC analyses showed that budgerigars used a combination of single-channel energy and ENV cues for detection in NB noise. In contrast, for WB noise, energy cues carried less weight, and detection was based primarily on ENV cues. An alternative multiple-channel energy detector could also explain WB behavioral results. These findings highlight similarity between budgerigars and humans in the use of energy and ENV cues to detect tones in noise.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health Grant Nos. R01-DC001641 to L.H.C. and R00-DC013792 and R01-DC017519 to K.S.H. Wilson Geisler provided code to analyze DVCs. The authors thank Douglas Schwarz for technical assistance and Brett Tingley and Margaret Youngman for collecting behavioral results.

- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). "Fitting linear mixed-effects models using lme4," *J. Stat. Softw.* **67**, 1–48.
- Biberger, T., and Ewert, S. D. (2017). "The role of short-time intensity and envelope power for speech intelligibility and psychoacoustic masking," *J. Acoust. Soc. Am.* **142**, 1098–1111.
- Carney, L. H., Ketterer, A. D., Abrams, K. S., Schwarz, D. M., and Idrobo, F. (2013). "Detection thresholds for amplitude modulations of tones in budgerigar, rabbit, and human," *Adv. Exp. Med. Biol.* **787**, 391–398.
- Davidson, S. A., Gilkey, R. H., Colburn, H. S., and Carney, L. H. (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, S. A., Gilkey, R. H., Colburn, H. S., and Carney, L. H. (2009). "An evaluation of models for diotic and dichotic detection in reproducible noises," *J. Acoust. Soc. Am.* **126**, 1906–1925.
- Dent, M. L., Dooling, R. J., and Pierce, A. S. (2000). "Frequency discrimination in budgerigars (*Melopsittacus undulatus*): Effects of tone duration and tonal context," *J. Acoust. Soc. Am.* **107**, 2657–2664.
- Dooling, R. J., Lohr, B., and Dent, M. L. (2000). "Hearing in birds and reptiles," in *Comparative Hearing: Birds and Reptiles*, edited by R. J. Dooling, R. R. Fay, and A. N. Popper (Springer, New York), pp. 308–359.
- Dooling, R. J., and Saunders, J. C. (1975). "Hearing in the parakeet (*Melopsittacus undulatus*): Absolute thresholds, critical ratios, frequency difference limens, and vocalizations," *J. Comp. Physiol. Psychol.* **88**, 1–20.
- Dooling, R. J., and Searcy, M. H. (1981). "Amplitude modulation thresholds for the parakeet (*Melopsittacus undulatus*)," *J. Comp. Physiol. A* **143**, 383–388.
- Early, S. J., Mason, C. R., Zheng, L., Evilsizer, M., Idrobo, F., Harrison, J. M., and Carney, L. H. (2001). "Studies of binaural detection in the rabbit (*Oryctolagus cuniculus*) with Pavlovian conditioning," *Behav. Neurosci.* **115**, 650–660.
- Evilsizer, M. E., Gilkey, R. H., Mason, C. R., Colburn, H. S., and Carney, L. H. (2002). "Binaural detection with narrowband and wideband reproducible noise maskers: I. Results for human," *J. Acoust. Soc. Am.* **111**, 336–345.
- Ewert, S. D., Paraouty, N., and Lorenzi, C. (2018). "A two-path model of auditory modulation detection using temporal fine structure and envelope cues," *Eur. J. Neurosci.* (published online).
- Farabaugh, S. M., Dent, M. L., and Dooling, R. J. (1998). "Hearing and vocalizations of wild-caught Australian budgerigars (*Melopsittacus undulatus*)," *J. Comp. Psychol.* **112**, 74–81.
- Fletcher, H. (1940). "Auditory patterns," *Rev. Mod. Phys.* **12**, 47–65.
- Gai, Y., Carney, L. H., Abrams, K. S., Idrobo, F., Harrison, J. M., and Gilkey, R. H. (2007). "Detection of tones in reproducible noise maskers by rabbits and comparison to detection by humans," *J. Assoc. Res. Otolaryngol.* **8**, 522–538.
- Gilkey, R. H., and Robinson, D. E. (1986). "Models of auditory masking: A molecular psychophysical approach," *J. Acoust. Soc. Am.* **79**, 1499–1510.
- Gilkey, R. H., Robinson, D. E., and Hanna, T. E. (1985). "Effects of masker waveform and signal-to-masker phase relation on diotic and dichotic masking by reproducible noise," *J. Acoust. Soc. Am.* **78**, 1207–1219.
- Gleich, O., and Klump, G. M. (1995). "Temporal modulation transfer functions in the European starling (*Sturnus vulgaris*): II. Responses of auditory-nerve fibers," *Hear. Res.* **82**, 81–92.
- Goldberg, J. M., and Brown, P. B. (1969). "Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: Some physiological mechanisms of sound localization," *J. Neurophysiol.* **32**, 613–636.
- Heinz, M. G., Colburn, H. S., and Carney, L. H. (2001). "Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve," *Neural Comput.* **13**, 2273–2316.
- Henry, K. S., Abrams, K. S., Forst, J., Mender, M. J., Neilans, E. G., Idrobo, F., and Carney, L. H. (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, K. S., Amburgey, K. N., Abrams, K. S., Idrobo, F., and Carney, L. H. (2017b). "Formant-frequency discrimination of synthesized vowels in budgerigars (*Melopsittacus undulatus*) and humans," *J. Acoust. Soc. Am.* **142**, 2073–2083.
- Henry, K. S., Neilans, E. G., Abrams, K. S., Idrobo, F., and Carney, L. H. (2016). "Neural correlates of behavioral amplitude modulation sensitivity in the budgerigar midbrain," *J. Neurophysiol.* **115**, 1905–1916.
- Johnson, D. H. (1980). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," *J. Acoust. Soc. Am.* **68**, 1115–1122.
- Jørgensen, S., Ewert, S. D., and Dau, T. (2013). "A multi-resolution envelope-power based model for speech intelligibility," *J. Acoust. Soc. Am.* **134**, 436–446.
- Joris, P. X., Bergevin, C., Kalluri, R., Mc Laughlin, M., Michelet, P., van der Heijden, M., and Shera, C. A. (2011). "Frequency selectivity in Old-World monkeys corroborates sharp cochlear tuning in humans," *Proc. Natl. Acad. Sci. U.S.A.* **108**, 17516–17520.
- Joris, P. X., Schreiner, C. E., and Rees, A. (2004). "Neural processing of amplitude-modulated sounds," *Physiol. Rev.* **84**, 541–577.
- Joris, P. X., and Yin, T. C. T. (1992). "Responses to amplitude-modulated tones in the auditory nerve of the cat," *J. Acoust. Soc. Am.* **91**, 215–232.
- Kidd, G., Mason, C. R., Brantley, M. A., and Owen, G. A. (1989). "Roving-level tone-in-noise detection," *J. Acoust. Soc. Am.* **86**, 1310–1317.
- Köppl, C. (1997). "Phase locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, *Tyto alba*," *J. Neurosci.* **17**, 3312–3321.
- Krishna, B. S., and Semple, M. N. (2000). "Auditory temporal processing: Responses to sinusoidally amplitude-modulated tones in the inferior colliculus," *J. Neurophysiol.* **84**, 255–273.
- Langner, G., and Schreiner, C. E. (1988). "Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms," *J. Neurophysiol.* **60**, 1799–1822.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Macmillan, N. A., and Creelman, C. D. (2005). *Detection Theory: A User's Guide*, 2nd ed. (Lawrence Erlbaum Associates, Inc., Mahwah, NJ).
- Mao, J., and Carney, L. H. (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., Koch, K. J., Doherty, K. A., and Carney, L. H. (2015). "Cues for diotic and dichotic detection of a 500-Hz tone in noise vary with hearing loss," *J. Assoc. Res. Otolaryngol.* **16**, 507–521.
- Mao, J., Vosoughi, A., and Carney, L. H. (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.
- Nelson, P. C., and Carney, L. H. (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.
- Niemiec, A. J., Yost, W. A., and Shofner, W. P. (1992). "Behavioral measures of frequency selectivity in the chinchilla," *J. Acoust. Soc. Am.* **92**, 2636–2649.
- Nienhuys, T. G., and Clark, G. M. (1979). "Critical bands following the selective destruction of cochlear inner and outer hair cells," *Acta Otolaryngol.* **88**, 350–358.
- Okanoya, K., and Dooling, R. J. (1987). "Hearing in passerine and psittacine birds: A comparative study of absolute and masked auditory thresholds," *J. Comp. Psychol.* **101**, 7–15.
- Okanoya, K., and Dooling, R. J. (1990). "Detection of gaps in noise by budgerigars (*Melopsittacus undulatus*) and zebra finches (*Poephila guttata*)," *Hear. Res.* **50**, 185–192.
- Paraouty, N., Ewert, S. D., Wallaert, N., and Lorenzi, C. (2016). "Interactions between amplitude modulation and frequency modulation processing: Effects of age and hearing loss," *J. Acoust. Soc. Am.* **140**, 121–131.
- Pickles, J. O. (1975). "Normal critical bands in the cat," *Acta Otolaryngol.* **80**, 245–254.
- Rhebergen, K. S., Versfeld, N. J., and Dreschler, W. A. (2006). "Extended speech intelligibility index for the prediction of the speech reception threshold in fluctuating noise," *J. Acoust. Soc. Am.* **120**, 3988–3997.
- Richards, V. M. (1992). "The detectability of a tone added to narrow bands of equal-energy noise," *J. Acoust. Soc. Am.* **91**, 3424–3435.
- Sachs, M. B., and Abbas, P. J. (1974). "Rate versus level functions for auditory-nerve fibers in cats: Tone-burst stimuli," *J. Acoust. Soc. Am.* **56**, 1835–1847.
- Sachs, M. B., Young, E. D., and Lewis, R. H. (1974). "Discharge patterns of single fibers in the pigeon auditory nerve," *Brain Res.* **70**, 431–447.

- Saunders, J. C., Denny, R. M., and Bock, G. R. (1978). "Critical bands in the parakeet (*Melopsittacus undulatus*)," *J. Comp. Physiol.* **125**, 359–365.
- Saunders, J. C., and Pallone, R. L. (1980). "Frequency selectivity in the parakeet studied by isointensity masking contours," *J. Exp. Biol.* **87**, 331–342.
- Saunders, J. C., Rintelmann, W. F., and Bock, G. R. (1979). "Frequency selectivity in bird and man: A comparison among critical ratios, critical bands and psychophysical tuning curves," *Hear. Res.* **1**, 303–323.
- Sebastian, S., Abrams, J., and Geisler, W. S. (2017). "Constrained sampling experiments reveal principles of detection in natural scenes," *Proc. Natl. Acad. Sci. U.S.A.* **114**, E5731–E5740.
- Sebastian, S., and Geisler, W. S. (2018). "Decision-variable correlation," *J. Vis.* **18**, 1–19.
- Watson, C. S. (1963). "Masking of tones by noise for the cat," *J. Acoust. Soc. Am.* **35**, 167–172.
- Yost, W. A., and Shofner, W. P. (2009). "Critical bands and critical ratios in animal psychoacoustics: An example using chinchilla data," *J. Acoust. Soc. Am.* **125**, 315–323.
- Zheng, L., Early, S. J., Mason, C. R., Idrobo, F., Harrison, J. M., and Carney, L. H. (2002). "Binaural detection with narrowband and wideband reproducible noise maskers: II. Results for rabbit," *J. Acoust. Soc. Am.* **111**, 346–356.