How canaries listen to their song: Species-specific shape of auditory perception

Adam R. Fishbein,[1,a),b)] Shelby L. Lawson,[2] Robert J. Dooling,[1,a)] and Gregory F. Ball[1,a)]

[1]Psychology Department, University of Maryland, 4094 Campus Drive, College Park, Maryland 20742, USA
[2]Department of Animal Biology, University of Illinois, 505 South Goodwin Avenue, Champaign, Illinois 61801, USA

The melodic, rolling songs of canaries have entertained humans for centuries and have been studied for decades by researchers interested in vocal learning, but relatively little is known about how the birds listen to their songs. Here, it is investigated how discriminable the general acoustic features of conspecific songs are to canaries, and their discrimination abilities are compared with a small parrot species, the budgerigar. Past experiments have shown that female canaries are more sexually responsive to a particular song element—the "special" syllables—and consistent with those observations, it was found that special syllables are perceptually distinctive for canaries. It is also shown that canaries discriminate the subtle differences among syllables and phrases using spectral, envelope, and temporal fine structure cues. Yet, while canaries can hear these fine details of the acoustic structure of their song, the evidence overall suggests that they listen at a more global, phrase by phrase level, rather than an analytic, syllable by syllable level, except when attending to some features of special syllables. These results depict the species-specific shape of auditory perception in canaries and lay the groundwork for future studies examining how song perception changes seasonally and according to hormonal state. © 2019 Acoustical Society of America.
https://doi.org/10.1121/1.5087692

[AMS] Pages: 562–574

I. INTRODUCTION

Songbirds have long been models for the comparative study of vocal learning and complex vocal communication (Fee and Scharf, 2010; Catchpole and Slater, 2008). Much is known about how songbirds learn and produce their songs and how species vary in these learned vocalizations (e.g., Marler, 2004), but relatively less is known about how they perceive species-specific song (Ball and Hulse, 1998). Students of birdsong, as well as casual listeners, have historically been drawn to the musical qualities of birdsong and have relied on rather broad musical and speech terms to describe and analyze the complex songs of birds such as pitch, tempo, tours, syllables, notes, phrases, etc., suggesting that humans, at least, may listen to birdsong in a somewhat global or synthetic mode (Helmholtz, 1863; Terhardt, 1974)—as we might listen to orchestral music. At the same time, the modern science of birdsong has long drawn attention to the fine acoustic details and precision with which birdsong is produced (e.g., Marler, 2004). Here, we marvel at the precision of produced notes, inter-note-intervals, frequency sweeps, and glides of species-typical song—a precision that suggests birds may listen to their song in a more analytic mode (Helmholtz, 1863; Terhardt, 1974), perhaps hearing details in their song to which the human ear is not sensitive (e.g., Prior et al., 2018). Thus, there are

unanswered questions as to how birds hear their song and whether this is similar to the way humans hear their song (Dooling and Prior, 2017). On the one hand, there are plenty of studies suggesting that female birds, listening over periods of days and weeks, are responsive to global features such as the overall number of songs and song complexity (e.g., Kroodsma, 1976; Nowicki and Searcy, 2004). On the other hand, there are studies indicating that birds may hear the fine details in elements of their song to which the human ear is not particularly sensitive such as the "special" syllable in canary song or the temporal fine structure in harmonic zebra finch vocalizations (Dooling and Prior, 2017). These are not mutually exclusive possibilities, but it is important to understand what acoustic information in conspecific song is perceived by songbirds and how auditory sensitivities shape the song that the birds learn to produce. Here, we examined the sensitivity of canaries to the fine details of their species-specific song.

Among songbirds, the domestic canary (*Serinus canaria*) has long been a favorite for behavioral and neuroanatomical investigations of vocal learning and production (Waser and Marler, 1977; Nottebohm et al., 1976). These birds have been bred as much for their melodious song as for their plumage for hundreds of years (Marler, 2004). Male canaries produce a courtship song that is acquired by a combination of flexible imitation and innate song constraints (Gardner et al., 2005). The song is produced seasonally and consists of basic units, syllables (typically 20–200 ms in duration), repeated in groups to form broader units, phrases (typically 500 ms–3 s), which are flexibly sequenced to form

a)Also at: Neuroscience and Cognitive Science Program, University of Maryland, College Park, MD 20742, USA.
b)Electronic mail: afishbei@terpmail.umd.edu

songs (typically 5–15 s long; Markowitz *et al.*, 2013). As "open-ended" learners (Beecher and Brenowitz, 2005), they acquire a set of song syllables during a juvenile sensitive period and can later integrate new syllables into their repertoire as adults (Lehongre *et al.*, 2006). The repertoire of an adult male canary can have as many as 25–35 distinct syllable types (Güttinger 1985; Markowitz *et al.*, 2013). Moreover, Vallet and Kreutzer (1995) showed that females are particularly sexually responsive to a class of syllables that they termed *A*-syllables, referred to here as special syllables or special phrases. These frequency-modulated phrase elements are composed of two notes with a wide bandwidth of about 4 kHz, extending to low frequencies, and a high syllable repetition rate exceeding 15 syllables per second (syll/s; Vallet and Kreutzer, 1995; Vallet *et al.*, 1998; Suthers *et al.*, 2012). These syllables, articulated using opposite sides of the syrinx, are hypothesized to be indicators of male fitness as they are difficult to produce (Suthers *et al.*, 2012).

Behavioral experiments have shown that both male and female canaries are sensitive to the acoustic differences between special and non-special phrases. Females will increase copulation solicitation displays (CSDs) when repetition rate and bandwidth are increased, when a two-note syllable structure is heard rather than a one-note structure, and when low frequencies are included (Draganoiu *et al.*, 2002; Pasteau *et al.*, 2004; Pasteau *et al.*, 2007). Playback experiments also show that male canaries call less in response to song containing special phrases compared to song lacking those phrases (Parisot *et al.*, 2002). In addition, male canaries sing longer special phrases when in the presence of either a male or female conspecific compared to when alone, suggesting that these vocalizations function in both inter- and intra-sexual communication (Kreutzer *et al.*, 1999). Thus, both male and female canaries hear the differences between special and non-special phrases, and the attractive acoustic features seem to carry biologically important information.

Neurophysiological studies have investigated how conspecific song is produced and encoded in the canary brain. Del Negro *et al.* (2000) found that neurons in the song control nuclei HVC (proper name) of sexually receptive females decrease in spike rate in response to sexually attractive song phrases. HVC neurons in reproductively active males, on the other hand, showed no sensitivity to the special phrases, but did fire synchronously with the onset of phrases, independent of phrase type. Electrophysiological recordings from neurons in the caudomedial nidopallium (NCM), a telencephalic auditory area that is selective for conspecific vocalizations in songbirds, found narrower tuning and more transient responses in canaries than in zebra finches, which may allow enhanced processing of the acoustic features in canary song (Terleph *et al.*, 2007). The previous study showed no sex differences in NCM responses in canaries, indicating that the auditory systems of both sexes are shaped to process canary song similarly. The sex differences in HVC neural coding could be due to an absence of tonic responses in the HVC neurons of reproductively active female canaries (Del Negro and Edeline, 2001). These findings indicate how the canary brain may be tuned for conspecific song, but less is known

about canary perception of song from a psychoacoustic perspective.

As far as we know, there have not been any psychophysical investigations of how canaries hear the acoustic details of their song, though a considerable amount is known about basic hearing capacities in canaries from such approaches. Canaries, like most birds, hear best at around 3 kHz (with a threshold of 5–15 dB) over a range that extends from about 2 to 5 kHz (Okanoya and Dooling, 1987a). Also, canaries of the Belgian Waterslager strain, along with hybrid Waterslager-Roller canaries, have elevated absolute thresholds at high frequencies (Okanoya and Dooling, 1987b; Okanoya *et al.*, 1990). Even though most canary vocalizations are largely tonal, most tests also reveal that canaries (and other birds) are not especially sensitive to frequency changes compared to other vertebrates, including humans (Lauer *et al.*, 2007). By contrast, it is quite clear that canaries, and other birds, tested in discrimination experiments show a much more precise temporal resolution for complex sounds than do humans (Dooling *et al.*, 2002; Dooling and Prior, 2017).

The purpose of the following experiments is to explore with a rigorous psychophysical approach how well canaries perceive the acoustic features of their songs, including special syllables. We test the hypothesis that special syllables are perceptually distinctive for canaries and determine whether canaries listen to their song in more of an analytic or synthetic mode, which is a framework with a long history of use in characterizing how humans listen to music and speech (Helmholtz, 1863; Terhardt, 1974; Tsoumani and Postma-Nilsenova, 2013; Hermann, 2002; Supper and Bijsterveld, 2015). These results may then provide a basis for future investigations as to whether song perception changes seasonally and according to hormonal state as is the case for song production in male canaries.

II. GENERAL METHODS

A. Subjects

All canaries tested in this study were of the American Singer strain. In experiments 1 and 2, eight canaries (four females and four males) were tested. In experiments 3 and 4, six canaries (four females and two males) were tested. In experiment 5, six canaries (three females and three males) were tested. In experiments 1–4, two budgerigars were tested as reference species. In experiment 5, one budgerigar and one zebra finch were tested as reference species. Birds were housed in an avian vivarium and kept on a winter photoperiod (8 h light, 16 h dark). Birds were maintained at 90%–95% of their free feeding weights and given free access to water.

B. Equipment

All psychoacoustic experiments took place in a wire cage anchored inside of a sound-attenuated chamber (Industrial Acoustics Company, Bronx, NY, model IAC-3) lined with acoustic foam. Birds sat on a perch and had access to food through an opening in the floor of the cage. Millet

was delivered through a food hopper, which was brought up in the food opening through activation of a solenoid. Two response keys, each consisting of a light-emitting diode (LED) attached to a microswitch, were mounted to the wall of the cage directly in front of the perch.

Stimuli were stored as wav files on an Intel Core 2 Duo computer (Mid Atlantic Data Systems, Gaithersburg, MD), which controlled all experiments. The computer operated a Tucker Davis Technologies System 3 DD1 stereo analog interface (Alachua, FL), sent signals through a Crown D-75 amplifier (Elkhart, IN), and finally to an Orb full range point source speaker (Model Mod1, Orb Audio, Sherman Oaks, CA) which was placed 40 cm above the bird's head when standing on the perch.

C. Stimuli

In these experiments, the test stimuli consisted of individual phrases that were either natural or composed of synthetic or manipulated syllables. Phrases and syllables were extracted from song recordings using Adobe Audition (ver: 2015.2, Adobe, San Jose, CA) and MATLAB (MathWorks, Natick, MA). We defined a syllable as a collection of notes (a note is a continuous trace on a spectrogram) and a phrase as a repetition of the same syllable type (Catchpole and Slater, 2008). Synthesized stimuli were generated using MATLAB and Adobe Audition. Stimuli included natural phrases, synthetic phrases containing repeated syllables, synthetic phrases containing reversed individual syllables, and synthetic phrases where inter-syllable interval was manipulated. Natural vocalizations were recorded using an AE3300 microphone (Audio-Technica, Stow, OH) and a Marantz portable solid-state recorder (Model PMD670, Mahwah, NJ) at a sampling rate of 48 kHz from the songs of two male canaries, one of which was also a subject in this study. For every experiment, background and target sounds were attenuated to be at the same sound level [∼60 dBA sound pressure level (SPL), peak amplitude]. Sound level was measured with the $\frac{1}{4}$ in. microphone of a BK precision sound level meter (#732, Brüel & Kjær, Nærum, Denmark) placed 7 cm above the bottom of the cage, at approximately the location of the bird's head when it was positioned in front of the observation key. Recordings were resampled at 24 414 Hz for playback in the psychoacoustic setup as appropriate for the Tucker Davis System 3 equipment.

D. Testing procedure

The procedures have been described in detail previously (Vernaleo and Dooling, 2011). Birds were initially trained on a detection task to peck the observation key (the left LED) for a random interval of 2–7 s in order to present a tone (i.e., the target). After the tone plays, if the bird pecks the report key (the right LED) within 3 s, they receive 2 s access to a food reward. This response is recorded as a hit. Failure to peck within this time is recorded as a miss and is not rewarded. Birds were tested on both low- and high-frequency pure tones at several different levels to ensure that they could hear sounds of 60 dBA SPL. Low-frequency

thresholds were normal, while high-frequency thresholds were more variable.

Next, birds were trained on a discrimination task. Here, the birds were trained to peck the observation key while a background sound repeats 2/s, until they hear a new sound (i.e., the target), which alternates with the background. If birds peck the report key (the right LED) within 3 s, they receive 2 s access to a food reward. This is recorded as a hit. Failure to peck within this time is recorded as a miss and is not rewarded. The target sound alternates with the background sound during the response window, which was set for all experiments in this study so that the bird could hear the target sound repeated twice (the response window varied between 3 and 3.5 s for each experiment depending on the length of the stimuli).

In a session, birds generally ran 100 trials consisting of 10 10-trial blocks. Three of the trials within each block are sham trials in which the background sound is presented as a target. If the bird pecks the response key during the response window of a sham trial, it is recorded as a false alarm and all of the lights in the chamber are extinguished during a short blackout period (3–10 s based on each bird's propensity for false responding). If birds correctly withhold pecking during a sham trial, this is recorded as a correct rejection but is not rewarded. Birds usually ran two sessions a day, each lasting approximately 20–40 min.

E. Data analysis

In these experiments, the raw data consisted of both percent correct values and response latency values. Since birds varied somewhat in their false alarm rates, d' was used as a measure of sensitivity in this study in order to equalize the performance of conservative and liberal subjects. d' is defined in terms of the z-score of hit rate and false alarm, converting these values to standard deviation units (Macmillan and Creelman, 2005; Dooling and Okanoya, 1995):

$$d' = z(\text{hit rate}) - z(\text{false alarm rate}).$$

To avoid infinite values in cases of 100% or 0% rates for hits or false alarms, the values were converted to $1 - 1/(2N)$ and $1/(2N)$, respectively, where N is the number of trials on which the percentage was based (Macmillan and Creelman, 2005). When observers cannot discriminate at all, $d' = 0$. When, for example, the hit rate is 100% ($N = 70$) and the false alarm rate is 0% ($N = 30$), $d' = 4.28$. An arbitrary d' of 1.0 is often taken as a reference well above chance to evaluate performance (Tu and Dooling, 2012).

For statistical analyses, t-tests and analyses of variance (ANOVAs) were performed in SPSS (version 2.3, Chicago, IL). Cohen's d was calculated to determine effect size for t-tests, and partial eta squared was calculated using SPSS to determine effect size for ANOVAs.

III. EXPERIMENT 1: PERCEPTION OF NATURAL SONG PHRASES

There is considerable variation among different syllable types in canary song and song generally consists of a sequence of approximately 1 s long phrases, each made up

of the repetition of a single syllable type. This raises several perceptual questions such as whether subtle variations between similar syllables within a phrase are discriminable, and whether some phrases sound more similar than other phrases to canaries. For instance, previous behavioral studies have drawn a distinction between special syllables and other non-special syllable types (Vallet and Kreutzer, 1995; Vallet et al., 1998; Suthers et al., 2012). In contrast to simpler one-note non-special syllables, these previous studies have defined special syllables as being composed of two notes with a wide bandwidth of about 4 kHz, extending to low frequencies, and a high syllable repetition rate exceeding 15 syll/s (Vallet and Kreutzer, 1995; Vallet et al., 1998; Suthers et al., 2012). All of the various syllables the canaries produce are discriminable by spectrographic analysis and to casual human listeners, but the extent to which these syllables sound different or similar to each other for canaries is unknown.

In this experiment, we used multidimensional scaling (MDS) to produce a map of canary perception of these natural sound signals. These techniques have been used to understand the perceptual organization of complex acoustic stimuli in other animals, such as budgerigars (Dooling et al., 1987), great tits (Pohl et al., 2015), mice (Neilans et al., 2014), and humans (Yang and Fox, 2014; Kisenwether and Prosek, 2014). Basically, the technique involves collecting response data from all pairwise stimulus comparisons in a discrimination task and then analyzing such measures by MDS to arrive at the perceptual distances between stimulus objects. The resulting spatial map from MDS enables one to see the bird's perceived similarity among phrases and can provide clues about the acoustic dimensions upon which the birds are making their discriminations. Here, we asked how canaries perceptually group different phrases, what acoustic features underlie these distinctions, and if any sex differences or species differences exist.

A. Stimuli

Birds were tested on a set of eight natural phrases extracted from the songs of two male canaries (four from each). One of the male canaries was also a subject in this experiment. The phrases were manually classified as being special or non-special through inspection of the spectrograms and time waveforms using the general parameters described above from Vallet and Kreutzer (1995), Vallet et al. (1998), and Suthers et al. (2012). Four of the phrases were classified as special [two from each bird; Fig. 1(a), phrases C and D; Fig. 1(b), phrases G and H]. The four other phrases were classified as non-special [again, two from each bird; Fig. 1(a), phrases A and B; Fig. 1(b), phrases E and F]. Though phrases G and H are just below a repetition rate of 15 syll/s, they were still classified as special because of their wide bandwidth and two-note structure. Although phrase F has a similar repetition rate, it was classified as non-special because of its narrower bandwidth and lack of apparent two-note structure. Acoustic measurements for the eight phrases are summarized below (Table I).

B. Procedure

All subjects were tested on the same set of eight natural phrases. Each phrase served as the background while the other seven phrases were played as targets until all possible pairwise comparisons were available. Birds were tested on each phrase as a background and the others as the targets until their performance across five successive sessions stabilized (defined here as when neither false alarm rate nor hit rate changed by more than 15% between each successive set of 100 trials). These sessions (500 trials) were then used for analysis. Once each phrase had served as the background stimulus, there was a matrix consisting of all possible pairings of the set of eight song phrases. This matrix of response latency scores was then analyzed using the ALSCAL MDS algorithm in SPSS (Chicago, IL).

For ALSCAL, the individual differences Euclidean distance scaling model was used to produce both a group space for all subjects of each species and a set of subject weights reflecting the salience of each group dimension for each listener (Yang and Fox, 2014). Response latencies were converted into dissimilarity scores by subtracting the original data from the highest possible latency in this experiment, 3.5 s (Giguére, 2006). A square asymmetric shape of the data was chosen since scores for the same pairwise comparisons were different when background and target were switched. "Ratio" was chosen as the level of measurement since response latency measurements are on a ratio scale. And "unconditional" was chosen for conditionality since data from each individual's matrix can be meaningfully compared to each other (Giguére, 2006).

For MDS, stress scores, a measure of the "badness-of-fit" of the representational map, were used to choose the optimal dimensionality of the stimulus space (Giguére, 2006). The general rule of thumb is to choose the solution for which further increases in dimensions does not decrease stress (Borg and Groenen, 2005). This optimal dimensionality can be identified as the elbow in the curve of a scree plot, which plots the stress score for each dimensionality (see supplementary material[1]). Based on these factors, two dimensions were chosen for analysis of the canary perceptual map. A three-dimensional solution was also produced but it did not seem to add a meaningful dimension to the analysis. The units in the MDS perceptual maps are arbitrary.

The following acoustic measurements for the eight natural phrases were computed in Raven Pro 1.4 (Cornell Lab of Ornithology, Ithaca, NY) using a fast Fourier transform (FFT) of 1024 in a Hann window with 50% overlap: first quartile frequency, third quartile frequency, aggregate entropy, average entropy, bandwidth 90%, center frequency, frequency 5%, frequency 95%, peak frequency. Syllable repetition rate was computed by dividing the duration of each phrase by the number of syllables in each phrase. Syllable durations were extracted for each phrase using Avisoft-SASLab Pro (Glienicke, Germany).

C. Results and discussion

Across all trials used for analysis, canaries discriminated phrases at an average d' of 2.95, showing that these phrases

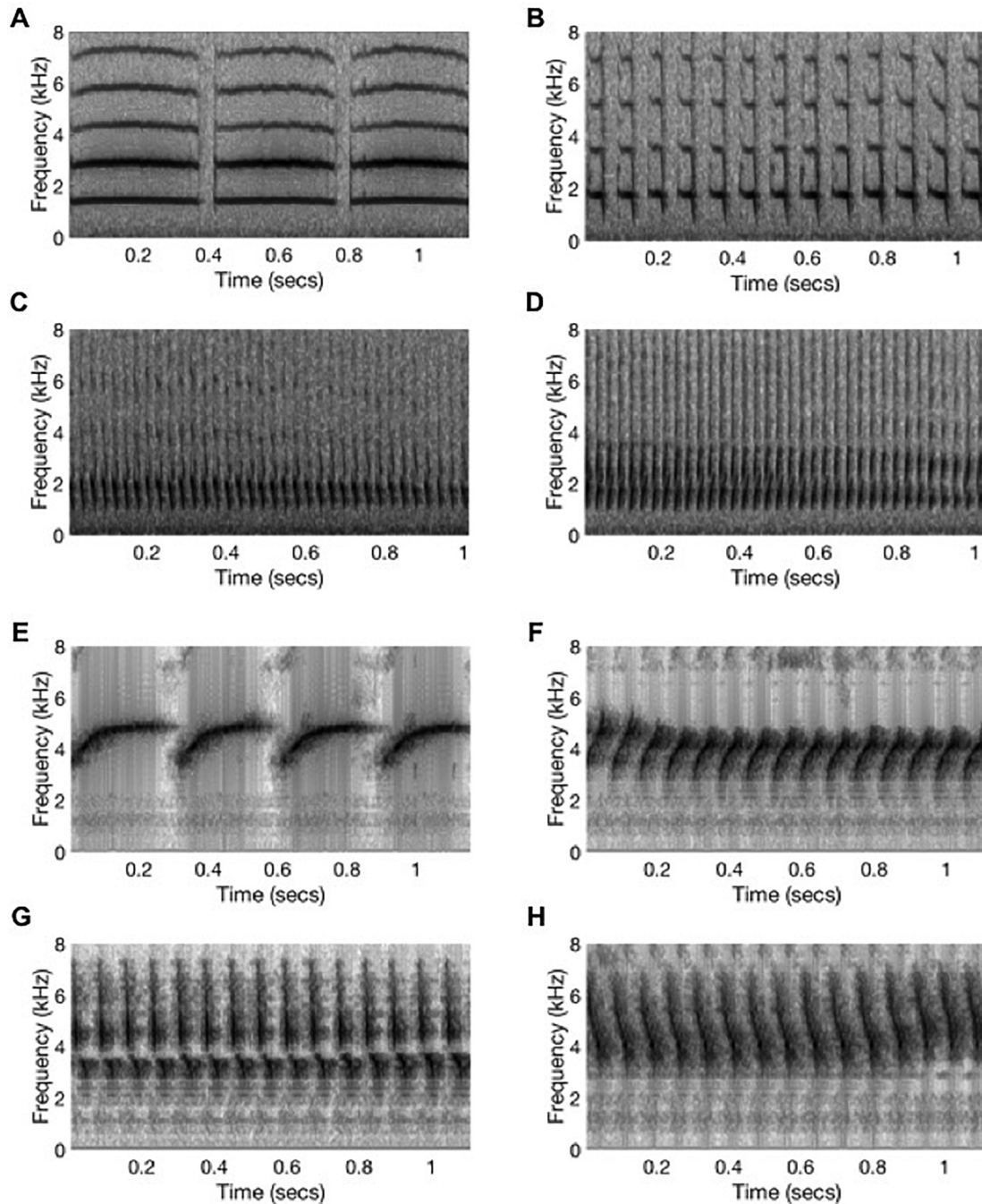J. Acoust. Soc. Am. **145** (1), January 2019

Fishbein et al.   565

FIG. 1. (a) Four song phrases from one male canary. Phrases C and D display characteristics of special syllables, while phrases A and B do not and are thus classified as non-special. Phrases A–D here are A1–A4 in MDS plots in Figs. 3 and 4. A syllable is a collection of notes (a note is a continuous trace on a spectrogram) and a phrase is a repetition of the same syllable type. (b) Four song phrases from another male canary. Phrases G and H display characteristics of special syllables, while phrases E and F do not and are thus classified as non-special. Phrases E–H here are B1–B4 in MDS plots in Figs. 3 and 4.

were highly discriminable. Budgerigars were even more sensitive to overall differences in phrases, with an average d' of 5.21. This species difference was significant as calculated using an independent samples t-test [$t(8) = -3.87, p = 0.005, d = 3.058$]. The subject weight plot (Fig. 2) for the group MDS solution also indicates that the species are weighing acoustic features differently. There was no significant sex difference for canary performance on phrase discriminations [$t(6) = 1.20, p = 0.276, d = 0.846$], though the subject weight plot suggests that males and females may be weighing acoustic features somewhat differently.

The two-dimensional (2-D) canary perceptual map [Fig. 3(a)] shows that special phrases from the same bird are more similar than are non-special phrases from the same bird (average Euclidean distance: 0.12 vs 1.82). Figure 3(b) shows this is not the case for budgerigars (average Euclidean distance: 1.01 vs 0.85). In other words, canaries perceive special phrases of individual birds as more similar to each other than do budgerigars, consistent with the idea that special phrases are indeed perceptually distinctive for canaries.

Fitting the dimensions of the perceptual space with the acoustic features of the phrases can indicate what auditory

TABLE I. Acoustic measurements for all eight natural phrases used in experiment 1. First quartile frequency, third quartile frequency, aggregate entropy, average entropy, bandwidth 90%, center frequency, frequency 5%, frequency 95%, and peak frequency were computed in Raven Pro. Syllable durations were extracted using Avisoft (Glienicke, Germany). SD = Standard Deviation.

	Syllable repetition rate (syll/s)	Average syllable duration (s)	Average interval duration (s)	Q1 frequency (Hz)	Q3 frequency (Hz)	Aggregate entropy (u)
Average special	24.310	0.038	0.013	2610.7	3331.9	5.82
SD special	12.215	0.020	0.008	1282.5	1421.8	0.31
Average non-special	7.987	0.176	0.035	3182.9	3403.4	4.117
SD non-special	5.656	0.137	0.027	1144.0	1346.8	1.166
	Average entropy (u)	BW 90% (Hz)	Center frequency (Hz)	Frequency 5% (Hz)	Frequency 95% (Hz)	Peak Frequency (Hz)
Average special	5.202	1478.2	2992.2	2336.5	3814.7	2759.7
SD special	0.3	430.5	1466.1	1303.4	1590.7	1114.8
Average non-special	3.334	923.9	3331.9	2628.6	3552.4	3361.7
SD non-special	0.856	511.1	1293.1	1320.5	1415.5	1344

properties the birds may be using to discriminate the sounds. Here, Pearson product moment correlations were performed to fit the coordinates of each phrase in each dimension with the acoustic properties of the phrases (see supplementary material[1]). For the canary plot, dimension 1 is significantly correlated with aggregate entropy ($r = 0.829$, $p = 0.011$) and frequency 95% ($r = 0.798$, $p = 0.018$). Also, the special phrases for each bird and the non-special phrases from bird A are each grouped together on this dimension suggesting that it could capture the global features of bird identity and phrase type. On dimension 2, canary performance is significantly correlated with syllable repetition rate ($r = -0.933$, $p = 0.001$), first quartile frequency ($r = 0.847$, $p = 0.008$), third quartile frequency ($r = 0.708$, $p = 0.05$), center frequency ($r = 0.788$, $p = 0.02$), frequency 5% ($r = 0.731$, $p = 0.039$), and peak frequency ($r = 0.808$, $p = 0.015$). These results suggest that overall canaries are using a combination of particular acoustic features along with global perceptual properties to discriminate song phrases. For the budgerigar plot, dimension 1 is significantly correlated with first quartile frequency ($r = -0.871$, $p = 0.005$), third quartile frequency ($r = -0.853$, $p = 0.007$), center frequency ($r = -0.851$, $p = 0.007$), frequency 5% ($r = -0.898$, $p = 0.002$), frequency 95% ($r = -0.888$,

$p = 0.003$), and peak frequency ($r = -0.805$, $p = 0.016$), and dimension 2 is not significantly correlated with any of the acoustic features tested, suggesting that the budgerigars are using particular frequency features to discriminate the song phrases.

2-D perceptual maps were also separately produced for male [Fig. 4(a)] and female [Fig. 4(b)] canaries. They both show clustering of special phrases from the different birds and also subtle sex differences in the arrangement of the phrases. Property-fitting offers an explanation for these differences. The stimulus location on the first dimension for both male and female canaries is significantly correlated with first quartile frequency, third quartile frequency, center frequency, frequency 5%, frequency 95%, and peak frequency. On the second dimension for male canaries, the stimulus location correlates significantly with average syllable duration ($r = -0.766$, $p = 0.027$), aggregate entropy ($r = 0.910$, $p = 0.002$), and average entropy ($r = 0.921$, $p = 0.001$). The stimulus plot of female canaries on the second dimension significantly correlates with bandwidth 90% ($r = 0.814$, $p = 0.014$). These results show that while there is no significant difference in male and female performance in discriminating song phrases, the two sexes may be focusing on different acoustic features to perform the task (at least while in non-breeding condition). Overall, though, these perceptual maps suggest that male and female canaries are perceiving song phrases similarly.

IV. EXPERIMENT 2: DISCRIMINATION OF REVERSED SONG SYLLABLES

The previous experiment showed that canaries hear differences among their song phrases and identified some of the acoustic features used in discriminations. While the syllables within a song phrase look and sound very similar, close inspection of sonograms and other signal analysis reveals that no two syllables are identical. The question is whether canaries are sensitive to differences among syllables within a phrase or if, instead, they listen to the phrase in a more global, synthetic mode of listening. This experiment focuses on discrimination at the syllable level using altered natural syllables. When a syllable is temporally reversed, the overall
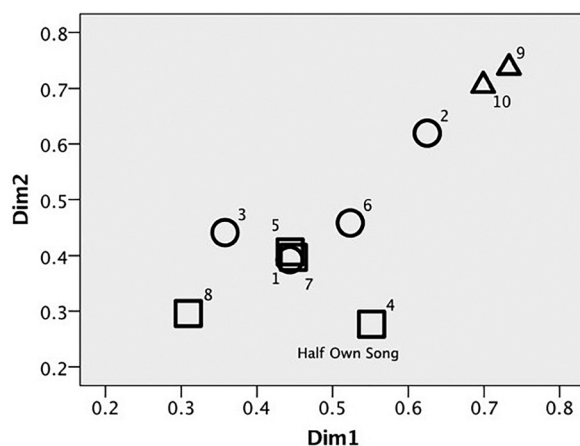

FIG. 2. Subject weight plot for 2-D MDS solution for all birds. Triangles = Budgerigars, Circles = Female Canaries, Squares = Male Canaries. Half of the stimuli were from the song of bird 4 who was also a subject in the experiment.

J. Acoust. Soc. Am. **145** (1), January 2019
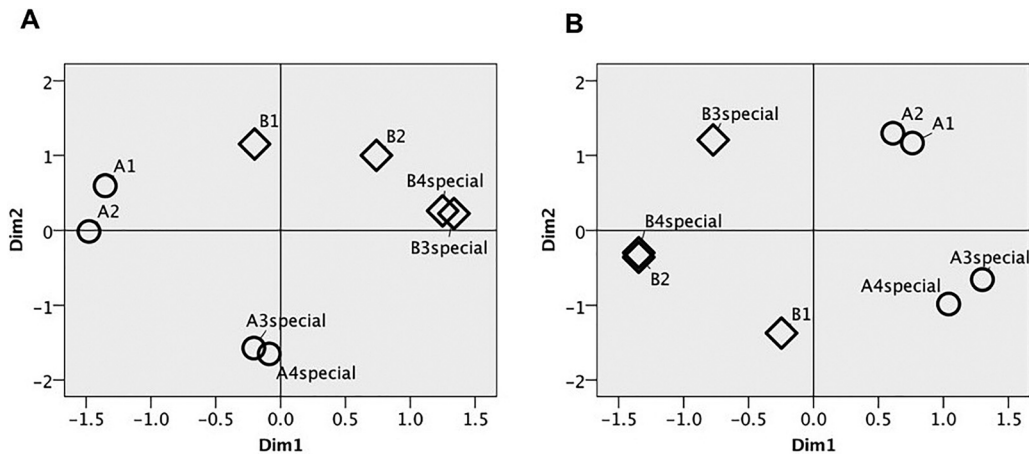
Fishbein et al.   567

FIG. 3. (a) Canary 2-D MDS solution. (b) Budgerigar 2-D MDS solution. Circles indicate phrases taken from recordings of bird A. Diamonds indicate phrases taken from recordings of bird B.

spectrum and amplitude remains the same while the temporal details of the fine structure within a syllable and the shape of the envelope are changed. Past psychoacoustic experiments have shown that zebra finches are very sensitive to syllable reversals within a motif no matter where they occur—probably the result of an unusual sensitivity to changes in temporal fine structure in harmonic sounds (Vernaleo and Dooling, 2011). Here the question is how sensitive canaries are to similar manipulations in syllable structure within a phrase.

A. Stimuli

Individual syllables were reversed in five synthetic phrases. Synthetic phrases were composed of a single syllable extracted randomly from the natural phrase and repeated at a constant rate so that the synthetic phrase matched the duration of the natural phrase. Each stimulus set contained seven target phrases, each of which contained a single syllable reversed at a different location while the order of syllables remained intact. Birds were tested on two sets of stimuli containing special syllables and three sets containing non-special syllables (see supplementary material[1]).

B. Procedure

All subjects were tested on the same five sets of eight stimuli. A synthetic phrase consisting of syllables repeated in the natural forward direction was played in the background. Seven synthetic phrases, each the same as the background except for a single syllable reversal, were played as targets. Birds were run until their performance across two successive sessions stabilized (defined here as when neither false alarm rate nor hit rate changed by more than 15% between each successive set of 100 trials). The final 100 trials were used for analysis.

C. Results and discussion

Across all phrases and reversal locations, canaries discriminated single syllable reversals at an average d' of 1.90. Budgerigars performed significantly better at the same task as a mixed model ANOVA with phrase as a within-subjects factor and species as a between-subjects factor showed a significant main effect of species $[F(1,8) = 28.49, \ p = 0.001, \ \eta^2 = 0.966]$. In Vernaleo and Dooling (2011), zebra finches discriminated single syllable reversals in zebra finch motifs greater than 90% of the time, with a false alarm rate no
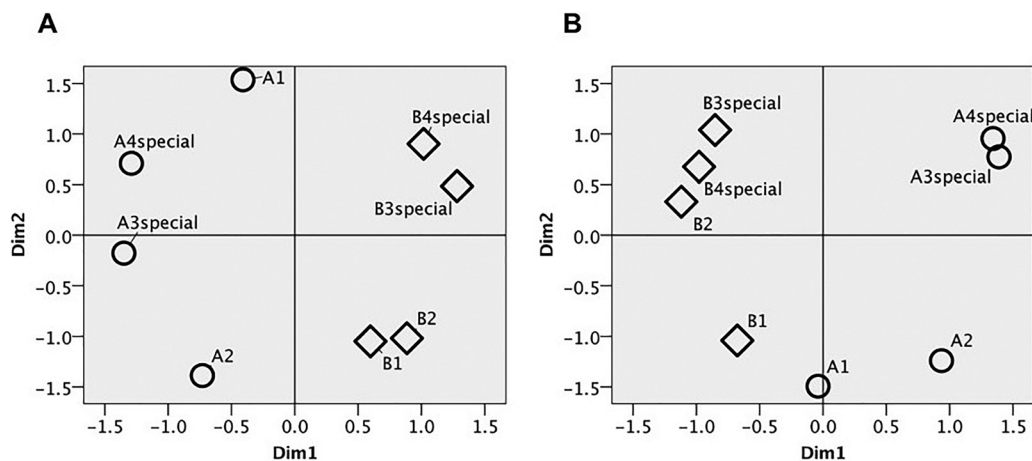

FIG. 4. (a) Female canaries 2-D MDS solution. (b) Male canaries 2-D MDS solution. Circles indicate phrases taken from recordings of bird A. Diamonds indicate phrases taken from recordings of bird B.

greater than 20%. Thus, like zebra finches, canaries and budgerigars showed sensitivity to temporal fine structure and amplitude envelope shape. Also, a mixed model ANOVA with phrase as a within-subjects factor and sex as a between-subjects factor showed no sex difference in canary discrimination of single syllable reversals across phrases $[F(1,6) = 0.006, p = 0.94, \eta^2 = 0.001]$.

Interestingly, canaries performed significantly better on reversals of non-special syllables compared to reversals of special syllables (Fig. 5) $[t(7) = 4.96, p = 0.002, d = 1.751]$, though performance on both well exceeded chance. Budgerigars, on the other hand, showed no significant difference between those syllable types $[t(1) = 0.398, p = 0.759, d = 0.282]$. This performance difference for canaries may be due to differences in syllable duration between special and non-special syllables used, as there was a positive correlation between syllable duration and d' for canaries $[r = 0.936, n = 5$ phrases, $p = 0.019]$. No other acoustic features tested correlated with d' for canaries. Budgerigars did not show the same correlation of d' and syllable duration $[r = 0.359, n = 5$ phrases, $p = 0.553]$, nor correlations with any of the other acoustic features tested. Thus, fine structure and envelope cues are less salient for canaries in special phrases than in non-special phrases, suggesting a trade-off between decreased spectral resolution and increased temporal resolution when perceiving special syllables.

V. EXPERIMENT 3: DISCRIMINATION OF SPECTRAL-TEMPORAL VARIATION IN NATURAL SONG PHRASES

Experiment 2 showed that canaries can perceive relatively gross changes (i.e., a reversal) in a single syllable within a song phrase. The following experiment investigated discrimination of the finer acoustic changes in syllables within a phrase. Repeated syllables within a phrase are not acoustically identical and the quiet intervals between syllables within a phrase are also similar in duration but not identical. This experiment asks whether these kinds of spectral-temporal variations are perceptually salient to canaries. If such spectral-temporal variation is discriminable it could provide a substrate for the birds to encode important biological information in song communication. If not, then one could argue that birds may be listening to these phrases at a
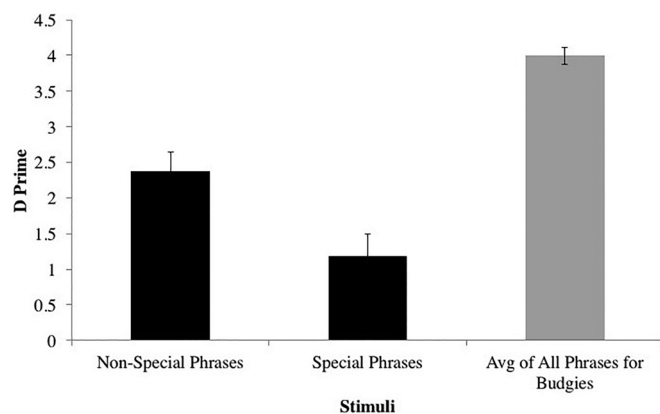

FIG. 5. Canary perception of single syllables reversed at various locations in phrases. Black = Canaries, Gray = Budgerigars.

more global or synthetic level and not attending to the acoustic variability within song phrases.

A. Stimuli

Birds were tested on five pairs of natural and synthetic phrases. As in experiment 2, synthetic phrases were composed of a single syllable extracted from the natural phrase and repeated at a constant rate so that the synthetic phrase matched the duration of the natural phrase. Thus, synthetic phrases lacked the slight variation in syllable spectral content and interval timing that exists within natural phrases, while other acoustic features were kept constant. In all, birds were tested on two special phrase pairs and three non-special phrase pairs (see supplementary material[1]).

B. Procedure

All subjects were tested on the same five stimulus pairs. For the first phase of experiment 3, the natural phrase was played in the background while the synthetic phrase was presented as the target. In phase two, the synthetic phrase served as the background and the natural phrase as the target. For all stimuli, the same sound was presented as each of the seven targets within a block of ten trials. Birds were run until their performance across two successive sessions stabilized, as defined in experiment 2. The second set of 100 trials was used for analysis.

C. Results and discussion

When synthetic phrases were the targets, canaries discriminated these synthetic targets from natural backgrounds at an average d' of 1.11, which is well above chance. Canaries performed even better when the natural phrases were the targets and the synthetic targets the backgrounds, at an average d' of 2.25, and this difference was significant by a paired-samples t-test $[t(5) = -4.30, p = 0.008, d = 1.757]$. Results for the two budgerigars, by contrast, when natural phrases were the background and when they were the targets were not significantly different $[t(1) = -1.03, p = 0.492, d = 0.725]$. The fact that canaries performed better when the natural phrases were the targets suggests the operation of something like the "feature positive effect" (Sainsbury, 1971), where a distinctive feature is more readily discriminated when positively displayed (i.e., a feature added to create the target) than when negatively displayed (i.e., a feature taken away to create the target). If so, this kind of result does suggest that syllable variation may be a biologically relevant acoustic feature for canaries listening to song.

It is interesting that in both phases of this experiment, budgerigars performed significantly better than canaries. For synthetic targets, a mixed model ANOVA with phrase as a within-subjects factor and species as a between-subjects factor showed a significant main effect of species for synthetic targets $[$(Fig. 6); $F(1,6) = 13.15, p = 0.01, \eta^2 = 0.687]$ and natural targets $[F(1,6) = 17.06, p = 0.006, \eta^2 = 0.740]$. Mixed model ANOVAs with phrase as a within-subjects factor and sex as a within-subjects factor showed no sex differences in the performance of the canaries [for synthetic
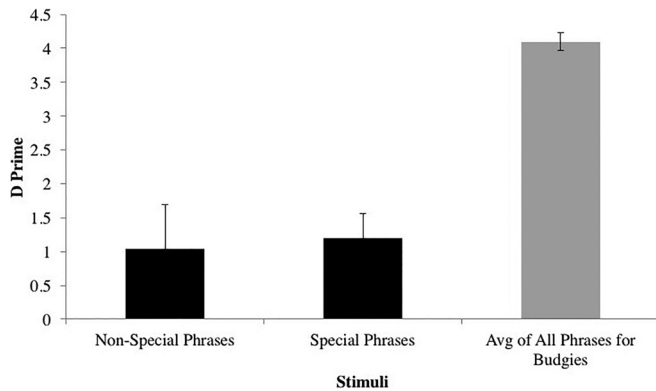
FIG. 6. Canary perception of spectral-temporal syllable variation (synthetic target). Black = Canaries, Gray = Budgerigars.

targets: $F(1,4) = 0.580$, $p = 0.49$, $\eta^2 = 0.127$; for natural targets: $F(1,4) = 0.122$, $p = 0.74$, $\eta^2 = 0.03$]. Unlike in experiment 2, there was no significant effect of the type of phrase (special or non-special) on canary performance [(Fig. 6); for synthetic targets: $t(5) = -0.179$, $p = 0.87$, $d = 0.073$; for natural targets: $t(5) = -0.254$, $p = 0.81$, $d = 0.103$].

Taken together, these results show that subtle spectral-temporal variation among syllables within phrases is discriminable to canaries, which means this syllable variation could code important biological information such as the identity of the singing male. On the other hand, the fact that the canaries were no better at discriminating variation in special phrases than in non-special phrases also suggests that such variation does not play a role in the attractiveness of the phrases. Overall, budgerigars were more sensitive to the syllable spectral-temporal variation than the canaries. This suggests that budgerigars may be listening in a more "analytic" mode while canaries may listen in a more global, phrase by phrase, "synthetic" manner. These species differences in perception are consistent with species differences in song characteristics: canary song is repetitive at the syllable level and variable at the phrase level, while budgerigar warble song is quite variable at the syllable level and has no obvious phrasal patterns (Tu and Dooling, 2012).

VI. EXPERIMENT 4: DISCRIMINATION OF VARIATIONS IN THE TEMPORAL DELIVERY OF SONG SYLLABLES

A striking characteristic of canary phrases is the temporal precision at which syllables are produced. Special phrases are also defined in part by the fast temporal delivery of syllables, and behavioral experiments have shown that female canaries increase CSDs when syllable repetition rate is increased (Draganoiu et al., 2002). Since there have been no psychoacoustic tests of this phenomenon, it is not known whether canaries show a particularly acute sensitivity for discriminating phrases that differ in repetition rate, whether increases or decreases in tempo are more salient, or whether there are sex or species differences in performance.

A. Stimuli

In this experiment, the temporal delivery of all syllables was manipulated for five synthetic phrases. For each

stimulus set, a phrase consisting of repetitions of a single syllable and constant 15 ms inter-syllable intervals (the lower border of special syllable repetition rate according to Vallet and Kreutzer, 1995) was played as the background. The seven target phrases consisted of the same syllable but with varying constant intervals: 0 ms, 5 ms, 10 ms, 20 ms, 25 ms, 30 ms, and 35 ms. The stimuli were constructed so that the total duration of the phrases was held constant. Three of the sets of stimuli contained special syllables, while the other two did not (see supplementary material[1]).

B. Procedure

All birds were tested on the same five stimulus sets where the targets varied in syllable repetition rate and the background phrase contained 15 ms inter-syllable intervals. Birds were run until their performance across two successive sessions stabilized, as defined in experiment 2, and the final 100 trials were used for analysis.

C. Results and discussion

Across all phrases and targets of different intervals, canaries discriminated phrases varying only in syllable repetition rate at an average d' of 1.04. Budgerigars performed significantly better at the same task, with an average d' of 3.07, as tested in a mixed model ANOVA with phrase as a within-subjects factor and species as a between-subjects factor [$F(1,6) = 23.19$, $p = 0.003$, $\eta^2 = 0.794$]. A mixed model ANOVA with phrase as a within-subjects factor and sex as a between-subjects factor showed no sex difference in canaries for average d' across phrases [$F(1,4) = 1.10$, $p = 0.353$, $\eta^2 = 0.216$]. In sum, results show that canaries can discriminate phrases on the basis of syllable repetition rate but that budgerigars are even more sensitive to this song feature than canaries.

Importantly, canaries performed significantly better on special syllables compared to non-special syllables [$t(5) = -2.88$, $p = 0.035$, $d = 1.176$], while budgerigars showed no significant difference between the phrase types [$t(1) = -1.43$, $p = 0.389$, $d = 1.001$]. In addition, there was an interesting asymmetry for canaries in performance on targets that were presented at a faster and slower syllable rate than the background phrase. One would expect that performance on targets equally slower or faster than the background sound (i.e., 10 ms slower vs 10 ms faster) would be the same. In fact, for both special and non-special phrase types, budgerigars showed no significant performance difference on targets slower and faster than the background sound [non-special: $t(3) = -0.068$, $p = 0.95$, $d = 0.034$; special: $t(5) = 1.01$, $p = 0.36$, $d = 0.410$]. Canaries, on the other hand, showed no significant performance difference on targets slower and faster than the background for non-special syllables [$t(11) = 1.37$, $p = 0.20$, $d = 0.396$]. However, for stimuli with special syllables, they performed significantly better on targets faster than the background compared to targets slower [Fig. 7; $t(17) = 8.84$, $p < 0.001$, $d = 2.084$]. Mixed model ANOVAs with speed of the target (faster vs slower than the background) as a within-subjects factor and sex as a between-subjects factor showed no sex differences
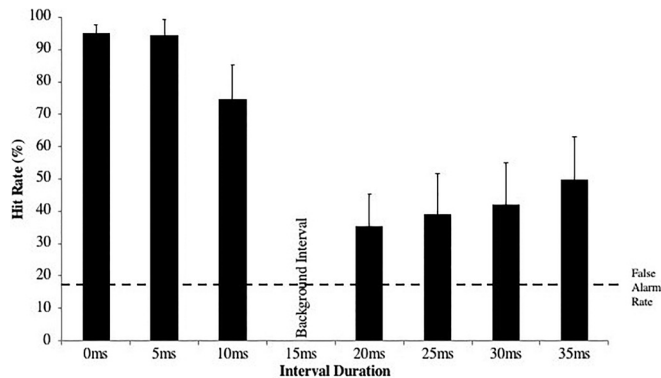
FIG. 7. Canary discrimination of phrases based on the temporal delivery of special song syllables. Birds performed significantly better on targets faster than the background than on slower targets.

in sensitivity to the speed of the target for special syllables $[F(1,16) = 2.02, \ p = 0.174, \ \eta^2 = 0.112]$ or for non-special syllables $[F(1,10) = 0.692, p = 0.425, \eta^2 = 0.065]$.

In sum, canaries show greater sensitivity to increases in the syllable repetition rate for special syllables but not for non-special syllables. This is consistent with the notion that the repetition rate of special syllables is an honest signal for singer quality while the repetition rate of non-special syllables is probably not. The canary perceptual system, then, may be especially tuned to tracking repetition rate when it is processing special syllables.

VII. EXPERIMENT 5A: DISCRIMINATING GAPS BETWEEN NOTES OF SYNTHETIC SPECIAL SYLLABLES

Results from the previous experiment suggested that the short intervals between special syllables may be a particularly salient acoustic feature to canaries. Thus, this experiment examines the temporal offset between notes within a syllable, which is thought to be an important acoustic feature for processing special syllables. Suthers et al. (2012) argued that the suite of traits characterizing special syllables provides an honest signal about the male's ability for bilateral coordination. They found that the two notes are produced sequentially rather than simultaneously, with the left syrinx producing the low-frequency note and the right syrinx the high-frequency note, resulting in an increase in bandwidth that could not be due to just one side of the vocal organ. The authors argue that the temporal offset between the two notes is a necessary cue to indicate that the notes were bilaterally produced, preventing the possibility of "cheating." Here, synthetic special syllables were created with varied gaps between the two notes in order to test directly the sensitivity to temporal offset within special syllable-like sounds.

A. Stimuli

A synthetic special syllable was created in MATLAB based on the third syllable in Fig. 2(A) of Suthers et al. (2012). Each note was 10 ms long. The first note rose from 2300 Hz to 5800 Hz with a 2 ms cosine taper in the beginning and 3 ms cosine taper at the end. The second note fell from 3200 Hz to 1800 Hz with a 2 ms cosine taper in the beginning

and 4 ms cosine taper at the end. Phrases were constructed of 14 synthetic syllables with 15 ms silent intervals between the two-note syllables. The background trill had 1 ms gaps between the two notes (see supplementary material[1]). In each target phrase, the gap was increased at equal steps of either a 1 ms, 2 ms, or 4 ms depending on the performance of the bird (Fig. 8).

B. Procedure

Six canaries (three male and three female) were tested in this experiment. One budgerigar and one zebra finch were also tested. Birds were run until they reached a threshold that was consistent (i.e., within 1 step size) over 200 successive trials (2 consecutive sessions) and where the false alarm rate was no greater than 20%. Performance over these last 200 trials was then averaged to create a psychometric function. The threshold was defined as the gap duration resulting in a hit rate of 50%. Since the background notes had 1 ms gaps, 1 ms was subtracted from the gap duration of the threshold to obtain the true gap detection threshold.

C. Results and discussion

The canary gap detection threshold was 5.18 ± 1.59 ms (standard error of the mean) for gaps between notes of the synthetic special syllables. The budgerigar and zebra finch had a mean threshold of 4.15 ± 1.85 ms. In summary, canaries can detect small changes in the gaps between special syllable-like notes, and could use this capability to identify special syllables and evaluate the degree of bilateral coordination. However, since two other species of birds performed similarly on this task, it is unlikely that special perceptual abilities of canaries are being recruited in processing this acoustic feature.

VIII. EXPERIMENT 5B: DISCRIMINATING NOTE OVERLAP IN SYNTHETIC SPECIAL SYLLABLES

Suthers et al. (2012) also argued for the importance of temporal offset as a cue for special syllables, but close inspection of natural canary song suggests that the notes of special syllables may also sometimes overlap. Since greater overlap may be a signal of poor bilateral coordination, it is also possible that degree of overlap may be used as a cue for male song quality. Changes in overlap could be expected to affect the temporal fine structure of sounds, which birds are
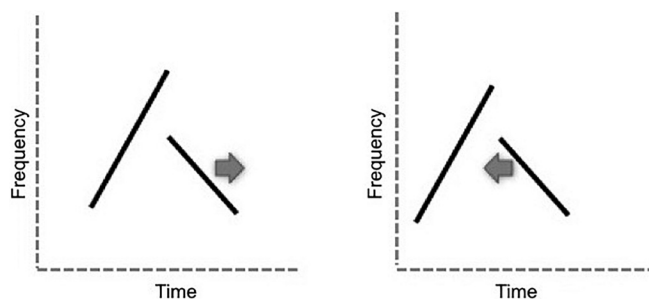

FIG. 8. Schematic of the synthetic special syllables produced for experiment 5a (left) in which the gap between notes was manipulated and for experiment 5b (right) in which the degree of overlap of notes was manipulated.

J. Acoust. Soc. Am. **145** (1), January 2019

Fishbein et al.   571

especially sensitive to, as previously shown (Dooling and Prior, 2017). Here, we varied the degree of overlap for the same synthetic notes used in experiment 5a in order to measure the birds' sensitivity to changes in the overlap of notes in special syllable-like sounds.

A. Stimuli

The same synthetic notes were used as in experiment 5a, but here there was no gap between the notes in the background phrase. The syllables in the target phrases were constructed by decreasing the onset of the second note so that the two notes overlapped (Fig. 8). A range of stimuli was created and the overlap was increased at equal steps of either a half ms, 1 ms, or 2 ms depending on the bird's sensitivity as determined by preliminary testing.

B. Procedure

The procedure was the same as in experiment 5a. Psychometric functions were constructed, and the threshold value was the duration of overlap resulting in a hit rate of 50%.

C. Results and discussion

Canaries discriminated changes in the amount of overlap between notes of the synthetic special syllables at a mean threshold of 2.03 ± 0.42 ms. The budgerigar and zebra finch discriminated the overlap at a mean threshold of 1.35 ± 0.25 ms. Thus, while canaries can discriminate very small changes in the overlap between special syllable-like notes and could use that cue to identify special syllables and evaluate the degree of bilateral coordination, the fact that the budgerigar and zebra finch performed similarly to canaries on this experiment suggests that canaries are probably not relying on special processing abilities in discriminating this acoustic feature. But the fact that the canaries were even better at detecting changes in note overlap than changes in the gap between notes suggests that the acoustic result of this tonal overlap provides a more informationally rich cue for evaluating special syllables. This is slightly at odds with the claim by Suthers et al. (2012) that the temporal offset between the two notes is a necessary cue to indicate that the notes were bilaterally produced. Instead, it is possible that the canaries could be using either cue and the degree of overlap in notes might be more informative.

IX. GENERAL DISCUSSION

Canaries have been a favorite of birdsong aviculturists for a long time and have served as important songbird models for the study of auditory perception and vocal learning (e.g., Waser and Marler, 1977; Nottebohm et al., 1976; Nowicki and Searcy, 2004). Except for behavioral evidence for the attractiveness of special syllables in song, nothing is known about what or how information is conveyed by each syllable and phrase. Our results show that canaries can discriminate the following: the subtle spectral-temporal differences among syllables within phrases, amplitude envelope and fine structure cues at the level of individual syllables,

tempo differences across repeated syllables, and the temporal offset between notes within syllables. Previous studies examining canary song production have shown that variation in these features is important for characterizing differences in song between males and between song bouts of an individual male (Güttinger et al., 1978; Güttinger 1985; Vallet and Kreutzer, 1995; Gardner et al., 2005; Markowitz et al., 2013). Thus, it is likely that all of these features could be used by canaries to convey information in their vocal communication system. The auditory sensitivities measured here also likely shape the song that males learn to produce—i.e., they would not learn to sing repeated syllables at a fairly fixed rate if they could not hear tempo differences across repeated syllables.

Importantly, we found several differences between canaries and budgerigars regarding sensitivities to changes in special syllables. First, canaries perceived more similarity among special phrases from the same bird than did budgerigars. Second, canaries performed worse at discriminating reversed special syllables than at discriminating reversed non-special syllables, an effect that seems due to syllable duration. Third, canaries also performed better at discriminating increases in tempo compared to discriminating decreases in tempo for special syllables—an effect that did not hold when discriminating non-special syllables. Taken together, these results suggest a trade-off in canaries for increased temporal resolution and decreased spectral resolution when perceiving special syllables, which is consistent with studies in other birds showing species-specific trade-offs in temporal and spectral resolution (Henry et al., 2011). Thus, when a canary transitions to its special syllables during song, temporal information may gain salience for both males and females in indicating the quality of a song, as compared to its importance in conveying information at other parts of the song.

All together, these experiments show how the canary perceptual system is suited to receive messages about special syllables. Although experiments 2–5 show no significant sex differences between canaries, the results of experiment 1 indicate that there may still be subtle sex differences in how canaries perceive song phrases. As these birds were kept on short days (non-breeding condition), it is possible that these sex differences are enhanced when birds are in breeding condition (i.e., long days) and higher-order levels of acoustic processing may be involved. In other words, applying the distinction between analytic and synthetic listening (Helmholtz, 1863; Terhardt, 1974), on an analytical level, male and female canaries may be perceiving acoustic features in the same way, but on a synthetic level, male and female canaries may be perceiving song units holistically in subtly different ways. It should be noted that since the sample sizes in these experiments were small, as is typical in psychophysical studies with animals, we could only detect significant results for large effect sizes (Sullivan and Feinn, 2012). With a larger sample, it is possible that smaller effects, such as sex differences in canary performance, could be detected.

It would be interesting to compare these perceptual results from canaries in a non-breeding state to canaries with

elevated hormone levels that are preparing to breed. Recent work on Border strain canaries focused on song production has shown that testosterone differentially affects motor nuclei in the song system along with distinct acoustic aspects of the song (Alward *et al.*, 2013). The present results provide a baseline of performance on the perception of different song features. Estradiol is known to regulate the processing of conspecific vocal signals in songbirds (e.g., Maney and Pinaud, 2011). It is thought to act at multiple levels in the brain and the periphery (Brenowitz and Remage-Healy 2016). There is already evidence for anatomical specificity of the actions of estrogens in the auditory forebrain on specific aspects of song perception as measured by immediate early gene expression (Sanford *et al.*, 2010). Thus, studies of estradiol action at multiple sites in the auditory system might reveal mechanisms of song perception in the brain and periphery.

Our comparisons to the performances of other birds also suggest that canaries listen primarily to global or synthetic, phrase by phrase changes in their song rather than to detailed, syllable by syllable changes. Yet, they seem to shift to listening at an analytical level for perceiving details about the special phrases. Thus, in a way, canaries may listen to their song like we listen to an orchestral symphony, hearing the melody and rhythm of the whole piece, integrating the contributions of each instrument, and not zooming in on the performance of a single instrument except during an especially impressive solo.

X. CONCLUSION

While canaries have long been popular songbird models for the study of vocal learning, little is known about how they listen to their song. In this study, we show that the birds can hear subtle temporal and spectral differences between syllables and notes, but that they may be primarily listening at the more global phrase level. We also demonstrate that perception of special syllables, which females are particularly sexually responsive to, is distinctive for canaries. These results show how the canary auditory system has been shaped to receive messages from conspecific song.

ACKNOWLEDGMENTS

[1]See supplementary materials at https://doi.org/10.1121/1.5087692 for scree plot; for correlations between phrase coordinates and acoustic features; for spectrograms of stimuli for experiment 2; for spectrograms of stimuli for experiment 3; for spectrograms of stimuli for experiment 4; and for spectrograms of stimuli for experiment 5.

Alward, B. A., Balthazart, J., and Ball, G. F. (**2013**). "Differential effects of global versus local testosterone on singing behavior and its underlying neural substrate," Proc. Natl. Acad. Sci. U.S.A. **110**, 19573–19578.

Ball, G. F., and Hulse, S. H. (**1998**). "Birdsong," Am. Psychol. **53**(1), 37–58.

Beecher, M. D., and Brenowitz, E. A. (**2005**). "Functional aspects of song learning in birds," Trends Ecol. Evol. **20**(3), 143–149.

Borg, I., and Groenen, P. J. F. (**2005**). *Modern Multidimensional Scaling: Theory and Applications*, 2nd ed. (Springer Science and Business Media, Inc., New York).

Brenowitz, E. A., and Remage-Healey, L. (**2016**). "It takes a seasoned bird to be a good listener: Communication between the sexes," Curr. Opin. Neurobiol. **38**, 12–17.

Catchpole, C. K., and Slater, P. J. B. (**2008**). *Bird Song: Biological Themes and Variations*, 2nd ed. (Cambridge University Press, Cambridge).

Del Negro, C., and Edeline, J. M. (**2001**). "Differences in auditory and physiological properties of HVc neurons between reproductively active male and female canaries (*Serinus canaria*)," Eu. J. Neurosci. **14**(8), 1377–1389.

Del Negro, C., Kreutzer, M., and Gahr, M. (**2000**). "Sexually stimulating signals of canary (*Serinus canaria*) songs: Evidence for a female-specific auditory representation in the HVc nucleus during the breeding season," Behav. Neurosci. **114**(3), 526–542.

Dooling, R. J., Leek, M. R., Gleich, O., and Dent, M. L. (**2002**). "Auditory temporal resolution in birds: Discrimination of harmonic complexes," J. Acoust. Soc. Am. **112**(2), 748–759.

Dooling, R. J., and Okanoya, K. (**1995**). "The method of constant stimuli in testing auditory sensitivity in small birds," in *Methods in Comparative Psychoacoustics*, edited by G. M. Klump, R. J. Dooling, R. R. Fay, and W. C. Stebbins (Birkhäuser, Basel), pp. 161–169.

Dooling, R. J., Park, T. J., Brown, S. D., and Okanoya, K. (**1987**). "Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulates*): II. Vocal signals," J. Comp. Psychol. **101**, 367–381.

Dooling, R. J., and Prior, N. H. (**2017**). "Do we hear what birds hear in birdsong?," Anim. Behav. **124**, 283–289.

Draganoiu, T. I., Nagle, L., and Kreutzer, M. (**2002**). "Directional female preference for an exaggerated male trait in canary (*Serinus canaria*) song," Proc. R. Soc. London **269**, 2525–2531.

Fee, M. S., and Scharff, C. (**2010**). "The songbird as a model for the generation and learning of complex sequential behaviors," ILAR J. **51**(4), 362–377.

Gardner, T. J., Naef, F., and Nottebohm, F. (**2005**). "Freedom and rules: The acquisition and reprogramming of a bird's learned song," Science **308**, 1046–1049.

Giguére, G. (**2006**). "Collecting and analyzing data in multidimensional scaling experiments: A guide for psychologists using SPSS," Tutorials Quant. Methods Psychol. **2**, 26–37.

Güttinger, H. R. (**1985**). "Consequences of domestication on the song structures in the canary," Behaviour **94**, 254–278.

Güttinger, H. R., Wolffgramm, J., and Thimm, F. (**1978**). "The relationship between species specific song programs and individual learning in songbirds: A study of individual variation in songs of canaries, greenfinches, and hybrids between the two species," Behaviour **65**(1–2), 241–262.

Henry, K. S., Gall, M. D., Bidelman, G. M., and Lucas, J. R. (**2011**). "Songbirds tradeoff auditory frequency resolution and temporal resolution," J. Comp. Psychol. A **197**(4), 351–359.

Hermann, T. (**2002**). "Sonification for exploratory data analysis," Ph.D. thesis, Bielefeld, University, Bielefeld, Germany.

Kisenwether, J. S., and Prosek, R. A. (**2014**). "The effect of experience on perceptual spaces when judging synthesized voice quality: A multidimensional scaling study," J. Voice **28**(5), 548–553.

Kreutzer, M., Beme, I., Vallet, E., and Kiosseva, L. (**1999**). "Social stimulation modulates the use of the 'A' phrase in male canary songs," Behaviour **136**(10), 1325–1334.

Kroodsma, D. E. (**1976**). "Reproductive development in a female songbird: Differential stimulation by quality of male song," Science **192**, 574–575.

Lauer, A. M., Dooling, R. J., Leek, M. R., and Poling, K. (**2007**). "Detection and discrimination of simple and complex sounds by hearing-impaired Belgian Waterslager canaries," J. Acoust. Soc. Am. **122**(6), 3615–3627.

Lehongre, K., Lenouvel, P., Draganoiu, T., and Del Negro, C. (**2006**). "Long-term effect of isolation rearing conditions on songs of an 'open-ended' song learner species, the canary," Anim. Behav. **72**, 1319–1327.

Macmillan, N. A., and Creelman, C. D. (**2005**). *Detection Theory: A User's Guide*, 2nd ed. (Lawrence Erlbaum Associates, Mahwah, NJ).

Maney, D., and Pinaud, R. (**2011**). "Estradiol-dependent modulation of auditory processing and selectivity in songbirds," Front. Neuroendocrinol. **32**, 287–302.

Markowitz, J. E., Ivie, E., Kligler, L., and Gardner, T. J. (**2013**). "Long-range order in canary song," PLoS Comput. Biol. **9**(5), e1003052.

Marler, P. R. (**2004**). "Science and birdsong: The good old days," in *Nature's Music: The Science of Birdsong*, edited by P. R. Marler and H. Slabbekoorn (Elsevier, San Diego, CA), pp. 1–37.

Neilans, E. G., Holfoth, D. P., Radziwon, K. E., Portfors, C. V., and Dent, M. L. (**2014**). "Discrimination of ultrasonic vocalizations by CBA/CaJ mice (*Mus musculus*) is related to spectrotemporal dissimilarity of vocalizations," PloS One **9**(1), e85405.

Nottebohm, F., Stokes, T. M., and Leonard, C. M. (**1976**). "Central control of song in the canary, *Serinus canarius*," J. Comp. Neurol. **165**: 457–486.

Nowicki, S., and Searcy, W. A. (**2004**). "Song function and the evolution of female preferences: Why birds sing, why brains matter," Ann. N. Y. Acad. Sci. **1016**, 704–723.

Okanoya, K., and Dooling, R. J. (**1987a**). "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. (**1987b**). "Strain differences in auditory thresholds in the canary (*Serinus canarius*)," J. Comp. Psychol. **101**(2), 213–215.

Okanoya, K., Dooling, R. J., and Downing, J. D. (**1990**). "Hearing and vocalizations in hybrid Waterslager-Roller canaries (*Serinus canarius*)," Hearing Res. **46**(3), 271–275.

Parisot, M., Vallet, E., Nagle, L., and Kreutzer, M. (**2002**). "Male canaries discriminate among songs: Call rate is a reliable measure," Behaviour **139**(1), 55–63.

Pasteau, M., Nagle, L., and Kreutzer, M. (**2004**). "Preferences and predispositions for intra-syllabic diversity in female canaries (*Serinus canaria*)," Behaviour **141**, 571–583.

Pasteau, M., Nagle, L., and Kreutzer, M. (**2007**). "Influences of learning and predispositions on frequency fevel preferences on female canaries (*Serinus canaria*)," Behaviour **144**, 1103–1118.

Pohl, N. U., Klump, G. M., and Langemann, U. (**2015**). "Effects of signal features and background noise on distance cue discrimination by a songbird," J. Exp. Biol. **218**(7), 1006–1015.

Prior, N. H., Smith, E., Lawson, S., Ball, G. F., and Dooling, R. J. (**2018**). "Acoustic fine structure may encode biologically relevant information for zebra finches," Sci. Rep. **8**, 6212.

Sainsbury, R. S. (**1971**). "Effect of proximity of elements on the feature-positive effect," J. Exp. Anal. Behav. **16**, 315–325.

Sanford, S. E., Lange, H. S., and Maney, D. L. (**2010**). "Topography of estradiol modulated genomic responses in the songbird auditory forebrain," Dev. Neurobiol. **70**(2), 73–86.

Sullivan, G. M., and Feinn, R. (**2012**). "Using effect size—Or why the *p* value is not enough," J. Grad. Med. Educ. **4**(3), 279–282.

Supper, A., and Bijsterveld, K. (**2015**). "Sounds convincing: Modes of listening and sonic skills in knowledge making," Interdiscip. Sci. Rev. **40**(2), 124–144.

Suthers, R., Vallet, E., and Kreutzer, M. (**2012**). "Bilateral coordination and the motor basis of female preference for sexual signals in canary song," J. Exp. Biol. **215**, 2950–2959.

Terhardt, E. (**1974**). "Pitch, consonance, and harmony," J. Acoust. Soc. Am. **55**, 1061–1069.

Terleph, T. A., Mello, C. V., and Vicario, D. S. (**2007**). "Species differences in auditory processing dynamics in songbird auditory telencephalon," Dev. Neurobiol. **67**(11), 1498–1510.

Tsoumani, O., and Postma-Nilsenova, M. (**2013**). "Perceiving sounds: Analytic and synthetic listening, global-local processing and possible links with empathy and self-construal," Proc. Annu. Meet. Cognit. Sci. Soc. **35**, 3587–3592.

Tu, H-W., and Dooling, R. J. (**2012**). "Perception of warble song in budgerigars (*Melopsittacus undulates*): Evidence for special processing," Anim. Cognit. **15**, 1151–1159.

Vallet, E., Beme, I., and Kreutzer, M. (**1998**). "Two-note syllables in canary songs elicit high levels of sexual display," Anim. Behav. **55**, 291–297.

Vallet, E., and Kreutzer, M. (**1995**). "Female canaries are sexually responsive to special song phrases," Anim. Behav. **49**, 1603–1610.

Vernaleo, B. A., and Dooling, R. J. (**2011**). "Relative salience of envelope and fine structure cues in zebra finch song," J. Acoust. Soc. Am. **129**, 3373–3383.

von Helmholtz, H. (**1863**). *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, 3rd English ed. (Longmans, Green, and Co., New York), translated by A. J. Ellis, 1895.

Waser, M. S., and Marler, P. (**1977**). "Song learning in canaries," J. Comp. Physiol. Psychol. **91**(1), 1–7.

Yang, J., and Fox, R. A. (**2014**). "Perception of English vowels by bilingual Chinese-English and corresponding monolingual listeners," Lang. Speech **57**, 215–237.