Journal of Comparative Psychology

Zebra Finches (Taeniopygia guttata) Can Categorize Vowel-Like Sounds on Both the Fundamental Frequency ("Pitch") and Spectral Envelope

Merel A. Burgering, Jean Vroomen, and Carel ten Cate

Online First Publication, September 20, 2018. http://dx.doi.org/10.1037/com0000143

CITATION

Burgering, M. A., Vroomen, J., & ten Cate, C. (2018, September 20). Zebra Finches (Taeniopygia guttata) Can Categorize Vowel-Like Sounds on Both the Fundamental Frequency ("Pitch") and Spectral Envelope. *Journal of Comparative Psychology*. Advance online publication. http://dx.doi.org/10.1037/com0000143

American Psychological Association
Journal of Comparative Psychology

0735-7036/18/$12.00
http://dx.doi.org/10.1037/com0000143

Zebra Finches (*Taeniopygia guttata*) Can Categorize Vowel-Like Sounds on Both the Fundamental Frequency ("Pitch") and Spectral Envelope

Merel A. Burgering
Tilburg University and Leiden University

Jean Vroomen
Tilburg University

Carel ten Cate
Leiden University

Humans can categorize vowels based on spectral quality (vowel identity) or pitch (speaker sex). Songbirds show similarities to humans with respect to speech sound discrimination and categorization, but it is unclear whether they can categorize harmonically structured vowel-like sounds on either spectrum or pitch, while ignoring the other parameter. We trained zebra finches in two experimental conditions to discriminate two sets of harmonic vowel-like sounds that could be distinguished either by spectrum or fundamental frequency (pitch). After the birds reached learning criterion, they were tested on new sounds that were either noise-vocoded versions of the trained sounds (sharing the spectral envelope with the trained sounds but lacking fine spectral detail from which pitch could be extracted) or sounds lacking the amplified harmonics (sharing only pitch with the trained sounds). Zebra finches showed no difference in the number of trials needed to learn each stimulus–response mapping. Birds trained on harmonic spectrum generalized their discrimination to vocoded sounds, and birds trained on pitch generalized their discrimination to harmonic sounds with a flat spectrum. These results demonstrate that, depending on the training requirements, birds can extract either fundamental frequency or spectral envelope of vowel-like sounds and use these parameters to categorize new sounds.

Keywords: pitch perception, categorization, harmonic sounds, songbird, spectral envelope

Supplemental materials: http://dx.doi.org/10.1037/com0000143.supp

Driven by an interest in speech sound perception and its evolution, multiple studies investigated to what extent nonhuman animals (hereafter: animals) are able to discriminate and categorize speech sounds. In particular, the discrimination and/or categorization of vowel-like sounds has been examined and shown for a range of species, such as ferrets (Bizley, Walker, King, & Schnupp, 2013; Town, Atilgan, Wood, & Bizley, 2015), cats (Dewson, 1964), budgerigars (Dooling & Brown, 1990), European starlings (Kluender, Lotto, Holt, & Bloedel, 1998), and zebra finches (Kriengwatana, Escudero, Kerkhoven, & ten Cate, 2015; Kriengwatana, Escudero, & ten Cate, 2015; Ohms, Escudero, Lammers, & ten Cate, 2012; Ohms, Gill, Van Heijningen, Beckers, & ten Cate, 2010).

Essential for speech perception is the ability to categorize speech sounds (Goudbeek, Swingley, & Kluender, 2007; Goudbeek, Swingley, & Smits, 2009; Holt & Lotto, 2006, 2010). Humans can categorize the same speech sound on different dimensions, for example, vowel and speaker identity or speaker sex. Speaker sex categorization is mostly driven by sex differences in pitch, which are related to differences in fundamental frequency (f0; the human analogue of pitch) and its harmonic spectrum. In contrast, vowel categorization is driven by differences in timbre, called "formants" for speech sounds, characterized by different amplitude peaks in the harmonic spectrum (Fuller et al., 2014; Holt & Lotto, 2010). Songbirds are also able to discriminate between speech sounds (Kluender et al., 1998; Ohms et al., 2010, 2012) and speakers (Dooling, 1992), suggesting they can identify the most important parameters that humans use, that is, pitch and spectral envelope, to categorize vowel-like sounds.

Despite these similarities, pitch processing is often assumed to differ between humans and birds. In tone sequences and music, humans are known to perceive relative pitch: We attend to the relationship or intervals between successive sound elements to recognize and discriminate sound sequences (Bregman, Patel, & Gentner, 2012). This ability has also been shown by several mammal species, including Japanese (Izumi, 2001) and macaque

---

Merel A. Burgering, Tilburg School of Social and Behavioral Sciences, Department of Cognitive Neuropsychology, Tilburg University, and Institute of Biology Leiden, Leiden University; Jean Vroomen, Tilburg School of Social and Behavioral Sciences, Department of Cognitive Neuropsychology, Tilburg University; Carel ten Cate, Institute of Biology Leiden and Leiden Institute for Brain and Cognition, Leiden University.

This research was supported by Gravitation Grant 024.001.006 of the Language in Interaction Consortium from Netherlands Organization for Scientific Research.

Correspondence concerning this article should be addressed to Merel A. Burgering, Tilburg School of Social and Behavioral Sciences, Department of Cognitive Neuropsychology, Tilburg University, Warandelaan 2, P.O. Box 90153, 5000 LE Tilburg, the Netherlands. E-mail: m.a.burgering@tilburguniversity.edu

monkeys (Brosch, Selezneva, Bucks, & Scheich, 2004) and ferrets (Walker, Schnupp, Hart-Schnupp, King, & Bizley, 2009; Yin, Fritz, & Shamma, 2010). In contrast, songbirds, including starlings (Hulse & Cynx, 1985; Hulse, Cynx, & Humpal, 1984; MacDougall-Shackleton & Hulse, 1996; Page, Hulse, & Cynx, 1989), mockingbirds and cowbirds (Hulse & Cynx, 1985), white-throated sparrows (Hurly, Ratcliffe, & Weisman, 1990), black-capped chickadees (Weary & Weisman, 1991), and zebra finches (Weisman, Njegovan, & Ito, 1994; Weisman et al., 1998), seem to attend primarily to the absolute pitch of sound stimuli to make their perceptual decisions. Absolute pitch is the capacity to distinguish different pitches without an external referent (Friedrich, Zentall, & Weisman, 2007). After training on tone sequences, songbirds can eventually learn to generalize for relative pitch (Cynx, 1995; Page et al., 1989). However, the birds do not transfer their strategies to sequences outside of the training frequency range, and they need the same amount of trials to learn the new discrimination as they need to learn the first discrimination (Cynx, 1995; Hulse & Cynx, 1985). A cross-species mammal comparison (between rats and humans) led Weisman to hypothesize that there is general difference in processing of absolute and relative pitch between mammals (including humans, rats, and ferrets) and songbirds (Weisman, Njegovan, Williams, Cohen, & Sturdy, 2004).

However, a recent study showed that starlings are able to categorize pitch-shifted conspecific songs, including songs that were shifted outside the frequency range of the trained songs (Bregman et al., 2012), although they could not do this for piano tone sequences. Starling songs are spectrotemporally complex, and Bregman et al. (2012) suggested that the observed generalization across frequency-shifted songs reflects the birds' ability to detect spectrotemporal changes over time independent of absolute frequency. Bregman, Patel, and Gentner (2016) subsequently investigated how starlings perceive tone sequences that varied over time both in pitch and spectral timbre. With these experiments, they showed that the birds are able to use spectral shape information, the so-called "spectral envelope," that is, the overall pattern of amplitude across frequency bands. To what extent this is a general feature for songbirds or specific for starlings, is still unknown.

The study by Bregman et al. (2016) also provides a novel perspective on the perception of vowel spectra by birds. Do birds discriminate vowel-like sounds by relying on fine spectral details, or do they rely on differences in the spectral envelope? In the current study, we further examined the acoustic parameters that zebra finches can use to discriminate and categorize artificial sounds differing acoustically in fundamental frequency (hereafter: pitch, which is related to speaker sex differences) and in formant-like variation in spectral envelope. Zebra finches are known to be skilled at discriminating absolute frequencies (Cynx, 1995; Lee, Charrier, Bloomfield, Weisman, & Sturdy, 2006), differences in harmonic spectra (Lohr & Dooling, 1998; Uno, Maekawa, & Kaneko, 1997; Vignal & Mathevon, 2011), and natural and artificial vowels (Kriengwatana, Escudero, Kerkhoven, et al., 2015; Ohms et al., 2010, 2012), as well as known to be very sensitive to fine temporal structure of sounds (Dooling, 1992; Dooling & Prior, 2017; Lohr, Dooling, & Bartone, 2006). Unknown, however, is whether they can use the spectral envelope of harmonically structured sounds, which lacks information about pitch and fine spectral details, to discriminate sounds. Also, although zebra finches may have a natural tendency to be more sensitive to some parameters, for example, temporal fine structure, than to others, for example, spectral envelope (Dooling & Prior, 2017), some experiments have suggested that they can learn to use both speaker sex and vowel quality for categorization (Burgering, ten Cate, & Vroomen, 2018). Here, we studied in a systematic way whether zebra finches are perceptually sensitive to the same parameters, pitch and spectral envelope, which humans use for vowel categorization and for speaker sex categorization. Birds were trained to classify six sounds into two categories using a two-alternative forced choice paradigm with corrective feedback. In one condition, the birds had to discriminate artificial sounds based on differences in pitch while variation in spectral envelope was irrelevant to obtain reinforcement. In the other condition, the birds had to discriminate sounds based on differences in spectral envelope while variation in pitch was irrelevant. By differential reinforcement, the birds could potentially learn to focus on one parameter (e.g., pitch) and to ignore the other (e.g., spectral envelope) to succeed in correct classification.

In the test phase, we examined whether the birds generalized the learned discrimination to new sounds using the relevant parameter irrespective of the value of the other (irrelevant) parameter. One set of test sounds had a flat spectrum, but contained information about f0 in the harmonics (pitch). The other test set consisted of so-called noise-vocoded (hereafter: vocoded) versions of the trained sounds. The vocoding technique replaces the detailed harmonic spectrum with noise distributed over a limited set of frequency bands, thus removing harmonic detail from which fundamental frequency or pitch can be derived, but maintaining the gross spectral envelope (Dorman, Loizou, & Rainey, 1997; Roberts, Summers, & Bailey, 2011).

If the zebra finches were simply memorizing the individual trained sounds (exemplars), we expected them to show limited generalization to the new sounds. On the other hand, if the finches extracted the relevant cue (pitch or spectral envelope) from the training sounds, we expected them to generalize to those new test sounds that share that cue.

Method

Subjects

We used 14 adult zebra finches (*Taeniopygia guttata*; eight males and six females) from the Leiden University breeding colony. At the start of the experiment, all birds were between 136 days and 672 days post hatching (398 ± 151 days). Before the experiment, birds were housed in single-sex groups with a maximum of 15 animals, and they were kept on a 13.5 L: 10.5 D schedule at 20–22°C. Birds always had free access to water and feeders with seeds. Furthermore, they received twice a week some vegetables and fruit (grated carrot and apple) and egg food (mashed boiled eggs). All zebra finches were physically examined and weighted before the start of the experiment. Throughout the experiment, the birds were carefully observed. During the experiment, drinking water, cuttlebone, and grit were available ad libitum. Food was used as reinforcement and therefore only available after a correct response. If, for some reason, a bird was not able to obtain food for 18 consecutive hours, the food hatch opened automatically for unlimited time. None of the birds had experience with similar behavioral testing. All animal procedures were approved by the Leiden Committee for animal experimentation

(DEC (dierexperimentencommissie (in Dutch: animal experimentation committee) number 14,229).

Apparatus

Zebra finches were individually housed in an operant conditioning cage (Skinnerbox) in a sound-attenuated room (Figure 1, adapted from Burgering et al., 2018). Three horizontally aligned pecking sensors in the back wall of the cage, a fluorescent lamp, a food hatch, and a speaker were connected to an operant conditioning controller that registered all sensor pecks. Pecking the middle sensor elicited a sound and illuminated the LED light of the left and right sensor. Pecking the left sensor after hearing a sound from Category A and pecking the right sensor after hearing a sound from Category B resulted in food access for 8–10 s. An incorrect response led to 1–12 s darkness depending on the experimental phase. The intertrial interval was 2 s. The cage was built from wire mesh walls and one foamed PVC back wall. A fluorescent lamp (Phillips Master TL-D 90 DeLuxe 18W/965, The Netherlands) was placed on top of the Skinnerbox and functioned as the only light source. The same light/dark schedule as in the breeding colony was applied. Sound stimuli were played at ~70 dB (SPL meter, Rion NL 15; Rion, Kokubunji, Tokyo, Japan) through a speaker (Vifa MG10SD09–08; Vifa, Viborg, Denmark) 1 m above the cage.

Stimulus Material

Training sounds. We created two stimulus sets of artificial harmonic sounds that varied with respect to fundamental frequency and spectrum (i.e., the relative amplitude of the harmonics; Figures 2 and 3). Stimulus creation started with flat-spectrum source sounds (i.e., without amplifications of a specific set of harmonics) differing in fundamental frequency and number of harmonics using a PRAAT script from Shigeto Kawahara (http://user.keio.ac.jp/~kawahara/). All sounds had a duration of 0.8 s and were ramped on and off with 0.2 s cosine ramps. Per stimulus set, three different f0 values were chosen (100, 165, and 230 Hz for Set 1; 105, 170, and 235 Hz for Set 2; see Figure 2), based on natural f0 values of male and female speakers (Adank, van Hout, & Smits, 2004). The maximum frequency was set to 6000 Hz, and the number of harmonics of the source sounds depended on f0. For stimulus Set 1, the flat spectrum source sound with an f0 of 100 Hz contained 60 harmonics, the f0 of 165 Hz contained 36 harmonics, and the source sound with f0 of 230 Hz contained 26 harmonics. For stimulus Set 2, the source sounds with an f0 of 105 Hz contained 57 harmonics, an f0 of 170 Hz contained 35 harmonics, and an f0 of 235 Hz contained 26 harmonics. All formant-like frequency bands had a bandwidth of 190 Hz.

To create the training stimuli, for each of the source sounds three spectral variants were created by amplifying a different frequency band (with 30–38 dB) centered around either 1450,

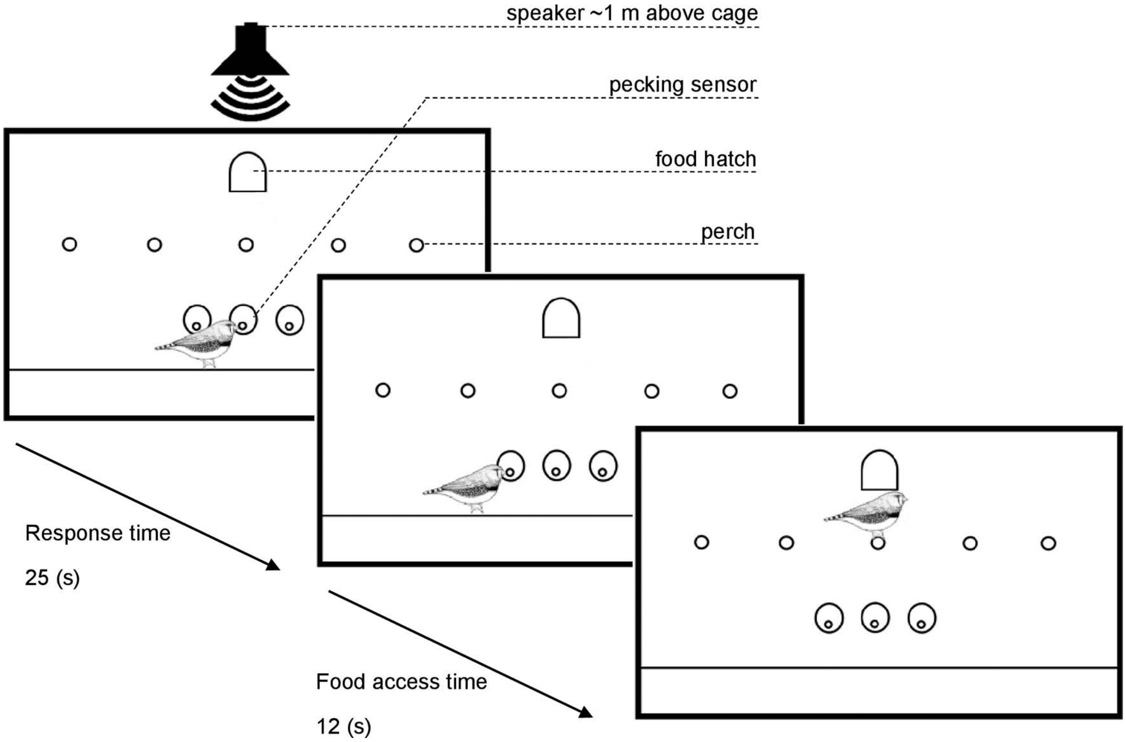


Figure 1. The operant conditioning chamber is 70 (l) × 30 (d) × 45 (h) cm. Three horizontally aligned pecking sensors are depicted at the bottom just above the long horizontal perch. In the first display, the zebra finch is pecking the middle sensor, which causes a sound to be played from the speaker above. The bird has 25 s to peck the left (or right) sensor (depicted in the second display). In the last display, the bird is sitting in front of the open food hatch (adapted with permission from Burgering, ten Cate, & Vroomen, 2018).

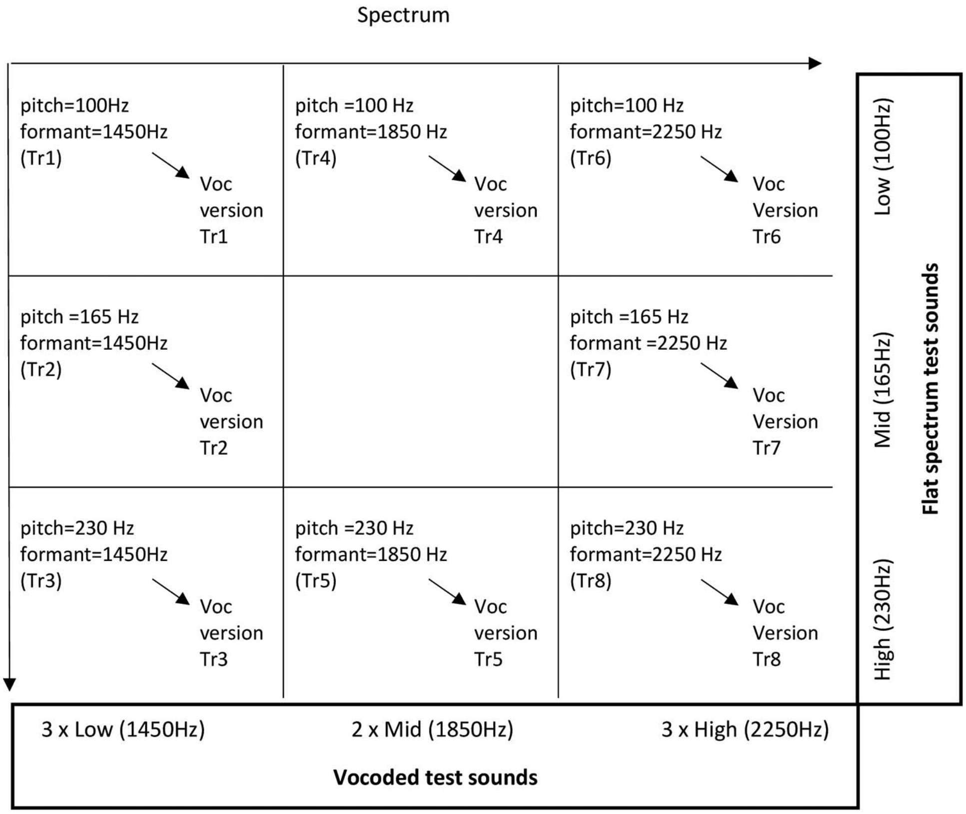


Figure 2. All birds were trained to categorize six sounds either based on spectrum (Tr1, Tr2, and Tr3 vs. Tr6, Tr7, and Tr8) or based on pitch (Tr1, Tr4, and Tr6 vs. Tr3, Tr5, and Tr8). The pitch values and formant values for Set 1 are given in this figure. The values for Set 2 were +5 Hz for all pitch values and −10 Hz for all formant values. Flat spectrum test sounds are depicted in the box on the right, and each vocoded test sound is depicted in the spectrum/pitch matrix. The vocoded test sounds were clustered in a low, mid, and high group, depicted in the box at the bottom.

1850, or 2250 Hz for Set 1; and 1440, 1840, or 2240 Hz for Set 2 using Adobe Audition (see Figure 2). Our amplified frequency bands resemble the second formant (F2) in natural vowels. For example, for Dutch vowels, frequencies for F2 are roughly between 800 and 2400 Hz (Adank et al., 2004). The overall intensity of all training sounds had same overall amplitude of 71 dB.

Test sounds. To create the vocoded test sounds, the training sounds were noise vocoded into 16 log-spaced frequency bands over the frequency range of 100 Hz to 6 kHz using a script from Matthew Winn (http://www.mattwinn.com/praat.html). To the human ear, noise vocoding distorts the fine spectral detail from which the pitch is derived, but the gross spectral envelope remains intact. For the test sounds with a flat spectrum, we used the original source sounds on which the training sound were based. Note that these source sounds had fine spectral details from which pitch could be derived, but their spectral envelope was flat (Figures 2 and 3).

Training phase. We used two different six-to-two stimulus–response (SR) mappings called “pitch” training and “spectrum” training. Six sounds could be categorized in training based on pitch (Tr1, Tr4, and Tr6 vs. Tr3, Tr5, andTr8), and six sounds could be categorized in training based on spectrum (Tr1, Tr2, and Tr3 vs. Tr6, Tr7, and Tr8).

Transition phase. In this phase, we presented birds with 12 amplitude-modulated versions of the trained sounds (±2 dB from original training sound). By doing so, we could exclude the possibility that birds were categorizing the sounds solely based on loudness. These sounds were not reinforced in 20% of the trials.

Test phase. This stage consisted of two stages. The birds either received the eight vocoded test sounds and next the three flat spectrum sounds, or the other way around.

Design

The birds were randomly assigned to one of the different SR mappings (SR mapping was between subjects). Every SR mapping was completed by seven birds. Per SR mapping, three birds were tested with stimulus Set 1 and four birds were tested with stimulus Set 2.

Procedure

All birds were trained to categorize six training sounds into two categories. Each experiment consisted of a shaping, training, transition, and test phase.

During the **shaping phase**, the bird had time to acclimate to the cage, learn where to find food, and learn how to use the pecking

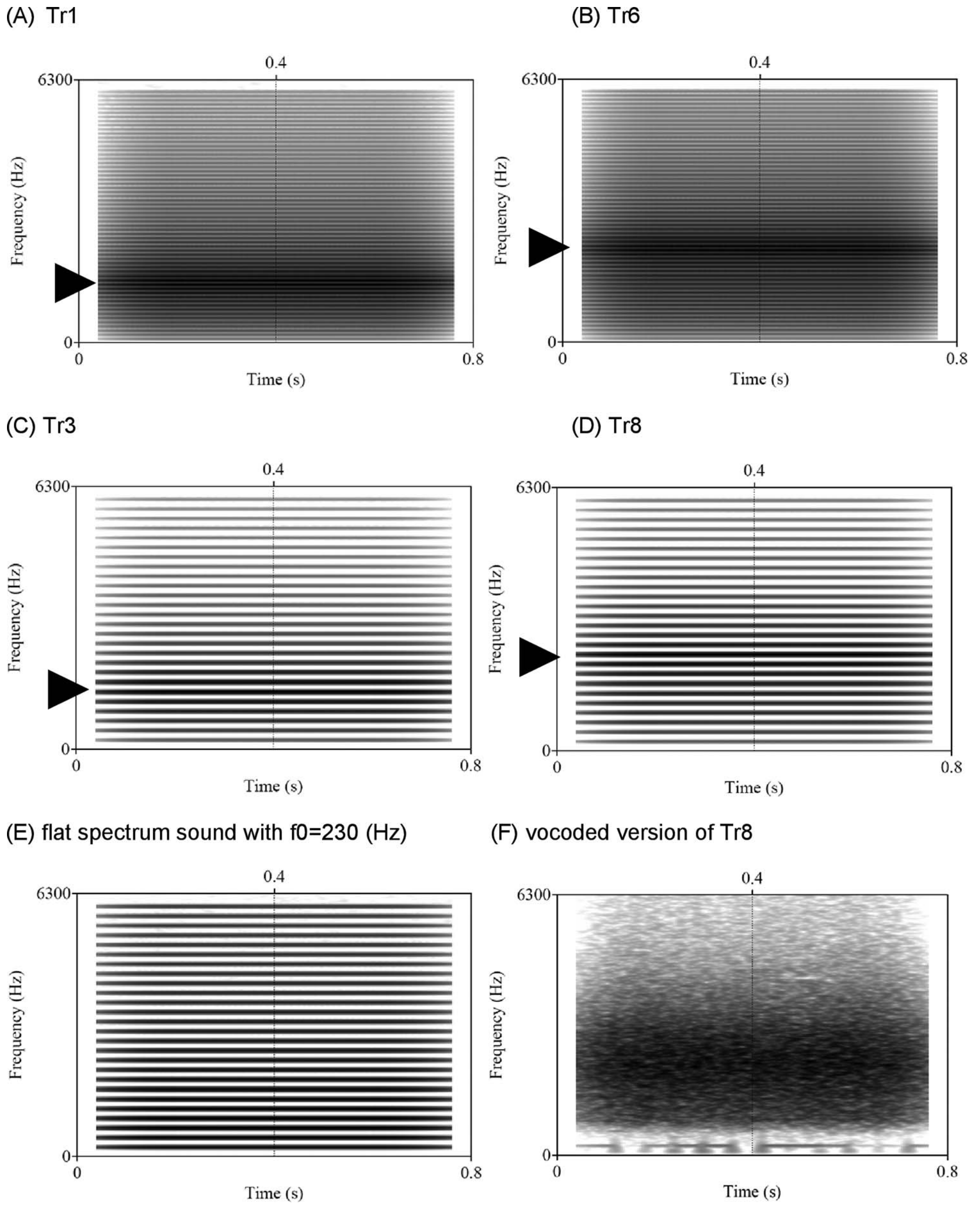
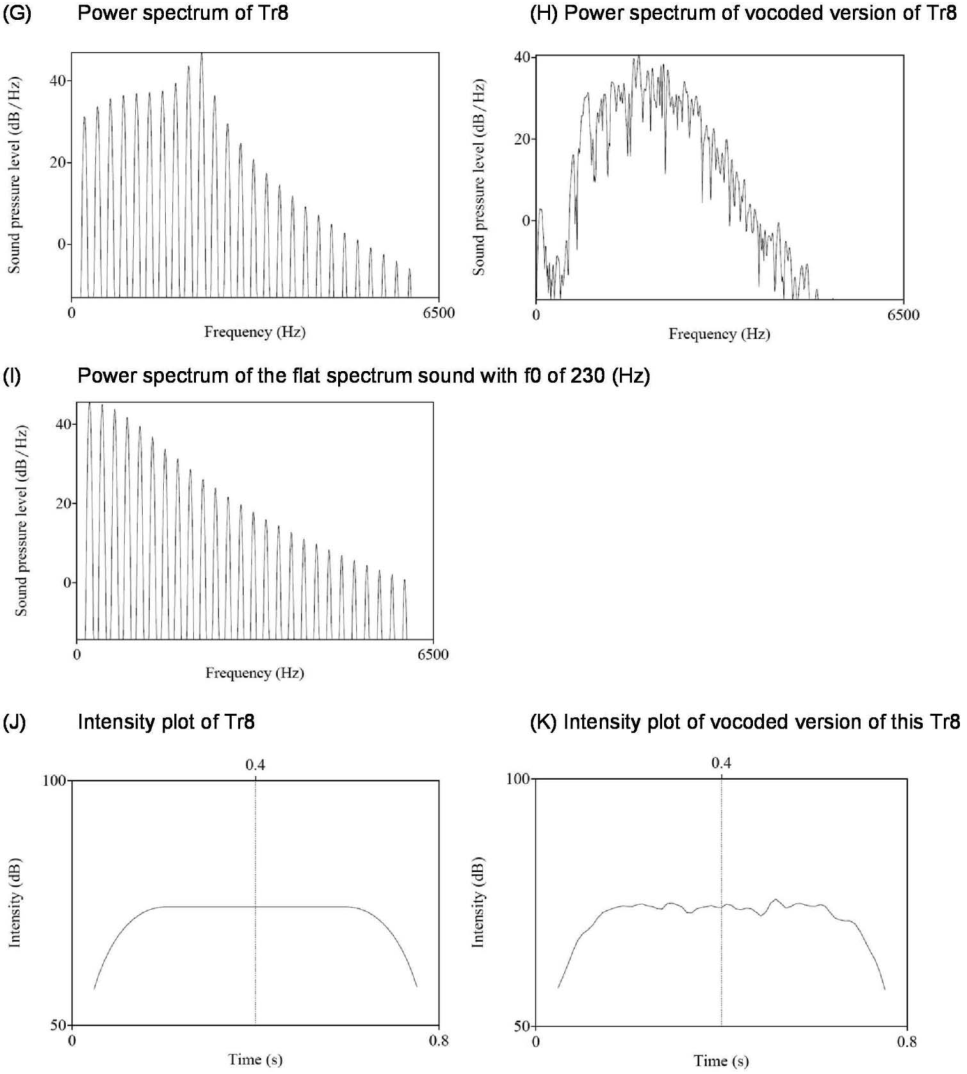


Figure 3. Spectrograms for the training sounds ([A] Tr1; [B] Tr6; [C] Tr3; [D] Tr8) for pitch training from Set 1 and two test sounds ([E] flat spectrum sound with pitch of 230 Hz; [F] vocoded version of Tr8). Power spectra for Tr8 (G), the vocoded version of Tr8 (H), and the flat spectrum sound with f0 = 230 Hz (I). Intensity plots for Tr8 (J) and the vocoded version of Tr8 (K).

sensors. The bird had to learn to initiate its own trial by pecking the middle sensor first, which elicited one of the two unfamiliar zebra finch songs (A and B). Next, the bird had to respond to the played sound by pecking the left or right sensor (see Burgering et al., 2018, for further details). The birds' responses were differentially reinforced, for example, when Song A was played, pecking the left sensor resulted in food access whereas pecking the right sensor resulted in a preset time of darkness and vice versa for Song B. For each day, the discrimination between the stimuli by each bird was calculated as the proportion of correct responses out of all sounds that birds responded to. After 3 days performing at $>$0.75 overall and $>$0.60 per category, the bird was promoted to the **training phase**, during which the bird was trained on six training sounds according to the type of training, either pitch or spectrum.

When a bird had an overall discrimination score of $>$0.75 and a score of $>$0.60 for each category for 3 consecutive days, the bird was promoted to the **transition phase**, during which the six training stimuli and the 12 amplitude-modulated versions of these stimuli were not reinforced in 20% of the trials of 1 day. The bird had to maintain its discrimination score for another day. This transition phase prepared the bird for the **test phase**. During the two stages of the test phase, 11 new test sounds were introduced:


(G) Power spectrum of Tr8

(H) Power spectrum of vocoded version of Tr8

(I) Power spectrum of the flat spectrum sound with f0 of 230 (Hz)

(J) Intensity plot of Tr8

(K) Intensity plot of vocoded version of this Tr8

Figure 3. (continued)

three flat spectrum sounds and eight vocoded versions of all trained sounds. Test sounds were randomly played interspersed between training sounds, and the test sounds were never reinforced. During each stage of the test phase, 20% of all trials were new sounds and 80% of all trials were trained sounds. Each test sound was presented 40 times. Therefore, birds were presented with 320 vocoded test sounds and 120 flat spectrum test sounds, respectively. When both stages of the test phase were completed, the experiment was finished and the bird was returned to the aviary.

Analyses

Training phase. We measured the number of training trials needed before the overall proportion correct was >0.75 and the discrimination for both left and right was >0.60 on 3 consecutive days. Due to our small sample size, we decided to treat the data as nonparametric, and we submitted the data to a Mann–Whitney test to examine whether training type had a significant effect on number of trials required to reach criterion. Due to the design, birds had the possibility to withhold their responses, here defined as no-responses. Therefore, we also calculated the percentage of “no responses” for all birds during the entire training phase.

Test phase. We calculated the percentage of no responses for trained and test sounds per test stage. If this percentage for each test sound was higher than 25%, we excluded the data points of that individual from the analysis. On the basis of this criterion, we excluded one bird (ID 675), which had >75% of no responses to each test sound. For all trials that got a response, the proportion of responses for each sound type was calculated as proportions of responses to the sensor for category low pitch (for the birds trained on the pitch mapping) or low spectrum (for the birds trained on the spectrum mapping).

If the birds passed the training phase, they already demonstrated to be able to categorize the trained sounds. Therefore, we combined the proportion of responses into one score for the three trained sounds within each category (pitch_low vs. pitch_high and spectrum_low vs. spectrum_high). Next, we tested whether the proportions of responses for the three vocoded versions of the trained sounds with a low frequency (voc_low) were significantly different. If not, we combined the proportion of responses into one score for these three vocoded sounds. We did the same for the two vocoded versions with a middle frequency (voc_mid) and the three vocoded versions with a high frequency (voc_high). To test whether the proportions of responses for the five different sound types were significantly different, we submitted these proportions for each test phase to a separate Friedman test with sound type (two trained sound types and three flat spectrum sound or three vocoded sound types) as a fixed effect. When the main analysis revealed significant effects, three post hoc pairwise comparisons between the proportions of responses for the different test sound types were performed using a Wilcoxon signed-ranks test. These reported *p* values are uncorrected. All analyses were performed with IBM SPSS Statistics 22 (Armonk, New York, United States of America).

Results

Training Phase

For both the spectrum and pitch mapping, seven different individual birds completed the training phase. The average number of trials with standard deviation are reported. To reach criterion, birds required on average 1,801 ± 619 trials in the spectrum training and 2,858 ± 2,063 trials in the pitch training (Figure 4; for individual data, see Table S1 in the online supplemental materials). To compare the number of trials needed to reach criterion in the two experimental conditions, we ran a Mann–Whitney test, which showed that the number of trials was not significantly different between the spectrum (median = 1,651) and pitch training (median = 1,851; U = 20 and $p = .620$). The average of no responses was 6.3% for birds assigned to the spectrum mapping and 7.7% for birds assigned to pitch mapping.

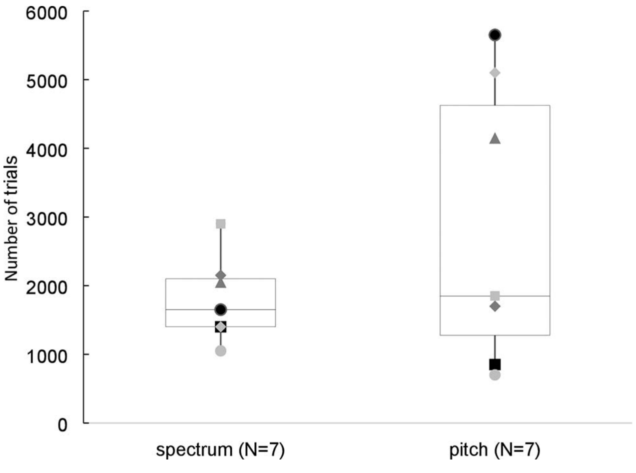


Figure 4. The number of trials for two experimental conditions: spectrum training (left) and pitch training (right). Individual data points are indicated with different black and gray symbols. Boxplots display median, interquartile range, and full range.

Performance on Test Trials After Spectrum Training

Figure 5 displays boxplots with the median, interquartile range, and full range of the proportion of responses to the sensor associated with the low spectrum sounds for the test phase with flat spectrum test sounds (Figure 5A) and the vocoded test sounds (Figure 5B; for individual data, see Tables S2a and S2b in the online supplemental materials). The birds had on average 208 ± 53 trials per day. We found no significant differences in performance of birds tested with Set 1 and those tested with Set 2.

The percentage of no responses in the test on flat spectrum sounds was 5.8% for the trained test sounds and 18.3% for the new test sounds (see Figure S1 in the online supplemental materials). To examine whether sound type had an effect on proportion of responses, a Friedman test was performed which rendered a chi-square value of 23.200 ($p < .001$), which showed that the effect of sound type on the proportion of responses to low spectrum sensor was significant. Post hoc Wilcoxon's tests showed that there was no significant difference in proportion of responses to the flat spectrum test sound with a low pitch and those to the flat spectrum test sound with a high pitch ($p = .237$) and those of a middle pitch ($p = .398$). The proportions of responses to the flat spectrum test sound with a middle pitch were not significantly different from those to the flat spectrum test sounds with a high pitch ($p = .499$).

The responses for the vocoded versions with different f0 but the roughly same frequency band (three low, two middle, and three high frequency) were not significantly different from each other. Therefore, we calculated average proportion of responses for these sound types (vocoded low-frequency sounds [voc_low], vocoded middle-frequency sounds [voc_mid], vocoded high-frequency sounds [voc_high]).

The percentage of no responses in the test on vocoded sounds was 4.2% for the trained test sounds and 16.4% for the new test sounds (see Figure S1 in the online supplemental materials). To examine whether sound type had an effect on proportion of responses, a Friedman test was performed which rendered a chi-square value of 26.857 ($p < .001$), which demonstrated that the effect of sound type on the proportion of responses to low spectrum sensor was significant. Post hoc Wilcoxon's tests showed that the proportions of responses to the vocoded low-frequency sounds (voc_low) were significantly different both from those to the vocoded high-frequency sounds (voc_high; $p = .018$) and also from those to the vocoded middle-frequency sounds (voc_mid; $p = .018$). The proportions of responses to the vocoded middle-frequency sounds (voc_mid) were significantly different from those to the vocoded high-frequency sounds (voc_high; $p = .018$).

Interim Conclusion for Spectrum Training

The zebra finches trained on spectrum differentiated among the vocoded test sounds with a low-, middle- and high-frequency spectrum. This indicates that zebra finches generalized for the parameter that the test sounds shared with the training sounds, namely, the spectral envelope. Furthermore, these birds did not show this generalization to the flat spectrum sounds, indicating that the feature that distinguishes these sounds (pitch value) was not relevant for them.

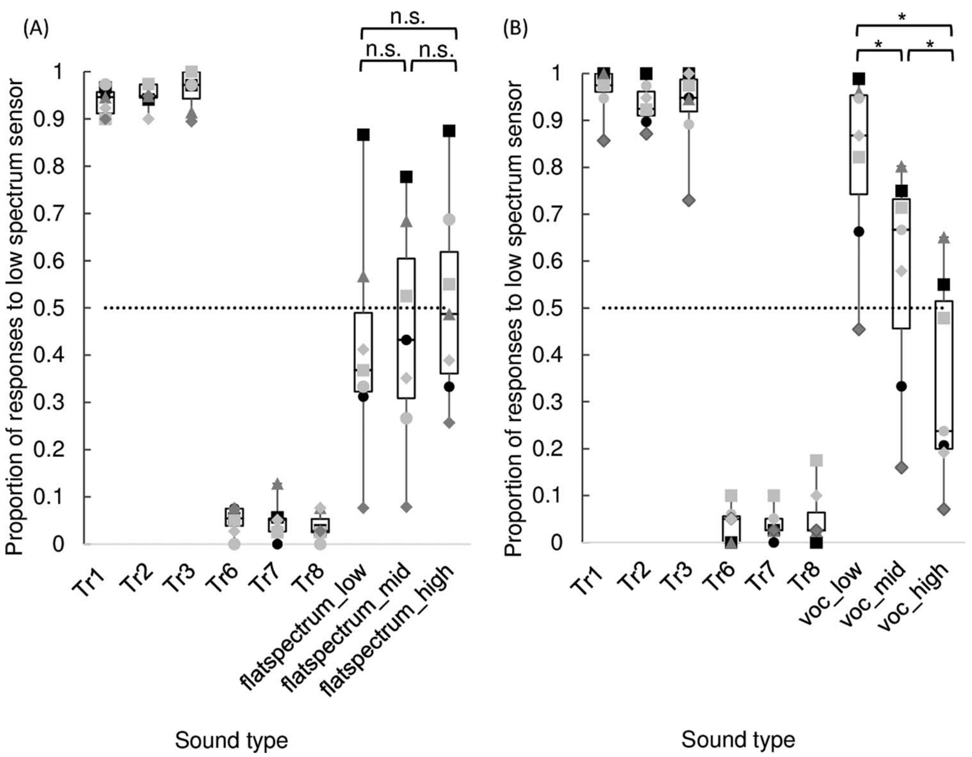


Figure 5. Spectrum training: The proportion of responses to the low spectrum sensor for each sound type for birds trained on spectrum and tested on flat spectrum sounds (A) and vocoded test sounds (B). Boxplots display median, interquartile range, and full range. Individual data points are indicated with symbols. Significant (* $p <$ .05) and nonsignificant differences (n.s.) among the test sounds are indicated at the top. A reference line marks the 0.5 proportion.

Performance on Test Trials After Pitch Training

Figure 6 displays boxplots with the median, interquartile range, and full range of the proportion of responses to the sensor associated with low-pitched sound for the test phase with the flat spectrum test sounds (Figure 6A) and the vocoded speech sounds (Figure 6B; for individual data, see Tables S3a and S3b in the online supplemental materials). The birds had on average 109 ± 57 trials per day. We found no significant differences in performance of birds tested with Set 1 and those tested with Set 2.

The percentage of no responses in the test on flat spectrum sounds was 4.2% for the trained test sounds and 22.1% for the new test sounds (see Figure S1 in the online supplemental materials). One individual bird (ID 675) showed a very high no response rate. Whereas the no response rate to the trained test sounds was 6.7%, it did not respond to 88.3% of the new test sounds. Therefore, we removed the data points of bird 675 from the analysis. For the six remaining individuals, the percentage of no responses in the test on flat spectrum sounds dropped to 3.8% of the trained test sounds and 11.1% of the new test sounds.

A Friedman test was performed which rendered a chi-square value of 22.267 ($p < .001$), which showed that the effect of sound type on the proportion of responses to the low-pitch sensor was significant. Post hoc Wilcoxon's tests showed that proportions of responses to the flat spectrum sounds with a low pitch (flatspectrum_low) were significantly different from those to the flat spectrum sounds with a high pitch (flatspectrum_high; $p =$ .028) and also from those to the middle pitch (flatspectrum_mid; $p =$ .028). Proportions of responses to the flat spectrum sounds with a middle pitch (flatspectrum_mid) were significantly different from those to the flat spectrum sounds with high pitch (flatspectrum_high; $p =$ .028).

The percentage of no responses in the test on vocoded sounds was 8% for the trained test sounds and 27.5% for the new test sounds (see Figure S1 in the online supplemental materials). The same individual bird as in the previous test (675) showed a very high no response rate. Whereas the no response rate to the trained test sounds was 6.5%, it did not respond to 82.8% of the new test sounds. Therefore, here we removed the data points of bird 675 from the analysis. For six remaining individuals, the percentage of no responses in the test on vocoded sounds dropped to 6.4% for the trained sounds and 18.3% for the new test sounds.

The responses for the vocoded versions with different f0 but the roughly same frequency band (three low, two middle, and three high frequency) were not significantly different from each other. Therefore, we calculated average proportion of responses for these sound types (vocoded low-frequency sounds [voc_low], vocoded middle-frequency sounds [voc_mid], vocoded high-frequency sounds [voc_high]).

A Friedman test was performed which rendered a chi-square value of 19.733 ($p =$ .001), which showed that the effect of sound type on

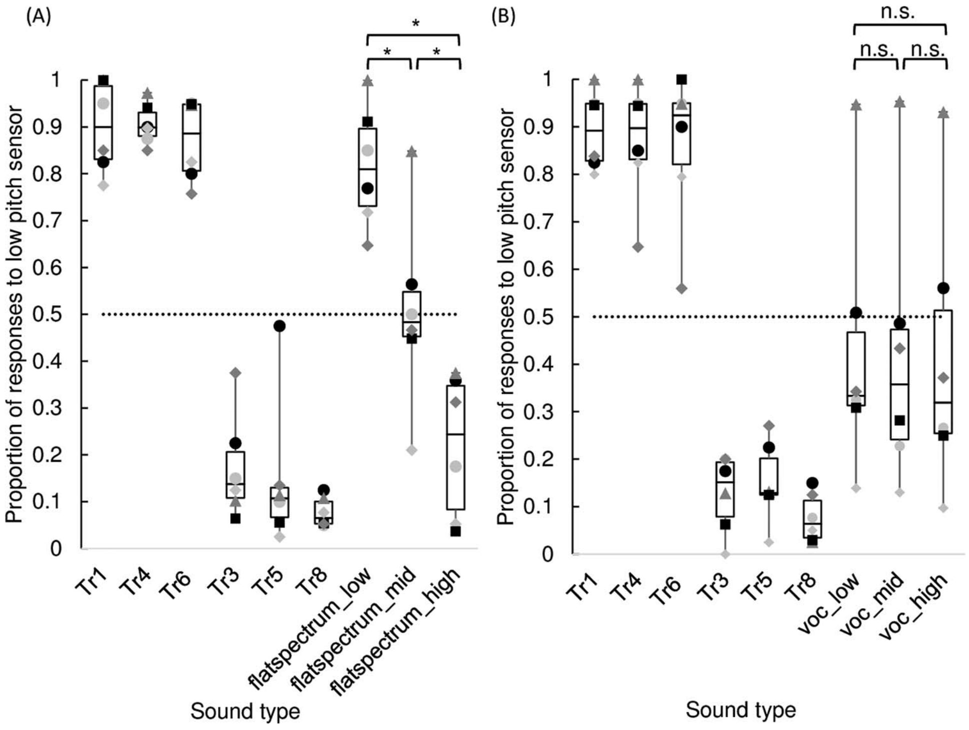


Figure 6. Pitch training: The proportion of responses to low pitch sensor for each sound type for birds trained on pitch mapping and tested on the flat spectrum sounds (A) and vocoded sounds (B). Boxplots display median, interquartile range, and full range. Individual data points are indicated with symbols. Significant (* $p < .05$) and nonsignificant differences (n.s.) among the test sounds are indicated at the top. A reference line marks the 0.5 proportion.

the proportion of responses to the low-frequency sensor was significant. Post hoc Wilcoxon's tests showed that the proportions of responses to the vocoded high-frequency sounds (voc_high) were not significantly different from those to the vocoded low-frequency sounds (voc_low; $p = .345$) and those to the vocoded middle-frequency sounds (voc_mid; $p = .917$). The proportions of responses for the vocoded low-frequency sounds (voc_low) were not significantly different from those to the vocoded middle-frequency sounds (voc_mid; $p = .345$).

Interim Conclusion for Pitch Training

The zebra finches discriminated between the flat spectrum sounds differing in pitch. Discrimination between low and high pitch was maintained after removing amplitude peaks from the harmonic spectrum. Proportion of responses to the flat spectrum sounds with middle pitch fell right between the low- and high-pitched test sounds. Responses to flat spectrum sound with low pitch were similar to responses to the trained sounds with low pitch and responses to flat spectrum sound with high pitch were similar to responses to trained sounds with high pitch. These findings show that zebra finches generalized for the shared parameter, namely, pitch. In contrast with birds trained on spectrum, birds trained on pitch do not react differently to the low versus high vocoded versions of the trained sound, indicating that the feature that distinguishes these sounds (formant-like frequency) was not relevant for them.

Discussion

Our results show that zebra finches can learn to discriminate artificial vowel-like sounds either by differences in fundamental frequency (pitch) or by differences in the spectral envelope. The parameter that zebra finches will extract from these speech-like sounds is experience dependent, in this case whether they were trained on pitch or spectrum.

Generalization for Spectral Envelope

Our study is the first showing that birds can generalize from harmonic sounds with formant-like amplified frequency bands to vocoded versions of the spectral envelope of their formants. Zebra finches are known to be very sensitive to fine spectral details of sounds (Dooling & Prior, 2017). It may therefore be surprising that they treat vocoded stimuli as being very similar to the original ones, despite substantial differences in spectral structure. However, the categorization of vocoded stimuli by zebra finches is not fully unexpected because other animal species also showed discrimination of vocoded speech-like stimuli. The identification of such stimuli has also been shown for chinchillas (Loebach & Wickesberg, 2006). In vivo recordings demonstrated that vocoded speech produced similar temporal representations in auditory nerves as produced by related natural stimuli, which could explain recognition of vocoded signals (Loebach & Wickesberg, 2006). Also Panzee, a common chimpanzee, responded significantly above chance when she was exposed to vocoded speech in a word-lexigram matching task (Heimbauer, Beran, & Owren, 2011).

Starlings, after being trained on ascending versus descending four-tone sequences differing in timbre, were also able to generalize discrimination to vocoded versions of the trained stimuli (Bregman et al., 2016). By comparing the birds' performance on the vocoded stimuli with piano-tone stimuli that shared absolute pitch but had a different timbre compared with the trained sounds, the authors showed that the starlings used the spectral envelope (maintained in vocoded stimuli) rather than pitch (maintained in piano-tone stimuli; Bregman et al., 2016). Patel (2017) has hypothesized that birds may generally be attending more to spectral envelopes of sounds than to pitch. If so, this indicates that birds may process complex spectral sounds more similar to how humans are processing speech than to how humans process music (Shannon, 2016), as vocoded speech sounds, in contrast to vocoded musical stimuli, are often still well recognizable by human listeners. We suggest that the ability of birds to recognize vowels spoken by human speakers differing in pitch of voice may be based on using similarities in spectral envelope.

Generalization for Pitch

That zebra finches can attend to pitch fits previous results on pitch perception in songbirds. Various studies demonstrated their ability to attend to absolute pitch (Cynx, 1995; Hulse & Cynx, 1985; Hulse et al., 1984; Lee et al., 2006; MacDougall-Shackleton & Hulse, 1996; Page et al., 1989; see also Hoeschele, 2017, for a review). Although starlings had difficulty generalizing the learned pitch cues to the piano tone stimuli (Bregman et al., 2016), our zebra finches were able to selectively attend to pitch and their recognition of pitch was not disrupted by the modification of spectral information in the flat spectrum sounds.

Our results have implications for the understanding of perception of human speaker sex by birds. Speaker sex differences are predominantly based on differences in voice pitch. Previous studies showed that budgerigars (Dooling & Brown, 1990) and zebra finches (Burgering et al., 2018) are sensitive to acoustic differences between vowels produced by males and females, although the budgerigars seem to find acoustic differences between different speakers less salient than the differences between vowels (Dooling & Brown, 1990). Results from the current study suggests that songbirds can identify pitch, or fundamental frequency, in speech-like harmonic sounds and that they can generalize this to harmonic sounds with different spectrum.

Mammal–Bird Differences?

Taken together, our studies suggest that pitch processing and processing of spectral envelope by songbirds might be more flexible than previously thought. Furthermore, as also suggested by others (Bregman et al., 2012; Hoeschele, 2017; Patel, 2017), it seems to depend on the context (in our case the training procedure) of which parameter is used to discriminate sounds. Our findings thus question the traditional idea that there is a fundamental difference in auditory processing of pitch between mammals and songbirds (Weisman et al., 2004). In general, it may be true that songbirds attend to absolute pitch by default (Hoeschele, 2017), but when another parameter is trained, they can categorize based on, for example, spectral envelope (Bregman et al., 2016). At the same time, studies on mammals indicate that relative pitch perception is not always their default strategy. For example, a study on pitch perception in ferrets showed that they needed additional training to expand the frequency range over which the relative pitch perception occurred (Yin et al., 2010). The authors concluded that although these mammals can also be trained to attend to relative pitch they might have the natural tendency to attend to absolute pitch (Yin et al., 2010).

Underlying Mechanisms

Generalization to new sounds that only share one sound parameter with the trained sounds could be facilitated by rule learning (Ashby & Gott, 1988; Smith et al., 2012; Smith, Zakrzewski, Johnson, Valleau, & Church, 2016), that is, by the birds discovering that all sounds in one category shared an underlying feature and next using this feature to also categorize novel sounds. This form of learning contrasts with an alternative way of distinguishing sounds, which is by exemplar learning, that is, memorizing all trained sounds. In our experiments, almost all birds readily generalized to the novel test stimuli using either the pitch or spectrum that defined the categories, supporting earlier suggestions of rule learning by some zebra finches (Burgering et al., 2018).

If the birds would only have memorized the trained sounds (exemplars), we would expect less discrimination of the test sounds and also few responses to these sounds. This is actually the pattern shown by the one bird (675) that deviated strongly in its behavior to the test sounds. It performed well in the training, but showed very few responses in the tests, suggesting it did not attend to general similarities between training and test sounds.

Conclusion

The current study shows that zebra finches cannot only learn to discriminate sounds based on pitch and spectrum depending on the context, but also show generalization for the relevant underlying parameter. This suggests that zebra finches can identify and generalize for both parameters independently to discriminate and categorize complex vowel-like sounds. It confirms the suggestion that songbirds are very flexible in their perceptual strategy and use the most relevant parameters to succeed in an auditory discrimination task (Bregman et al., 2012).

Unlike humans, birds do not have to deal on a daily basis with variability in human speech. However, pitch and spectral features do also vary among natural songs and calls of zebra finches, and might possibly vary independently to convey different types of information. The spectral envelope may be relevant for their communication because their conspecific calls differ primarily by this feature (Elie & Theunissen, 2016). Male testosterone levels were found to be related to pitch height in songs, which in turn might be one of the predictors for reproductive success for males in a wild population (Woodgate, Mariette, Bennett, Griffith, & Buchanan, 2012). Zebra finches may thus use both sound parameters for recognition of conspecific calls and songs or estimation of the sender's quality. Taken together, our study supports the hypothesis that flexible pitch and spectral processing might be a more general perceptual ability of many vertebrates and not unique for humans (Patel, 2017).

References

Adank, P., van Hout, R., & Smits, R. (2004). An acoustic description of the vowels of northern and southern standard Dutch. *The Journal of the Acoustical Society of America, 116,* 1729–1738. http://dx.doi.org/10.1121/1.1779271

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 33–53. http://dx.doi.org/10.1037/0278-7393.14.1.33

Bizley, J. K., Walker, K. M. M., King, A. J., & Schnupp, J. W. H. (2013). Spectral timbre perception in ferrets: Discrimination of artificial vowels under different listening conditions. *The Journal of the Acoustical Society of America, 133,* 365–376. http://dx.doi.org/10.1121/1.4768798

Bregman, M. R., Patel, A. D., & Gentner, T. Q. (2012). Stimulus-dependent flexibility in non-human auditory pitch processing. *Cognition, 122,* 51–60. http://dx.doi.org/10.1016/j.cognition.2011.08.008

Bregman, M. R., Patel, A. D., & Gentner, T. Q. (2016). Songbirds use spectral shape, not pitch, for sound pattern recognition. *Proceedings of the National Academy of Sciences of the United States of America, 113,* 1666–1671. http://dx.doi.org/10.1073/pnas.1515380113

Brosch, M., Selezneva, E., Bucks, C., & Scheich, H. (2004). Macaque monkeys discriminate pitch relationships. *Cognition, 91,* 259–272. http://dx.doi.org/10.1016/j.cognition.2003.09.005

Burgering, M. A., ten Cate, C., & Vroomen, J. (2018). Mechanisms underlying speech sound discrimination and categorization in humans and zebra finches. *Animal Cognition, 21,* 285–299. http://dx.doi.org/10.1007/s10071-018-1165-3

Cynx, J. (1995). Similarities in absolute and relative pitch perception in songbirds (starlings and zebra finch) and a nonsongbird (pigeon). *Journal of Comparative Psychology, 109,* 261–267. http://dx.doi.org/10.1037/0735-7036.109.3.261

Dewson, J. H., III. (1964). Speech sound discrimination by cats. *Science, 144,* 555–556. http://dx.doi.org/10.1126/science.144.3618.555

Dooling, R. J. (1992). *Perception of speech sounds by birds*. Paper presented at the Ninth International Symposium on Hearing: Auditory Physiology and Perception, Carcens, France.

Dooling, R. J., & Brown, S. D. (1990). Speech perception by budgerigars (*Melopsittacus undulatus*): Spoken vowels. *Perception and Psychophysics, 47,* 568–574. http://dx.doi.org/10.3758/BF03203109

Dooling, R. J., & Prior, N. H. (2017). Do we hear what birds hear in birdsong? *Animal Behaviour, 124,* 283–289. http://dx.doi.org/10.1016/j.anbehav.2016.10.012

Dorman, M. F., Loizou, P. C., & Rainey, D. (1997). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *Journal of the Acoustical Society of America, 102,* 2403–2411. http://dx.doi.org/10.1121/1.419603

Elie, J. E., & Theunissen, F. E. (2016). The vocal repertoire of the domesticated zebra finch: A data-driven approach to decipher the information-bearing acoustic features of communication signals. *Animal Cognition, 19,* 285–315. http://dx.doi.org/10.1007/s10071-015-0933-6

Friedrich, A., Zentall, T., & Weisman, R. (2007). Absolute pitch: Frequency-range discriminations in pigeons (*Columba livia*): comparisons with zebra finches (*Taeniopygia guttata*) and humans (*Homo sapiens*). *Journal of Comparative Psychology, 121,* 95–105. http://dx.doi.org/10.1037/0735-7036.121.1.95

Fuller, C. D., Gaudrain, E., Clarke, J. N., Galvin, J. J., Fu, Q. J., Free, R. H., & Başkent, D. (2014). Gender categorization is abnormal in cochlear implant users. *Journal of the Association for Research in Otolaryngology, 15,* 1037–1048. http://dx.doi.org/10.1007/s10162-014-0483-7

Goudbeek, M., Swingley, D., Kluender, K. R. (2007). The limits of multidimensional category learning. *Interspeech 2007: Eighth Annual Conference of the International Speech Communication Association, Vol. 1–4* (pp. 1301–1304). Retrieved from https://dblp.org/rec/conf/interspeech/GoudbeekSK07

Goudbeek, M., Swingley, D., & Smits, R. (2009). Supervised and unsupervised learning of multidimensional acoustic categories. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1913–1933. http://dx.doi.org/10.1037/a0015781

Heimbauer, L. A., Beran, M. J., & Owren, M. J. (2011). A chimpanzee recognizes synthetic speech with significantly reduced acoustic cues to phonetic content. *Current Biology, 21,* 1210–1214. http://dx.doi.org/10.1016/j.cub.2011.06.007

Hoeschele, M. (2017). Animal pitch perception: melodies and harmonies. *Comparative Cognition and Behavior Reviews, 12,* 5–18. http://dx.doi.org/10.3819/CCBR.2017.120002

Holt, L. L., & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *The Journal of the Acoustical Society of America, 119,* 3059–3071. http://dx.doi.org/10.1121/1.2188377

Holt, L. L., & Lotto, A. J. (2010). Speech perception as categorization. *Attention, Perception and Psychophysics, 72,* 1218–1227. http://dx.doi.org/10.3758/APP.72.5.1218

Hulse, S. H., & Cynx, J. (1985). Relative pitch perception is constrained by absolute pitch in songbirds (*Mimus, Molothrus,* and *Sturnus*). *Journal of Comparative Psychology, 99,* 176–196. http://dx.doi.org/10.1037/0735-7036.99.2.176

Hulse, S. H., Cynx, J., & Humpal, J. (1984). Absolute and relative pitch discrimination in serial pitch perception by birds. *Journal of Experimental Psychology: General, 113,* 38–54. http://dx.doi.org/10.1037/0096-3445.113.1.38

Hurly, T. A., Ratcliffe, L., & Weisman, R. (1990). Relative pitch recognition in white-throated sparrows (*Zonotrichia albicollis*). *Animal Behaviour, 40,* 176–181. http://dx.doi.org/10.1016/S0003-3472(05)80677-3

Izumi, A. (2001). Relative pitch perception in Japanese monkeys (*Macaca fuscata*). *Journal of Comparative Psychology, 115,* 127–131. http://dx.doi.org/10.1037/0735-7036.115.2.127

Kluender, K. R., Lotto, A. J., Holt, L. L., & Bloedel, S. L. (1998). Role of experience for language-specific functional mappings of vowel sounds. *Journal of the Acoustical Society of America, 104,* 3568–3582. http://dx.doi.org/10.1121/1.423939

Kriengwatana, B., Escudero, P., Kerkhoven, A. H., & Cate, C. T. (2015). A general auditory bias for handling speaker variability in speech? Evidence in humans and songbirds. *Frontiers in Psychology, 6,* 1243. http://dx.doi.org/10.3389/fpsyg.2015.01243

Kriengwatana, B., Escudero, P., & ten Cate, C. (2015). Revisiting vocal perception in non-human animals: A review of vowel discrimination, speaker voice recognition, and speaker normalization. *Frontiers in Psychology, 5,* 1543. http://dx.doi.org/10.3389/fpsyg.2014.01543

Lee, T. T. Y., Charrier, I., Bloomfield, L. L., Weisman, R. G., & Sturdy, C. B. (2006). Frequency-range discriminations and absolute pitch in black-capped chickadees (*Poecile atricapillus*), mountain chickadees (*Poecile gambeli*), and zebra finches (*Taeniopygia guttata*). *Journal of Comparative Psychology, 120,* 217–228. http://dx.doi.org/10.1037/0735-7036.120.3.217

Loebach, J. L., & Wickesberg, R. E. (2006). The representation of noise vocoded speech in the auditory nerve of the chinchilla: Physiological correlates of the perception of spectrally reduced speech. *Hearing Research, 213,* 130–144. http://dx.doi.org/10.1016/j.heares.2006.01.011

Lohr, B., & Dooling, R. J. (1998). Detection of changes in timbre and harmonicity in complex sounds by zebra finches (*Taeniopygia guttata*) and budgerigars (*Melopsittacus undulatus*). *Journal of Comparative Psychology, 112,* 36–47. http://dx.doi.org/10.1037/0735-7036.112.1.36

Lohr, B., Dooling, R. J., & Bartone, S. (2006). The discrimination of temporal fine structure in call-like harmonic sounds by birds. *Journal of*

Comparative Psychology, 120, 239–251. http://dx.doi.org/10.1037/0735-7036.120.3.239

MacDougall-Shackleton, S. A., & Hulse, S. H. (1996). Concurrent absolute and relative pitch processing by European starlings (*Sturnus vulgaris*). *Journal of Comparative Psychology, 110,* 139–146. http://dx.doi.org/10.1037/0735-7036.110.2.139

Ohms, V. R., Escudero, P., Lammers, K., & ten Cate, C. (2012). Zebra finches and Dutch adults exhibit the same cue weighting bias in vowel perception. *Animal Cognition, 15,* 155–161. http://dx.doi.org/10.1007/s10071-011-0441-2

Ohms, V. R., Gill, A., Van Heijningen, C. A. A., Beckers, G. J. L., & ten Cate, C. (2010). Zebra finches exhibit speaker-independent phonetic perception of human speech. *Proceedings of the Royal Society B-Biological Sciences, 277,* 1003–1009. http://dx.doi.org/10.1098/rspb.2009.1788

Page, S. C., Hulse, S. H., & Cynx, J. (1989). Relative pitch perception in the European starling (*Sturnus vulgaris*): Further evidence for an elusive phenomenon. *Journal of Experimental Psychology: Animal Behavior Processes, 15,* 137–146. http://dx.doi.org/10.1037/0097-7403.15.2.137

Patel, A. D. (2017). Why doesn't a songbird (the European starling) use pitch to recognize tone sequences? The informational independence hypothesis. *Comparative Cognition and Behavior Reviews, 12,* 19–32. http://dx.doi.org/10.3819/CCBR.2017.120003

Roberts, B., Summers, R. J., & Bailey, P. J. (2011). The intelligibility of noise-vocoded speech: Spectral information available from across-channel comparison of amplitude envelopes. *Proceedings of the Royal Society B-Biological Sciences, 278,* 1595–1600. http://dx.doi.org/10.1098/rspb.2010.1554

Shannon, R. V. (2016). Is birdsong more like speech or music? *Trends in Cognitive Sciences, 20,* 245–247. http://dx.doi.org/10.1016/j.tics.2016.02.004

Smith, J. D., Berg, M. E., Cook, R. G., Murphy, M. S., Crossley, M. J., Boomer, J., . . . Grace, R. C. (2012). Implicit and explicit categorization: A tale of four species. *Neuroscience and Biobehavioral Reviews, 36,* 2355–2369. http://dx.doi.org/10.1016/j.neubiorev.2012.09.003

Smith, J. D., Zakrzewski, A. C., Johnson, J. M., Valleau, J. C., & Church, B. A. (2016). Categorization: the view from animal cognition. *Behavioral Sciences, 6,* 24. http://dx.doi.org/10.3390/bs6020012

Town, S. M., Atilgan, H., Wood, K. C., & Bizley, J. K. (2015). The role of spectral cues in timbre discrimination by ferrets and humans. *The Journal of the Acoustical Society of America, 137,* 2870–2883. http://dx.doi.org/10.1121/1.4916690

Uno, H., Maekawa, M., & Kaneko, H. (1997). Strategies for harmonic structure discrimination by zebra finches. *Behavioural Brain Research, 89,* 225–228. http://dx.doi.org/10.1016/S0166-4328(97)00064-8

Vignal, C., & Mathevon, N. (2011). Effect of acoustic cue modifications on evoked vocal response to calls in zebra finches (*Taeniopygia guttata*). *Journal of Comparative Psychology, 125,* 150–161. http://dx.doi.org/10.1037/a0020865

Walker, K. M. M., Schnupp, J. W. H., Hart-Schnupp, S. M. B., King, A. J., & Bizley, J. K. (2009). Pitch discrimination by ferrets for simple and complex sounds. *The Journal of the Acoustical Society of America, 126,* 1321–1335. http://dx.doi.org/10.1121/1.3179676

Weary, D. M., & Weisman, R. G. (1991). Operant discrimination of frequency ratio in the black-capped chickadee (*Parus atricapillus*). *Journal of Comparative Psychology, 105,* 253–259. http://dx.doi.org/10.1037/0735-7036.105.3.253

Weisman, N., Njegovan, M., & Ito, S. (1994). Frequency ratio discrimination by zebra finches (*Taeniopygia guttata*) and humans (*Homo sapiens*). *Journal of Comparative Psychology, 108,* 363–372. http://dx.doi.org/10.1037/0735-7036.108.4.363

Weisman, R., Njegovan, M., Sturdy, C., Phillmore, L., Coyle, J., & Mewhort, D. (1998). Frequency-range discriminations: Special and general abilities in zebra finches (*Taeniopygia guttata*) and humans (*Homo sapiens*). *Journal of Comparative Psychology, 112,* 244–258. http://dx.doi.org/10.1037/0735-7036.112.3.244

Weisman, R. G., Njegovan, M. G., Williams, M. T., Cohen, J. S., & Sturdy, C. B. (2004). A behavior analysis of absolute pitch: Sex, experience, and species. *Behavioural Processes, 66,* 289–307. http://dx.doi.org/10.1016/j.beproc.2004.03.010

Woodgate, J. L., Mariette, M. M., Bennett, A. T. D., Griffith, S. C., & Buchanan, K. L. (2012). Male song structure predicts reproductive success in a wild zebra finch population. *Animal Behaviour, 83,* 773–781. http://dx.doi.org/10.1016/j.anbehav.2011.12.027

Yin, P., Fritz, J. B., & Shamma, S. A. (2010). Do ferrets perceive relative pitch? *The Journal of the Acoustical Society of America, 127,* 1673–1680. http://dx.doi.org/10.1121/1.3290988

Received March 8, 2018
Revision received May 30, 2018
Accepted June 26, 2018 ■