Octave equivalence perception is not linked to vocal mimicry: budgerigars fail standardized operant tests for octave equivalence

Bernhard Wagner^a, Dan C. Mann^{a,b}, Shahrzad Afroozeh^a,
Gabriel Staubmann^a and Marisa Hoeschele^{a,c,*}

^a Department of Cognitive Biology, University of Vienna, Vienna, Austria

^b Linguistics Program, The Graduate Center of the City University of New York,
New York, NY, USA

^c Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria

*Corresponding author's e-mail address: marisa.hoeschele@gmail.com

Received 10 January 2017; initial decision 5 March 2017; revised 9 March 2017;
accepted 27 March 2017; published online 21 January 2019

Abstract

Octave equivalence describes the perceived similarity of notes separated by an octave or a doubling in frequency. In humans, octave equivalence perception is used in vocal learning, enabling young children to approximate adult sounds where the pitch lies outside of their vocal range. This makes sense because the octave is also the first harmonic of any tonal sound including the human voice. We hypothesized that non-human animals may also need octave equivalence perception in vocal mimicry, the copying of other species or environmental sounds, to approximate sounds where the pitch lies outside their vocal range. Thus, in the current study, we tested budgerigars (*Melopsittacus undulatus*), a vocal mimicking species, for octave equivalence perception. Budgerigars were trained and tested in a go/no-go operant task previously verified in humans. Budgerigars did not show evidence of octave equivalence perception. This result suggests that vocal-mimicking does not necessarily facilitate or presuppose octave equivalence perception.

Keywords

octave equivalence, vocal mimicry, budgerigars, acoustics, comparative cognition, perception, music, language.

1. Introduction

Octave equivalence perception describes the perceived similarity of notes separated by an octave, which is a doubling in frequency (Burns, 1999; Patel, 2003). To humans, a note and another note with exactly twice the first note's frequency sound alike — in some cases more alike than the first note and a third note with a frequency closer to its own (see Allen, 1967; Kallman, 1982; Hoeschele et al., 2012a). While cultural differences in pitch perception have been shown (Hove, 2009), the octave as the basis of pitch perception is common to all musical cultures (Burns, 1999), possibly because the perception of octave equivalence may be rooted in physiology (Braun & Chaloupka, 2006). In humans, octave equivalence occurs spontaneously, such as when a young child imitates their singing father. If the father is singing in his normal range and not using a falsetto voice, it likely has a fundamental frequency that is too low for the child to produce. The child will spontaneously produce the octave and we perceive this as a successful imitation (see Hoeschele, 2017 for a review).

The perception of octave equivalence appears to facilitate language acquisition as well: Pre-verbal infants have been shown to respond less to tones an octave above a note they had previously been habituated to, compared to other tones (Demany & Armand, 1984). Also young children that are presented with non-words below their vocal range tend to imitate the words one octave higher. This is the case even if the octave of the presented non-word is still lower than their normal speaking range (Peter et al., 2008). These and additional data suggest that children (Peter et al., 2008, 2009) and adults (Peter et al., 2015) perceive the octave relationship as a salient aspect of similarity in speech and use it for an authentic sounding reproduction of the stimulus. Perhaps this is because if one cannot reproduce the same fundamental frequency, one can achieve maximal harmonic overlap by producing the octave: An octave above a note shares half the original note's harmonics (Hoeschele, 2017). Therefore, transposing an imitated sound by an octave results in the closest approximation of a sound in the case where producing the original fundamental frequency is not possible.

Based on the above, we might conclude that the fact that humans reproduce each other's vocalizations may be the reason our species has octave equivalence. One way to test this hypothesis is to look at other species. The ability to learn new vocalizations is known as vocal learning and is relatively

rare in the animal kingdom (Janik & Slater, 1997). Is octave equivalence related to whether or not an animal is a vocal learner?

Research regarding octave equivalence in non-human animals has been sparse as well as contradictory (Burns, 1999; Hoeschele et al., 2015, 2017). Two non-vocal learning species potentially have octave equivalence: rhesus monkeys, *Macaca mulatta* (Wright et al., 2000) and rats, *Rattus norvegicus* (Blackwell & Schlossberg, 1943). However, only two rhesus monkeys were tested in the first study so it is difficult to draw any definite conclusions. In addition, the latter study was criticized for not controlling for harmonics, which would have included octave information (Burns, 1999). A few vocal learning species were also studied: a single bottle-nosed dolphin, *Tursiops truncatus*, trained to imitate sounds, was able to transpose sounds outside her preferred vocal-range by an octave (Richards et al., 1984). A songbird, the European Starling, *Sturnus vulgaris*, failed to show octave equivalence using an operant conditioning procedure (Cynx, 1993). However, Hoeschele et al. (2012a) found that the procedure from Cynx (1993) also failed to produce octave equivalence perception in humans, so it is not possible to draw any inferences there either.

All of the above studies either have a very low sample size or have other methodological issues that make it very difficult to answer our question of whether or not vocal learning is relevant for octave equivalence. As a result, Hoeschele et al. (2012a) developed a standardized nonverbal operant conditioning go/no-go task that successfully showed octave equivalence in humans and could also be used to test non-human animals. This task has since been used to test black-capped chickadees, *Poecile atricapillus*, a vocal learning songbird. However, the chickadees failed this procedure, suggesting they do not perceive octave-equivalence (Hoeschele et al., 2013). This is a strike against the hypothesis that vocal learning is the important factor in human octave equivalence.

We were interested in why the black-capped chickadees failed the task. Our hypothesis was that vocal imitation outside of one's vocal range is important in octave equivalence perception. In humans, vocal ranges differ between individuals, so the octave relationship may be used to approximate an imitated sound from a conspecific as accurately as possible (Hoeschele et al., 2017). Unlike humans, the vocal range of the black-capped chickadee does not appear to change during development nor differ between the sexes (Ficken, Ficken & Witkin, 1978). In addition, although black-capped

chickadees are vocal learners (Nowicki, 1989), chickadees do not copy the vocalizations of other species or environmental sounds, an ability known as vocal mimicry (Kelley et al., 2008). Vocal mimicry might be important in octave equivalence perception for species that are vocal mimics (even if, like chickadees, they do not have individually differing vocal ranges), since mimicking environmental or biological sounds could require imitating sounds that occur outside of one's own vocal range. This could mean that a vocal mimicking species may show octave equivalence perception in an experimental design in which a non-vocal mimicking (but vocal learning) species did not.

As such, based on the nonverbal operant design by Hoeschele et al. (2012a), we tested budgerigars (*Melopsittacus undulatus*) for octave equivalence perception. Budgerigars are not only vocal learners (Tyack, 2008; Tu et al., 2011), with males imitating their female mates (Hile et al., 2005), but they are also vocal mimics (Gramza, 1970). For example budgerigars can mimic human speech (Gramza, 1970; Scanlan, 1999) which has a mean frequency of about 100–250 Hz (Berg et al., 2017). This range is well below budgerigars' own vocal range (approximately 1000–5700 Hz; Brittain-Powell & Dooling, 2002). If imitation of sounds outside one's own vocal range is relevant for octave equivalence, we should expect to find it in budgerigars.

2. Materials and methods

2.1. Subjects

We tested 12 adult budgerigars aged between two and three years (sexual maturity reached within the first 4–8 months; Brockway, 1964; maximum lifespan 21 years in Munshi-South & Wilkinson, 2010), that were housed at the Animal Care Facility in the Department of Cognitive Biology at the University of Vienna maintained on a 14/10 h light/dark cycle, adhering to the budgerigar's diurnal cycle (Wyndham, 1980). The birds were kept in two aviaries ($2 \times 1 \times 2$ m) containing groups of eight budgerigars each. Five of the budgerigars in the left aviary had prior experience with the operant box (see "apparatus" below). They were, however, naive to the current stimulus set.

The birds were kept on a diet of water (ad libitum, except during the experiment) and Avifood's "Harrison's Bird Food, Adult Lifetime Superfine —

Maintenance Formula for Small Birds” pellet food. Birds were trained five days a week and food was administered after the completion of the experiments each day. They also received vitamins (Combex V Multivitamin) mixed in with their regular daily ration of pellet food or with their water twice a week.

2.2. Apparatus

During the experiment individual budgerigars were separated from the other birds in a wooden operant box (see Figure 1). The birds could respond to visual stimuli on a CarrollTouch infrared touch screen (Elo Touch Solutions, USA). The box was lit by an LED houselight that simulated daylight (6500 Kelvin; Paulman IP67 special line, Paulman, Vancouver, BC, Canada). Acoustic stimuli were fed through a SMSL SA-36A digital amplifier (SMSL, ShuangMuSanLin Electronic, Fuyong, Shenzhen, P.R. China) and played via a Fostex FE108 full-range speaker (Fostex, Tokyo, Japan; frequency range 200–16 000 Hz) located next to the feeder. The food reward was a highly desirable mix of grains (Versele-Laga Budgies Prestige; Deinde, Belgium), administered to the birds through a Campden Instruments 80209 Pellet Dispenser (Campden Instruments, Loughborough, UK), which was controlled by a Mac mini computer (Apple, Cupertino, CA, USA). All stages of the experiment were programmed in Python using “Experimenter”, see: <https://github.com/cogbio/Experimenter/blob/master/Experimenter.py>.

Figure 1. Apparatus: a, light; b, speaker; c, food dispenser with tube leading to bowl; d, touchscreen; e, sliding door.

2.3. Stimuli

The acoustic stimuli were sine wave tones constructed at a standard 16-bit, 44.1-kHz sampling rate presented via the speaker in the operant box. Each tone lasted 440 ms and was at the onset ramped upwards and at the offset was ramped downwards for 5 ms. This was done to avoid transients, as in Hoeschele et al. (2012a; see also Hoeschele et al., 2013). The pitch of these sine wave tones were the 12 notes of octave four, beginning with C4 (“middle C”, 262 Hz) and ending with B4 (494 Hz), and the 12 notes of octave five, beginning with C5 (523 Hz) and ending with B5 (988 Hz). Each of these stimuli was intermittently played at two different levels of amplitude, categorized as “S” for soft at 70 dB and “L” for loud at 80 dB SPL (A weighting), to control for possible effects of amplitude and so birds would not confound loudness and pitch (Hoeschele et al., 2013). The frequency for each note can be calculated with the formula $F_x = 2^{(1/12)} \times F_{x-1}$, where C4 = 262 Hz (Hoeschele et al., 2013). These stimuli were the same as the ones used for both humans and chickadees in previous work (see Hoeschele et al., 2012a, 2013). These frequencies are all well within budgerigar hearing range at the amplitudes we used (see Okanoya & Dooling, 1987; Heffner et al., 2016) and have been successfully identified by budgerigars in previous operant work (see Dooling et al., 1995; Dent et al., 2000; Weisman et al., 2004).

2.4. Procedure

Birds were separated from their conspecifics in the wooden operant box described above, where they could respond to the visual and auditory stimuli on a touch screen. The operant box was located next to the aviary, so that birds could maintain acoustic contact with their flockmates, to improve their motivation. The chickadees completing this task in prior work were also kept in a room with other birds during the experiment (Hoeschele et al., 2013). The average background intensity in this experimental apparatus was measured to be approximately 35–55 dB SPL at the position of the bird depending on the vocal activity of the colony birds.

Analogous to the black-capped chickadee experiment by Hoeschele et al. (2013) the budgerigars were trained on several phases of shaping and preliminary training, two phases of discrimination training and tested in two different test phases as detailed below. The training, as well as the testing sessions, lasted 30 min per day for each bird regardless of the phase of the

Figure 2. Timeline of all phases (training and testing). Oct four = 12 notes of octave four, beginning with C4 (“middle C”, 262 Hz) and ending with B4 (494 Hz). Oct five = 12 notes of octave five, beginning with C5 (523 Hz) and ending with B5 (988 Hz).

experiment they were on. A summary of the training phases can be found in Figure 2.

2.5. Training phases

2.5.1. Phase 1 — millet shape training

This phase was conducted to encourage the birds to respond by pecking the touchscreen. On each trial, budgerigars were shown a picture of millet randomly either on the left or the right side of the screen. If they responded by pecking the picture of millet or if they did not respond and waited for 30 s, they received a food reward accompanied by a positive acoustic feedback cue (this cue can be heard at: <https://github.com/cogbio/Experimenter/tree/master/input/media>). The criterion for being transferred from phase 1 to phase 2 was: Completion of two sessions with at least 40 responses.

2.5.2. Phase 2 — white circle shape training

In this phase the training was the same as in phase 1 except that a white circle appeared on the touchscreen instead of the picture of millet. If the birds responded by pecking it they received a food reward accompanied by a positive acoustic feedback cue. The criterion for being transferred to phase 3 was: Completion of two sessions with at least 40 responses.

2.5.3. Phase 3 — two button training

On all trials, budgerigars were presented with a white circle on the left side of the screen (trial start key). If the budgerigars pecked this circle it was extinguished and another white circle on the right side of the screen appeared (response key). Responding by pecking the response key within 2 s produced a reward accompanied by a positive acoustic feedback cue. If the birds did not peck the response key within a 2-s window, the screen returned to its original state without further feedback. This was registered as “timeout”. The trial start key then reappeared and they could initiate a new trial after one second. The criterion for being transferred to phase 4 was: Completion of two sessions with at least 50 responses to the response key.

2.5.4. Phase 4 — sound versus no sound training

In phase 4 the birds were trained to only respond (by pecking the response key) when a sound played. This was done to ensure that the birds paid attention to acoustic stimuli and also were responding equally to future rewarded, unrewarded and test tones. The sounds were the notes from octave four and octave five. As in phase 3, pecking on a left white circle (trial start key) would produce another white circle to the right (response key), but now a note from octave four or octave five accompanied the appearance of the response key

on 50% of the trials. Notes were played randomly without replacement until birds had heard all 48 stimuli (12 notes $\times$ 2 octaves $\times$ 2 loudnesses). At that point, all stimuli were re-added to the stimulus pool. If the birds pecked on the response key within two seconds after a sound occurred, a positive feedback sound played from the speaker, and they were rewarded with food. If the birds pecked on the response key within two seconds after no sound was played, a negative feedback cue (see: <https://github.com/cogbio/Experimenter/tree/master/input/media>) played, the screen flashed red, and the lights as well as the touchscreen went black for a 30-s period, during which no trials could be initiated. As in phase 3, if the birds did not peck the response key within a two second window, the screen returned to its original state without further feedback. This was registered as “timeout”. The trial start key then reappeared and they could initiate a new trial after one second.

The criterion for being transferred to phase 5 was: Completion of two sessions with a discrimination ratio (DR) of 0.80 or higher, with a minimum of at least 20 trials initiated. The DR is a tool to assess a bird’s learning progress which was calculated for each day’s data as follows (“S+” are rewarded stimuli, “S–” are non-rewarded stimuli): $DR = (\% \text{ response to S+}) / ((\% \text{ response to S+}) + (\% \text{ response to S-}))$.

2.5.5. Phase 5 — discrimination training

Phase 5 was identical to phase 4, except that now the appearance of the response key was accompanied by a note from octave four on every trial. In addition only responses to a subset of the notes (1/3rd) were rewarded and the others were punished with lights out. The notes from octave four were divided into three ranges: Lower S– (unrewarded) range, containing the notes: C4, C#4, D4 and D#4, S+ (rewarded) range, containing the notes: E4, F4, F#4 and G4, and higher S– (unrewarded) range, containing the notes: G#4, A4, A#4 and B4. See Figure 3 for a visual representation. As in phase 4, notes were played randomly and without replacement until all notes had been heard and then all stimuli were re-added to the stimulus pool. These ranges were used to match what was done in the previous work with humans and chickadees (Hoeschele et al., 2012a,b, 2013). Ranges (rather than individual notes) had to be used with humans because otherwise the task would have been too difficult. The original experiment with humans was thoroughly counterbalanced (Hoeschele et al., 2012a). As the goal was to compare budgerigars with humans, we used only one version of the experiment to increase statistical power because it was not possible to test as many budgerigars as humans.

Figure 3. Rewarded and unrewarded ranges in phase 5 (discrimination training).

The criterion for being transferred to phase 6 was reaching: 18 blocks of 50 trials above chance (≥ 23 responses correct), with the last six blocks consecutive. 23 or more correct responses is given as above chance because the birds showed a strong “go-bias”, meaning that they almost always responded to stimuli. Responding was, however, only the correct response for 1/3rd of the stimuli. This means that untrained birds did not get 25 responses (50%) correct on average for each 50 trial block, but only about 17 (33.3%). Through training they learned to reduce their response rate for punished notes. 23 correct responses is significantly above 17 if 17 is assumed to be chance in a binomial distribution. If the birds had not reached criterion after 300 days on phase 5, they were removed from the experiment.

2.5.6. Phase 6 — discrimination training with reduced reward

This phase was identical to discrimination training, except that responses to S+ notes were rewarded intermittently with a probability of 85% to blur the distinction between unrewarded test trials and training trials in the following generalization testing phase (see phase 7). The criterion for being transferred to phase 7 was reaching either: 1) Completion of at least two blocks of 50 trials with a DR ≥ 0.8 . OR 2) six trial blocks above chance (≥ 23 responses correct).

2.6. Test phases

2.6.1. Phase 7 — generalization testing

This phase was the same as phase 6 except that probe trials using the 12 notes from octave five now occurred on 20% of trials. These octave five notes were neither reinforced nor punished. The other 80% of trials were the notes from octave four and were reinforced the same way as in phase 6 (S+ tones reinforced 85% of the time and S- tones always punished). As such,

Figure 4. Notes used in test phases 7 (generalization testing), 8a (non-reversed transfer training) and 8b (reversed transfer training).

all notes from both octave four and five were presented in this phase (see Figure 4).

This test continued until the birds had heard every probe note from octave five at least three times at each amplitude, meaning that birds were presented with a total of 72 probe trials. We ensured that the birds heard every note from octave five at both amplitudes once before they heard any for a second time and again before they heard them a third time.

From Hoeschele et al. (2012a) it can be expected that an octave equivalence perceiving species would generalize the rules regarding which notes are rewarded and which are not from the lower to the higher octave. This means that an octave equivalence perceiving species should respond similarly to octave five probe notes as to their octave four counterpart.

2.6.2. Phase 6 — discrimination training with reduced reward

Once birds had completed phase 7 they were trained on phase 6, discrimination training with reduced reward, for a second time to ensure they were still responding accurately to the training stimuli before testing them again. The criterion for completing this interposed re-training on phase 6 remained the same as given above.

2.6.3. Phase 8 — transfer of training test

All 24 notes from octave four and octave five were played equally in a random order. The stimuli from octave four had the same reward contingencies as in phase 5. Prior to entering this phase, the subjects were divided pseudo-randomly controlling for sex into two groups. In each group different sets of notes from octave five were rewarded. Birds finished this phase by completing 500 trials. See below for an explanation of the reward contingencies of each group. Figure 5 gives a visual representation of which notes were rewarded in the two groups.

Figure 5. S+ and S– ranges in phase 8a (non-reversed transfer training, top) and phase 8b (reversed transfer training, bottom). Dark ranges are unrewarded, light ranges rewarded.

2.6.4. Phase 8a — non-reversed transfer training

In this group the same notes were reinforced in octave five as they were in octave four. S+ notes from octave four were S+ notes in octave five and S– notes from octave four were S– notes in octave five. Based on Hoeschele et al. (2012a) it should be expected that this task comes naturally to species that perceive octave equivalence.

2.6.5. Phase 8b — reversed transfer training

In this group the notes that were reinforced in octave four were not reinforced in octave five but punished. Responses to the notes that were not reinforced in octave four were rewarded in octave five. In other words the S+ notes from octave four became S– notes in octave five and the S– notes from octave four became S+ notes in octave five. Based on Hoeschele et al. (2012a) it should be expected that this task would be very difficult for species that perceive octave equivalence, as it was difficult for humans.

3. Results

All procedures performed were in accordance with Austrian animal protection and housing laws and were approved by the ethical board of the behavioural research group in the faculty of Life Sciences at the University of Vienna (approval number 2015-005).

3.1. Training

Of the original 12 budgerigars, four individuals entered the test phases. Of the other eight birds, one bird had to be excluded from the experiment because of insufficient interaction with the touchscreen in phase 1. Another

seven birds had to be excluded after failing to reach criterion in phase 5, meaning they had not begun to discriminate between notes after 300 days of training. Of the four birds that entered the test phase, two were female (Annabelle and Phoenix), and two were male (Topaz and Zora). In total, these birds spent between 83 and 127 days on phase 5 and between eight and 29 days on discrimination training with reduced reward (phase 6) before entering the first test phase, phase 7.

The two males reached phase 6, as well as the testing phase, sooner than either female did, and both males fulfilled phase 6 criterion 1 (Completion of at least two blocks of 50 trials with a $DR \geq 0.8$ for at least 2 sessions). The females entered the testing phase by fulfilling the more lax phase 6 criterion 2 (6 trial blocks with ≥ 23 responses correct).

3.2. Generalization testing: all birds

With 61.8% go-responses, the bird's overall response rate to octave five stimuli did not differ significantly from the overall response rate to octave four stimuli (62.1% go-responses; $p = 0.9466$ Fisher's exact test).

Figure 6, which shows the percent response to ranges in each octave, reveals that the number of responses within ranges differed considerably across the two octaves. We pooled the data across all four birds for a total of 1621 trials, 1333 of which were octave four notes and 288 of which were

Figure 6. Data from phase 7 (generalization testing). Percent response for ranges across all birds. Left: Octave four, right: Octave five. Error bars represent 95% confidence interval for the estimated probability.

octave five probe tones (72 for each individual bird). The data was analysed using the programs SPSS 22 and RStudio and conducting logistic regressions with covariates range, (S+/S– lower/S– higher), octave (four/five), volume (loud/silent) and interaction terms for range and octave as well as range and volume. We additionally controlled for the individual birds by including “bird” as a nominal covariate.

In our first analysis we set the middle (S+) range and octave four as baselines. We found significant interactions: There were more responses to the middle (S+) range in octave four than middle range in octave five ($p < 0.001$). There were significantly fewer responses to the lower S– range as well as to the higher S– range when compared to the S+ range in octave four (both: $p < 0.001$). In octave 5 there were fewer responses to the middle range (one octave above the S+ range in octave 4) than to the lower range ($p = 0.005$) or the higher ($p < 0.001$), both of which were one octave above S– ranges in octave 4. Across all birds there was no significant difference between responses to louder (80 dB) “L”-stimuli and softer (70 dB) “S”-stimuli in either octave four nor in octave five ($p > 0.05$).

3.3. Generalization testing: individual birds

We ran logistic regression analyses specified as above for the data of each of the four birds separately. Our individual bird analyses showed similar results for three of the individual birds (Annabelle, Topaz and Zora) as can be seen in Figure 7.

We found significant interactions: There were significantly more responses to the S+ range in octave four than to S+ range in octave five for Annabelle, Topaz and Zora (all $p < 0.001$). In octave four there were significantly fewer responses to the lower S– range for Annabelle, Topaz and Zora (all $p < 0.001$) when compared to the S+ range, and also significantly fewer responses to the higher S– range (all $p < 0.001$) when compared to the S+ range. For Annabelle and Topaz there were significantly fewer responses to the S+ range when compared to the lower S– range in octave five (Annabelle: $p = 0.0460$; Topaz: $p = 0.009$; Zora: $p = 0.149$). For Topaz there were also significantly fewer responses to the S+ range when compared to the higher S– range (Annabelle: $p = 0.756$; Topaz: $p < 0.001$; Zora: $p = 0.083$).

For the fourth bird, Phoenix, none of the effects and interactions of octave and range were significant at $p = 0.05$. This lack of effect may be due to

Figure 7. Data from phase 7 (generalization testing). Percent response to ranges for individual birds. Left: Octave four, right: Octave five.

overall poor performance: Phoenix failed to show statistically significant discrimination of the ranges even in octave four, meaning that she failed to discriminate among the original training stimuli during testing.

3.4. Transfer of training test: groups

There was no significant effect of transfer test group on total number of responses: Descriptive statistics show that birds in non-reversed transfer responded to 77.6% of all trials, birds in reversed transfer responded to 74.5% of trials.

We analysed the data of all four birds using a logistic regression model with covariates octave (4/5), transfer group (non-reversed/reversed), range (low/middle/high) and the two and three-way interactions between octave, range and transfer group.

3.5. Group a, non-reversed transfer training (Phoenix and Zora):

There were significantly fewer responses to the lower S– range in octave four ($p = 0.018$) as well as to the higher S– range in octave four ($p = 0.039$) when compared to the S+ range in octave 4. The responses to the lower and higher S– range in octave five did not differ significantly from those to the octave five S+ range (lower S– range: $p = 0.954$; higher S+ range: $p = 0.652$).

However, because Phoenix did not discriminate significantly in octave four during generalization testing (octave four S+ range compared to lower

Figure 8. Data from phase 8a (non-reversed transfer training). Percent response to ranges for individual birds. Left: Octave four, right: Octave five.

octave four S– range: $p = 0.8159$; octave four S+ range compared to octave four high S– range: $p = 0.4229$) we also ran another logistic regression using only Zora’s data. The response pattern stayed the same except that responses to the higher S– range in octave five differed significantly from the responses to the octave five S+ range ($p = 0.0211$) with more responses to the higher S– range. This is the opposite direction than would have been expected if Zora had learned the task (see Figure 8).

3.6. Group b, reversed transfer training (Annabelle and Topaz):

There were significantly fewer responses to the lower S– range in octave four ($p < 0.001$) as well as to the higher S– range in octave four ($p < 0.001$) when compared to the S+ range in octave four. There was a significantly higher number of responses to the lower S+ range in octave five than to the middle octave five S– range ($p = 0.001$). There were also significantly more responses to the higher S+ range in octave five than to the middle octave five S– range: ($p < 0.001$). This result suggests that these birds were able to learn the reversed pattern across the two octaves (see Figure 9).

4. Discussion

The results of this study suggest that budgerigars, a vocal-mimicking species, do not perceive octave equivalence. In this study, we could therefore not provide support for the hypothesis that vocal mimicry and perception of octave

Figure 9. Data from phase 8b (reversed transfer training). Percent response to ranges for individual birds, 500 trials. Left: Octave four, right: Octave five.

equivalence are connected. This suggests that while octave equivalence perception is important in human vocal mimicry during language acquisition and singing, other species' vocal mimicry skills may be facilitated by mechanisms different from octave equivalence perception.

Species that perceive octave equivalence would be expected to respond in a similar manner to notes separated by an octave. This was the case for humans tested on the task we used in the current study (Hoeschele et al., 2012a). The budgerigars, however, seemed to reverse their response pattern to notes from octave four in octave five. In our generalization test they responded more to the notes of octave five that were unrewarded in octave four and less to the octave five notes that had been rewarded in octave four. In a second test we presented birds either with a matched or an unmatched condition where notes across two octaves were either rewarded in the same or in the opposite manner. We expected that, like the humans in Hoeschele et al. (2012a), the birds should have an easier time learning the matched than the unmatched condition. Once again we found the opposite.

Of course it has to be noted that birds in the reversed condition (phase 8b) received a higher percentage of food reward when responding randomly than birds in the non-reversed condition did (phase 8a). Because the budgerigars were go-biased, the reversed condition allowed for more correct responses by chance. As such, one might argue that this bias could explain the birds' success in the reversed condition. However the results from the generalization

test contradict this notion, because they showed the same reversed pattern of responding even without additional reward. Also, humans, who do perceive octave equivalence, found it very difficult in Hoeschele et al. (2012a) to learn the reversed pattern, sometimes even making mistakes in discriminating octave four notes during transfer training. In addition, the human experiments were thoroughly counter-balanced without any changes in the overall response patterns. In some of these counter-balancings, the reversed condition contained lower chances of reward (Hoeschele et al., 2012a).

It should also be noted that despite extensive training, only 1/3rd of our birds completed the task. This was surprising given previous work suggesting that budgerigars are good at frequency discrimination tasks, including tasks using sinewave tones in a similar range to the range presented here (Dooling et al., 1995; Dent et al., 2000; Weisman et al., 2004). Several studies show that the stimuli we used are well within budgerigar hearing range at the amplitudes we used. Budgerigars have a fairly large hearing range which has been shown by Okanoya & Dooling (1987) to extend from 250 Hz to 6000 Hz at an amplitude of 40 dB. However Okanoya & Dooling did not test below 250 Hz. Newer work has shown that budgerigars have a hearing range of 77 Hz to 7.6 kHz at an amplitude of 60 dB (Heffner et al., 2016). However, it might also be worthwhile to replicate the experiments within the budgerigar's best hearing range. Budgerigars' range of best hearing lies between 1000 and 5700 Hz, with their very best hearing sensitivity at 2860 Hz, the frequency which corresponds to the peak of budgerigar vocalizations (Brittain-Powell et al., 2002). We know that humans with absolute pitch perception fail to perform accurately at very high and low octaves outside the human vocal range (Baharloo et al., 1998), so it could also be the case that budgerigars would perform most accurately in the middle of their hearing range.

It is possible that had we used a choice paradigm rather than a go/no-go paradigm we would have been able to avoid problems due to the go-bias we found with our birds. This is an important difference to black-capped chickadees, who are generally more "no-go" biased (Hoeschele et al., 2013). This difference can be seen from the fact that chickadees responded less to the new stimuli from octave five during testing (Hoeschele et al., 2013), while the budgerigars responded almost equally as much to stimuli from either octave. A modified design with more rewarded notes also might be an effective way to collect more budgerigar data rather than keeping the

design the same across species. Another way to increase data collection in the budgerigars in future studies would be to allow the birds to perform in the experiment for more than 30 min per day. The chickadees were housed in the experimental apparatus (Hoeschele et al., 2013) and thus performed many more trials per day than the budgerigars in our study. Keeping the subjects in the experimental apparatus might speed up the procedure in future studies.

This study may provide another hint that the perception of octave equivalence could potentially be a mammalian trait. While evidence from one bottle-nosed dolphin (*Tursiops truncatus*) suggests that that species uses the octave relationship in vocal production when mimicking a sound outside their vocal range (Richards et al., 1984), dolphins have not, as of yet, been tested directly for octave equivalence perception. Aside from humans, only one experiment with two rhesus monkeys (*Macaca mulatta*) suggested that perception of octave equivalence may be found in other primate species (Wright et al., 2000; see Hoeschele, 2017 for review). Unlike dolphins, rhesus monkeys are not vocal learners (Owren et al., 1992), but both species are mammals. Further support for the hypothesis that octave equivalence may be a mammalian trait comes from brain organization: A tonotopic organization in the brain of the non-vocal-learning cat (*Felis catus*) with an underlying octave structure has been suggested (Imig & Morel, 1985). Although it has been suggested earlier that the avian tonotopic organization is very similar to that of mammals (Zaretsky & Konishi, 1976) to our knowledge this specific feature has not, as of yet, been shown in birds (Köppel, 2001). This could mean that octave equivalence perception may be a mechanism for classifying tones which is shared among mammals generally and used by a subset of mammals when vocally mimicking. We would therefore like to invite researchers who work with mammals to test this paradigm on their study species. Especially species like rats, ferrets and sea lions might be good choices given that they have been used in behavioural research on auditory perception (Cook et al., 2013; Reichmuth & Casey, 2014; Gold et al., 2015; Celma-Miralles & Toro, 2018).

Here we studied the perception of octave equivalence, but it would be interesting to see whether budgerigars ever produce octave equivalence. Budgerigars mimicking human speech do not seem to replicate the fundamental frequency in all cases (Scanlan, 1999), however this has not been systematically studied. It remains unclear whether they use octave equivalence, frequency modulation, other harmonics, formants, or some other strategy

and we hope that controlled experimentation and careful acoustic analysis could make this clearer. Perhaps birds use a mechanism different from octave equivalence perception for classifying tones (Hoeschele, 2013). This could explain why chickadees — North American songbirds — responded very similarly to the paradigm at hand, as did the only distantly related budgerigars — Australian parrots. What this mechanism of classifying tones may be, however, remains unclear at this point. One possibility explaining the reversal in response pattern in octave 5, suggested by Hoeschele et al. (2013) would be that the birds learned a rule along the lines of: “Groups of four non-rewarded neighbouring notes alternate with groups of four rewarded ones”. The response pattern derived from this rule could have been continued into the higher octave. Although this rule is fairly abstract, it might well be within the abilities of budgerigars because they have been shown to have highly accurate absolute pitch (Weisman et al., 2004), and their vocal repertoire is more complex than that of most oscine song birds (Kline & Dooling, 1984). Perhaps testing the birds with the counter-balanced variations of the task that were tested in humans could help determine whether such a rule was indeed being used. Testing a non-human mammalian species with the paradigm we used here could determine whether the responses of the chickadees and budgerigars are specific to avian species.

There are several other explanations that are in our view unsatisfactory for the peculiar response pattern reversal across the two octaves. The first explanation is that the birds have a simple preference for certain tones. This seems unlikely given that a very similar preference would have to then be shared by chickadees, a species whose classification differs in order from budgerigars. Still, studying a third avian species distantly related to both tested species, or even studying other mammalian species, could help clarify this issue. What seems clear is that the response pattern we saw here was not just random error, because it was shared with another species tested on this same paradigm.

A second unsatisfactory explanation for the peculiar response pattern is that, while the birds may attend to the harmonic series in vocal imitation, they may simply prefer harmonics other than the octave. Octave equivalence perception makes sense for animals that attend to the harmonic series, because the octave of a given note is also the first harmonic of that note (see Hoeschele, 2017 for review). However, in some cases animals attend to the second harmonic rather than the first (e.g., see Cator et al., 2009). Rather

than an octave relationship, the interval between a fundamental frequency and the second harmonic has a relationship close to the perfect fifth relationship of Western music. If animals treat notes separated by an interval close to a perfect fifth as being similar, rather than an octave, then we might expect a response pattern that differs from the humans tested with the current paradigm. Attending to a perfect fifth rather than an octave occurs in humans as well. It has also been shown that adult humans imitating sounds presented by a member of the opposite sex will sometimes transpose the fundamental frequency by an octave (1:2 ratio) but also sometimes by a perfect fifth (2:3 ratio) relationship (Peter et al., 2015). However, this explanation provides a clear prediction which stimuli the birds should have responded to most, and it does not match the data. For example, second harmonic for C4 (261 Hz), for example, would roughly correspond with G5, and for C#4 this corresponds to C#5 and so on. We have provided a figure with the percent response across these training notes and their corresponding second harmonic. As can be seen, there seems to be little correlation in percent response across these pairs (see Figure 10).

In general, it appears to be rare that avian species treat stimuli at different absolute pitches as being similar. In the current study, the only thing that separated rewarded and unrewarded notes was their absolute pitch. In other

Figure 10. Data from phase 7 (generalization testing): Percent response to single notes for pooled data across all birds. Left: Octave four, right: Octave five. Error bars represent 95% confidence interval for the estimated probability. Notes and roughly corresponding second harmonics as marked by colour: C4/G5, C#4/G#5, D4/A5, D#4/A#5, E4/B5.

studies where other information was available, the birds preferred to use absolute pitch to solve the task. For example, several songbird species (*Mimus*, *Molothrus* and *Sturnus*) that learned to discriminate ascending from descending tone sequences failed to generalize this pattern to novel pitches (Hulse & Cynx, 1985) instead focusing on the absolute pitch of the notes. In addition, chickadees that learned to respond to consecutive 12 tones that were the same and not to 12 tones that were different were able to generalize this rule to almost all stimuli with which they were tested. However, they failed to respond accurately when tested with tones outside of the training range (Hoeschele et al., 2012b). But in some cases with more naturalistic stimuli, birds are able to generalize across pitch changes. For example, zebra finches (*Taeniopygia guttata*) generalize spoken words regardless of the speaker's sex (Ohms et al., 2010), so it could be interesting whether they achieve this by octave generalization. However, because the words are discriminable in other ways besides pitch, such as their spectral shape, which was not the case for the experiments described above where the birds failed to generalize, it is more likely that the birds simply ignored the pitch information in this task. New research supports this idea: Burgering et al. (2018) showed that zebra finches are able to categorize vowel-like sounds not only by pitch but also by spectral envelope. In general, it seems likely that these birds and other birds such as starlings that have been shown to generalize across pitch shifted conspecific vocalizations (Bregman et al., 2012), likely are solving these tasks using spectral shape (see Bregman et al., 2016; Hoeschele, 2017; Patel, 2017).

Despite the lack of evidence for octave equivalence across avian species and uncertainties about if and when exactly birds tend to extract pitch in a complex acoustic signal, we plan to continue using this paradigm to test other species for octave equivalence. Because minimal research has been conducted in this area even in mammalian species, it is still possible that octave equivalence perception is not common to all or even the majority of mammals. Conducting experiments using the current paradigm at hand with more mammalian species, especially species where there is already tentative support for octave equivalence, may therefore increase our understanding of the origins of octave equivalence. It may also be interesting to use the paradigm to test a reptilian or amphibian species, perhaps one that vocalizes. If octave equivalence perception is, in fact, a purely mammalian trait, amphibians and reptilians should not show this ability. If they do, octave

equivalence perception may be a rather basal trait, which has been lost in birds only.

Acknowledgements

This research and M.H. were funded by a Lisa Meitner Postdoctoral Fellowship (M 1732-B19) from the Austrian Science Fund (FWF). We thank Riccardo Hofer and Jinook Oh for technical assistance with the budgerigar apparatus. We thank Felix Haiduk, Asha Stewart, Julie Carpenter for additional assistance conducting the experiments. We thank Alexandra Christian and the rest of the animal care staff for their dedication to the care of our research subjects.

References

- Allen, O. (1967). Octave discriminability of musical and non-musical subjects. — *Psychon. Sci.* 7: 421–422.
- Baharloo, S., Johnston, P.A., Service, S.K., Gitschier, J. & Freimer, N.B. (1998). Absolute pitch: an approach for identification of genetic and nongenetic components. — *Am. J. Hum. Gen.* 62: 224–231. DOI:10.1086/301704.
- Berg, M., Fuchs, M., Wirkner, K., Loeffler, M., Engel, C. & Berger, T. (2017). The speaking voice in the general population: normative data and associations to sociodemographic and lifestyle factors. — *J. Voice* 31: 257.e13–257.e24. DOI:10.1016/j.jvoice.2016.06.001.
- Blackwell, H.R. & Schlosberg, H. (1943). Octave generalization, pitch discrimination, and loudness thresholds in the white rat. — *J. Exp. Psychol.* 33: 407–419.
- Braun, M. & Chaloupka, V. (2006). Carbamazepine induced pitch shift and octave space representation. — *Hear. Res.* 210: 85–92.
- Bregman, M.R., Patel, A.D. & Gentner, T.Q. (2012). Stimulus-dependent flexibility in non-human auditory pitch processing. — *Cognition* 122: 51–60.
- Bregman, M.R., Patel, A.D. & Gentner, T.Q. (2016). Songbirds use spectral shape, not pitch, for sound pattern recognition. — *Proc. Natl Acad. Sci. USA* 113: 1–6.
- Brittain-Powell, E.F., Dooling, R.J. & Gleich, O. (2002). Auditory brainstem responses in adult budgerigars (*Melopsittacus undulatus*). — *J. Acoust. Soc. Am.* 112: 999–1008.
- Brockway, B.F. (1964). Ethological studies of the budgerigar (*Melopsittacus undulatus*): non-reproductive behavior. — *Behaviour* 22: 193–222.
- Burgering, M.A., Vroomen, J. & ten Cate, C. (2018). Zebra finches (*Taeniopygia guttata*) can categorize vowel-like sounds on both the fundamental frequency (“pitch”) and spectral envelope. — *J. Comp. Psychol.*, in press. DOI:10.1037/com0000143.
- Burns, E.M. (1999). Intervals, scales, and tuning. — In: *The psychology of music*, 2nd edn. (Deutsch, D., ed.). Academic Press, San Diego, CA, p. 215–264.
- Cator, L.J., Arthur, B.J., Harrington, L.C. & Hoy, R.R. (2009). Harmonic convergence in the love songs of the dengue vector mosquito. — *Science* 323: 1077–1079.

- Celma-Miralles, A. & Toro, J.M. (2018). Beat perception in a non-vocal learner: rats can identify isochronous beats. — In: Proc. 12th int. conf. evol. lang. (Evolang12), p. 76-79.
- Cook, P., Rouse, A., Wilson, M. & Reichmuth, C. (2013). A California sea lion (*Zalophus californianus*) can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic. — J. Comp. Psychol. 127: 412-442.
- Cynx, J. (1993). Auditory frequency generalization and a failure to find octave generalization in a songbird, the European starling (*Sturnus vulgaris*). — J. Comp. Psychol. 107: 140-146.
- Demany, L. & Armand, F. (1984). The perceptual reality of tone chroma in early infancy. — J. Acoust. Soc. Am. 83: 687-695.
- Dent, M.L., Dooling, R.J. & Pierce, A.S. (2000). Frequency discrimination in budgerigars (*Melopsittacus undulatus*): effects of tone duration and tonal context. — J. Acoust. Soc. Am. 107: 2657-2664.
- Dooling, R.J., Best, C.T. & Brown, S.D. (1995). Discrimination of synthetic full-formant and sinewave/ra-la/continua by budgerigars (*Melopsittacus undulatus*) and zebra finches (*Taeniopygia guttata*). — J. Acoust. Soc. Am. 97: 1839-1846.
- Ficken, M.S., Ficken, R.W. & Witkin, S.R. (1978). Vocal repertoire of the black-capped chickadee. — Auk 95: 34-48.
- Gold, J.R., Nodal, F.R., Peters, F., King, A.J. & Bajo, V.M. (2015). Auditory gap-in-noise detection behavior in ferrets and humans. — Behav. Neurosci. 129. DOI:10.1037/bne0000065.
- Gramza, A.F. (1970). Vocal mimicry in captive budgerigars (*Melopsittacus undulatus*). — Z. Tierpsychol. 27: 971-983.
- Heffner, H., Koay, G. & Heffner, R. (2016). Budgerigars (*Melopsittacus undulatus*) do not hear infrasound: the audiogram from 8 Hz to 10 kHz. — J. Comp. Physiol. A. 202. DOI:10.1007/s00359-016-1125-9.
- Hile, A.G., Burley, N.T., Coopersmith, C.B., Foster, V.D. & Striedter, G.F. (2005). Effects of male vocal learning on female behavior in the budgerigar, *Melopsittacus undulatus*. — Ethology 111: 901-923.
- Hoeschele, M. (2017). Animal pitch perception: melodies and harmonies. — Comp. Cogn. Behav. Rev. 12: 5-18.
- Hoeschele, M., Weisman, R.G. & Sturdy, C.B. (2012a). Pitch chroma discrimination, generalization and transfer tests of octave equivalence in humans. — Attent. Percept. Psychophys. 74: 1742-1760.
- Hoeschele, M., Cook, R.G., Guillette, L.M., Hahn, A.H. & Sturdy, C.B. (2012b). Auditory same/different concept learning and generalization in black-capped chickadees (*Parus atricapillus*). — PLoS ONE 7: e47691. DOI:10.1371/journal.pone.0047691.
- Hoeschele, M., Weisman, R.G., Sturdy, C.B., Hahn, A. & Guillette, L. (2013). Chickadees fail standardized operant tests for octave equivalence. — Anim. Cogn. 16: 599-609.
- Hoeschele, M., Merchant, H., Kikuchi, Y. & Ten Cate, C. (2015). Searching for the origins of musicality across species. — Phil. Trans. Roy. Soc. Lond. B: Biol. Sci. 360: 1-9.
- Hove, M.J., Sutherland, M.E. & Krumhansl, C.L. (2009). Ethnicity effects in relative pitch. — In: Proc. 31st ann. conf. cogn. sci. soc., p. 2256-2261.

- Hulse, S.H. & Cynx, J. (1985). Relative pitch perception is constrained by absolute pitch in songbirds (*Mimus*, *Molothrus*, and *Sturnus*). — J. Comp. Psychol. 99: 176–196.
- Imig, T.J. & Morel, A. (1985). Tonotopic organization in lateral part of posterior group of thalamic nuclei in the cat. — J. Neurophysiol. 53: 836–851.
- Janik, V. & Slater, P.J. (1997). Vocal learning in mammals. — Adv. Stud. Behav. 26: 59–99.
- Kallman, H.J. (1982). Octave equivalence as measured by similarity ratings. — Percept. Psychophys. 32: 37–49.
- Kelley, L.A., Coe, R.L., Madden, R. & Healy, S.D. (2008). Vocal mimicry in songbirds. — Anim. Behav. 76: 521–528.
- Kline, C.A. & Dooling, R.J. (1984). Acoustic analysis of the budgerigar vocal repertoire. — J. Acoust. Soc. Am. 75(S1): 27.
- Köppl, C. (2001). Tonotopic projections of the auditory nerve to the cochlear nucleus angularis in the barn owl. — J. Ass. Res. Otolaryngol. 2: 41–53.
- Munshi-South, J. & Wilkinson, G. (2010). Bats and birds: exceptional longevity despite high metabolic rates. — Age. Res. Rev. 9: 12–19. DOI:10.1016/j.arr.2009.07.006.
- Nowicki, S. (1989). Vocal plasticity in captive black-capped chickadees: the acoustic basis and rate of call convergence. — Anim. Behav. 37: 64–73.
- Ohms, V.R., Gill, A., Van Heijningen, C.A., Beckers, G.J.L. & Ten Cate, C. (2010). Zebra finches exhibit speaker-independent phonetic perception of human speech. — Proc. Roy. Soc. Lond. B: Biol. Sci. 277: 1003–1009.
- Okanoya, K. & Dooling, R.J. (1987). Hearing in passerine and psittacine birds: a comparative study of absolute and masked auditory thresholds. — J. Comp. Psychol. 101: 7–15.
- Owren, M.J., Dieter, J.M. & Seyfarth, R.M. (1992). ‘Food’ calls produced by adult female rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques, their normally-raised offspring, and offspring cross-fostered between species. — Behaviour 120: 218–231.
- Patel, A.D. (2003). Language, music, syntax and the brain. — Nature Neurosci. 6: 647–681.
- Patel, A.D. (2017). Why doesn’t a songbird (the European starling) use pitch to recognize tone sequences? The informational independence hypothesis. — Comp. Cogn. Behav. Rev. 12: 19–32.
- Peter, B., Stoel-Gammon, C. & Kim, D. (2008). Octave equivalence as an aspect of stimulus-response similarity during nonword and sentence imitations in young children. — In: Proc. 4th int. conf. on speech prosody, SP 2008, international speech communications association, Campinas, p. 731–734.
- Peter, B., Larkin, T. & Stoel-Gammon, C. (2009). Octave-shifted pitch matching in nonword imitations: the effects of lexical stress and speech sound disorder. — J. Acoust. Soc. Am. 126: 1663–1666.
- Peter, B., Foster, B., Haas, H., Middleton, K. & McKibben, K. (2015). Direct and octave-shifted pitch matching during nonword imitations in men, women, and children. — J. Voice 29: 260.e21–260.e30.
- Reichmuth, C. & Casey, C. (2014). Vocal learning in seals, sea lions, and walruses. — Curr. Opin. Neurobiol. 28: 66–71.

- Richards, D.G., Wolz, J.P. & Herman, L.M. (1984). Vocal mimicry of computer-generated sounds and vocal labeling of objects by a bottlenosed dolphin, *Tursiops truncatus*. — J. Comp. Psychol. 98: 10-28.
- Scanlan, J. (1999). The function and significance of inter-species acoustic cues in the transformation of budgerigar (*Melopsittacus undulatus*) sounds into “speech”. — Int. J. Comp. Psychol. 12: 111-152.
- Tu, H., Osmanski, M.S. & Dooling, R.J. (2011). Learned vocalizations in budgerigars (*Melopsittacus undulatus*): the relationship between contact calls and warble song. — J. Acoust. Soc. Am. 129: 2289-2299.
- Tyack, P.L. (2008). Convergence of calls as animals form social bonds, active compensation for noisy communication channels, and the evolution of vocal learning in mammals. — J. Comp. Psychol. 122: 319-331.
- Weisman, R.G., Njegovan, M.G., Williams, M.T., Chenc, J.S. & Sturdy, C.B. (2004). A behavior analysis of absolute pitch: sex, experience, and species. — Behav. Proc. 66: 289-307.
- Wright, A.A., Rivera, J.J., Hulse, S.H., Shyan, M. & Neiworth, J.J. (2000). Music perception and octave generalization in rhesus monkey. — J. Exp. Psychol. 129: 291-307.
- Wyndham, E. (1980). Diurnal cycle, behaviour and social organization of the budgerigar *Melopsittacus undulatus*. — Emu — Austral. Ornithol. 80: 25-33.
- Zaretsky, M.D. & Konishi, M. (1976). Tonotopic organization in the avian telencephalon. — Brain Res. 111: 167-171.