

Pitch shifting does not disrupt song recognition in the European Starling

Micah R. Bregman¹, Aniruddh D. Patel², Timothy Q. Gentner³

¹Dept. of Cognitive Science, UC San Diego; ²The Neurosciences Institute; ³Dept. of Psychology, UC San Diego

ABSTRACT

Musical absolute pitch is rare among humans, who attend primarily to the relationship between pitches even if they have some memory for absolute pitch. In contrast, previous research has suggested that songbirds perceive relative pitch only in limited contexts, and are only able to recognize novel stimuli on the basis of relative pitch within a range constrained by the absolute pitch of the training stimuli.

Our results suggest that the European starling, a species of songbird is able to recognize conspecific song even when the absolute pitch is shifted while preserving relative pitch. This ability to generalize was unconstrained by the training range. Strikingly, there was no difference in acquisition rate between a training set consisting of 64 stimuli (16 shifted versions of 4 song excerpts) and the training set containing only the original 4. Moreover, both groups of subjects were able to recognize novel pitch-shifted exemplars both within and outside of the training range. These results suggest that the full extent of avian pitch processing has not been fully explored, especially within the context of more natural stimuli.

1. Introduction

Many acoustic features allow us to recognize environmental and musical sounds. We use pitch, timbre, temporal cues and spatial location, often simultaneously. In the context of music and language, it is generally not the absolute pitch of sounds that drive recognition. We have no difficulty understanding a word spoken by two different individuals, even in a tonal language, since it is the pitch contour and relationship between pitches rather than the absolute frequency that is used to convey intonational and lexical meaning. With rare exception, adult humans are biased to attend to the relationship between pitches. Most adult humans are effortlessly able to recognize two versions of a melody as the same if all intervals between its notes are preserved, even when the starting pitch is different. In fact, even musically trained individuals with excellent relative pitch often lack the ability to provide the label (note name) of individual pitches. In contrast, previous studies in other species, particularly songbirds, suggest that absolute pitch processing is the dominant strategy. These studies suggest that relative pitch processing in songbirds is limited, and often constrained by the absolute pitch range of the training range (Hulse & Cynx, 1985; Cynx, Hulse, & Polyzois, 1986). Absolute pitch processing in other species is, it should be emphasized, most accurately compared not to musical AP, but rather to excellent long-term absolute pitch memory and representation.

In contrast to their facility using relative pitch to recognize melodies, relatively few humans possess reliable absolute pitch. The rarity of the ability to produce absolute pitch labels is puzzling because a subset of individuals possesses excellent musical absolute pitch (AP). These individuals are able to identify pitches without an external reference and yet do not seem to have any deficit in the processing of pitch relationships. Musical AP is almost never observed without childhood musical training and is rare, although more prevalent among speakers of tone languages (Deutsch, Henthorn, Marvin, & Xu, 2006).

Although musical AP does not occur among those without musical training, some form of absolute pitch memory is more widespread. It has been observed repeatedly in contexts where listeners have tens, or in some cases thousands of cases of repeated exposure to a stimulus over several months or years. When listeners are asked to hum or sing a very familiar song, the median of the distribution of starting tones produced by participants is close to the original source (Levitin, 1994). Similarly, individuals without AP are nonetheless able to recognize a transposed version of familiar television theme songs as being not at normal pitch. Participants achieved above chance performance with a one semitone shift and nearly 70 percent accuracy when the pitch was shifted by two semitones (Schellenberg & Trehub, 2003). Musically untrained listeners were able to identify whether a dial tone was at the correct pitch, too high or too low without any context (Smith & Schmuckler, 2008). Nevertheless, these participants did not possess musical absolute pitch, and are unable to attach specific pitch labels to the notes they recognize or produce.

Evidence for relative pitch processing in non-human species has been limited, although rhesus macaques (Hulse et al, 2000) and ferrets (Yin, Fritz, & Shamma, 2010) appear to recognize transposed tone sequences in some contexts. Pitch processing has been most studied among songbirds, which use song in the wild for mate attraction, territorial displays and conspecific recognition. Laboratory studies investigating songbird pitch perception using artificial tone sequences have suggested a limited ability to use relative pitch information for discrimination, but one which is often overridden by salient absolute pitch cues (Hulse & Cynx, 1985). Experiments using tone sequences have suggested, broadly, that songbirds are biased to use absolute pitch information under many conditions. Within a constrained training range, songbirds can eventually learn to generalize on the basis of relative pitch after many trials. However, even after doing so, they continue to readily generalize to repeated sequences of a single tone on the basis of absolute pitch. The relative pitch perception strategies that emerge after training do not transfer outside of the absolute frequency range of the training stimuli. After transferring subjects to stimuli outside of the training range, subjects took as long to learn the new discrimination as they did to learn the original

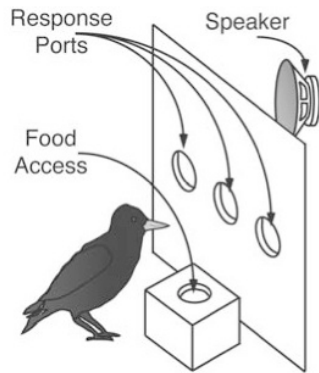


Figure 1: Schematic of the operant training panel.

sequence discrimination (Hulse & Cynx, 1985; Cynx, 1995). Pigeons, which are not songbirds, also seem to have a similar tendency. They are able to learn discriminations using either absolute or relative features, but have a dominant aptitude for absolute pitch processing (Murphy & Cook, 2008). When songbirds were trained to discriminate short tone sequences, they also failed to show any generalization to transpositions of an octave (Cynx, 1993). Perhaps this dominant bias arises because songbird natural communication signals tend to have stable absolute pitch across multiple bouts, making absolute pitch a useful cue for individual identity.

Although some evidence for relative pitch processing has been observed in songbirds, other data suggests relative pitch perception does not arise easily. European Starling subjects never successfully learned to discriminate between a large 64 exemplar set of ascending and descending tone sequences even after extensive training when absolute pitch cues were not informative—a task that would be very easy for humans. After successfully learning to classify a smaller set of stimuli, however, starling subjects were able to generalize to novel exemplars in the larger set (Page, Hulse, & Cynx, 1989).

Experiments investigating song production and perception have also found evidence suggesting a role for relative pitch in the vocal repertoires of some avian species (Weisman, Ratcliffe, Johnsruide, & Hurly, 1990; Christie, Mennill, & Ratcliffe, 2004). In the wild, Chickadees shift the absolute pitch of their call, but preserve the relative pitch between the “fee” and “bee” notes (Shackleton, Ratcliffe, & Weary, 1992).

To investigate the role of absolute and relative pitch perception in a more natural context, we used stimuli derived from recorded starling song rather than tone sequences. We trained subjects to recognize conspecific songs under one of two conditions: the first condition required similar discriminations to those made to perform individual recognition in the wild. For this task, song pitch was stable. We trained other subjects with the same set of songs shifted so that they were heard at a wide range of absolute pitch levels. After training, we tested their ability to maintain this recognition when songs were shifted to novel pitch levels. The results of these experiments suggest that songbirds, such as the

European Starling, do use pitch to some degree in song recognition but that their recognition performance remains high even when song is shifted. Because previous work suggested that any relative pitch ability in songbirds is tightly constrained by the pitches heard during training, we tested subjects on their ability to recognize novel stimuli shifted both within and outside of the training pitch range.

2. Stimuli

We tested subjects using several stimulus sets, each designed to investigate a different aspect of pitch perception. The method used to present each stimulus set was similar, and is described in the procedure section.

Pitch shifted song. To investigate the effect of altering the absolute pitch of song on recognition, we used a phase vocoding algorithm to digitally shift the pitch of starling song while keeping temporal patterning and the relative pitch between song notes constant. Our source material for these stimuli was four excerpts of starling song (each 14.8-15.0 seconds in length). The song excerpts were recorded in the laboratory from two individual males. We used these excerpts to create two stimulus sets. The first set contained all four exemplars presented at the natural recorded pitch. The second set consisted of 64 exemplars, 16 pitch-shifted versions of each of the four stimuli in the first training set shifted by up to $\pm 30\%$ in increments of 4%. Subjects were trained using either the 4 exemplar or 64 exemplar training set and were later tested to evaluate the degree to which shifts in pitch disrupted song recognition.

Probe stimuli consisted of untrained pitch-shifted song stimuli both from within the training range (within $\pm 30\%$ of the source recordings) and outside the training set (between $\pm 30\%$ and $\pm 40\%$ relative to the source recordings).

Absolute pitch stimuli. To investigate whether subjects are able to attend to the absolute pitch of conspecific song, we designed a stimulus set so that absolute pitch is the only cue available to achieve good discrimination performance. All exemplars were digitally pitch-shifted versions of the same 15 second song excerpt. Two of the exemplars were shifted by -24% and -28% relative to the original recording, while the other two were shifted by $+24\%$ and $+28\%$. In this task, absolute pitch was the only cue enabling subjects to discriminate between the two groups and the task was designed as a 2-alternative-choice (2AC) task.

Probe stimuli consisted of 11 exemplars spaced equally between -20% and 20% in increments of 4% relative to the source recording.

Transposed Melodies. To investigate the effect of altering absolute pitch on melody recognition, we trained subjects to discriminate between the opening phrase of two melodies, “Old Susanna” (3.78 seconds) and “London Bridge” (3.97 seconds) synthesized using a midi piano timbre. After learning this discrimination, subjects were tested using probe stimuli consisting of transposed versions of the learned melodies. These test stimuli were transposed by 1-6 semitones.

3. Subjects

A total of 7 wild-caught European Starlings (*Sturnus vulgaris*) participated in these experiments. After capture, subjects were housed with conspecifics in a large aviary. The photoperiod in both the aviary and experimental apparatus correspond to local sunrise and sunset times. All procedures were conducted as part of a protocol approved by the UCSD Institutional Animal Care and Use Committee.

Four subjects participated in the pitch shifted song experiments. One of these subjects, as well as an additional naïve animal participated in the transposed melody recognition task. Two naïve animals participated in the absolute pitch experiment.

4. Procedure

We placed subjects inside a wire cage located in a sound isolation chamber. One side of the cage allows access to an operant panel consisting of three response ports and a food hopper (Figure 1). In order to become familiar with the operant device, subjects are first trained to peck the center response port to receive four seconds of access to the food hopper. After learning to obtain food by pecking the response port, stimuli are introduced.

We trained subjects using two procedures: go-nogo (GNG) and 2-alternative-choice (2AC). For all subjects, the operant training procedure was separated into three sessions: initial training, transfer and probe testing. During initial training, a peck to the center response port initiated a trial and a stimulus was played from a speaker located behind the operant panel. After stimulus playback, a 4 second response window began, and pecks to the response ports during this window were recorded. For GNG training, a peck to the center response port following a “go” stimulus was immediately rewarded by two second of food access. A peck to the response port following a “no-go” stimulus was immediately punished by a brief period during which the cage

lights were extinguished and a trial cannot be initiated.

For the 2AC procedure, after stimulus playback subjects pecked either the left or right response ports. Each response was rewarded if the correct response port for a given stimulus was pecked and was punished when the incorrect port was pecked.

No responses were recorded or reinforced after the 4 second response window ended. After the response window ends, there is a 2 second inter-trial-interval during which a trial cannot be started.

During the transfer block, the trial procedure was identical, but reinforcement was reduced to occur randomly on 60% of trials. This increases the number of trials subjects perform per day and also introduces uncertain reward contingencies, important for later unreinforced probe trials to be presented undetected.

During the probe testing session, we presented two stimulus sets simultaneously. We continued to present the learned stimuli heard during training on 80% of trials and continued to reinforce correct and incorrect responses on 60% of these trials. On the other 20% of trials, we presented novel probe stimuli, which were not reinforced differentially for correct or incorrect responses. Our goal when introducing probe trials was to observe the subject’s classification of novel stimuli within the context of the classification task they learned during training. We reinforced probe trials randomly and with low probability, providing food reward on 10% of trials, time-outs on 10% of trials, and no reinforcement on 80% of trials.

5. Data Analysis

We evaluated subject’s discrimination performance using the d-prime measurement. D-prime provides a measure of discrimination that is unbiased by response rate and reflects both the correct responses to go stimuli and the false responses to no-go stimuli. Although statistically the threshold for above chance performance depends on response rate, d-prime values greater than one can be interpreted in these experiments as reflecting discrimination well above chance. When means are reported, we also report the standard error of the mean.

6. Results

Starlings rapidly learn to recognize pitch-shifted song.

Discrimination accuracy improved rapidly when measured using d-prime (Figure 2). For these four subjects, there was no difference in acquisition rate between subjects trained with the 64-exemplar stimulus set and those trained with the 4-exemplar set. In operant song recognition tasks it is common to observe an increase in acquisition time when the size of the stimulus set is increased. The observed lack of increase in acquisition time when introducing 16 times as many stimuli suggests that pitch shifting song has little impact on the difficulty of song recognition, either because starlings are naturally able to generalize and recognize song presented at novel shifted pitch levels, or because they are attending primarily to other (non-pitch) song features such as temporal modulation or timbre changes.

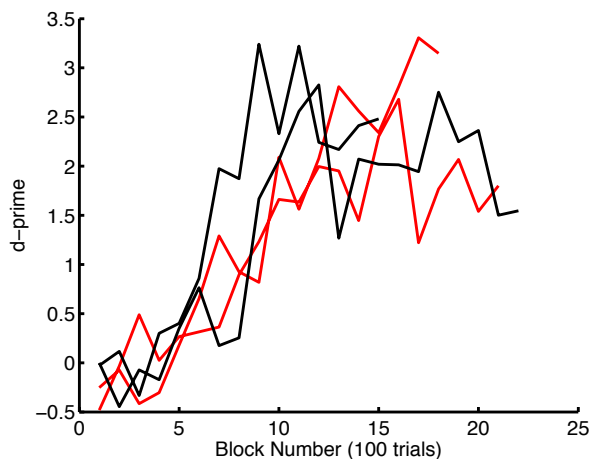


Figure 2: Acquisition rate for all four subjects. Subjects trained on 64-exemplar stimulus set in red, while subjects trained on 4-exemplars set in black.

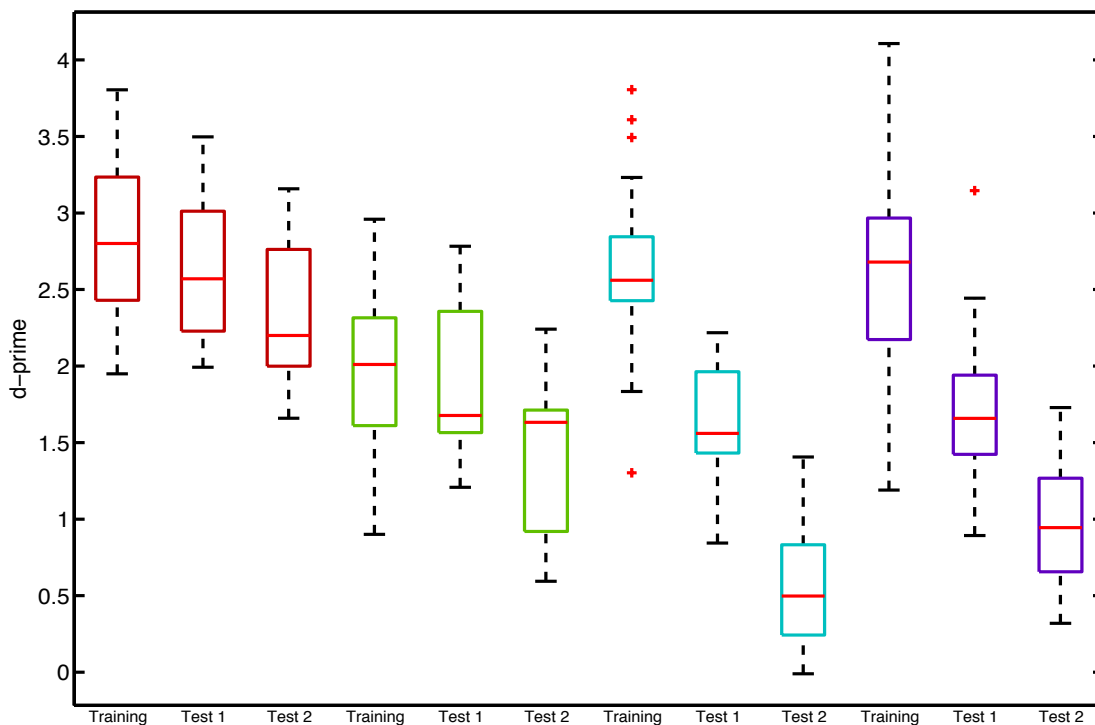


Figure 3: Generalization performance for four subjects on Test Set Data. Subjects recognize shifted starling songs both within (Test 1) and outside (Test 2) the absolute training range. Two subjects (colored red and green) were trained with the 64-exemplar set; Two subjects (colored blue and purple) were trained with the 4-exemplar stimulus set.

Subjects generalize to recognize novel pitch-shifted song stimuli, regardless of training. Subjects trained on both the 64-exemplar stimulus set and the 4-exemplar set recognized novel exemplars shifted by amounts up to $\pm 40\%$. Figure 3 shows performance on the learned baseline stimulus set compared to two novel test sets. Test set 1 contained novel stimuli shifted within the absolute pitch range range ($\pm 30\%$) of the training stimuli. Test set 2 consisted of stimuli shifted between 30% and 40%. Although discrimination performance is lower for stimuli shifted outside of the training range, all subjects were well above chance for all test sets, with the worst-performing subject on Test set 2 still achieving a mean d' well above zero (Mean=0.975, $t=14.07$, $df=34$, $p<10^{-8}$).

Generalization performance for novel stimuli is enhanced by familiarity with shifted song stimuli. Subjects trained with both 64-exemplar and 4-exemplar stimulus sets recognized shifted novel pitch-shifted stimuli with above chance. Subjects that had experience with the 64-exemplar set, however, performed more accurately on tests with novel stimuli than those subjects trained with a 4-exemplar stimulus set. This difference can be observed in the significant interaction between training condition (4-exemplar or 64-exemplar) and generalization performance to stimuli in novel pitch ranges ($df=2$, $F=22.49$, $p<.0001$). Main effects of Shift amount ($df=2$, $F=74.85$, $p<.0001$) and training condition ($df=1$, $F=19.12$, $p<.0001$) were also statistically significant.

Subjects are sensitive to absolute pitch cues in natural song. Our results showing successful generalization to novel pitch levels could be explained in two ways. Either subjects hear and attend to pitch differences, but are able to recognize that the absolute pitch differences are not meaningful predictors of song identity, or the

subjects are not attending to pitch features in natural stimuli at all. In order to verify that subjects are able to make use of the absolute pitch features present in song, we trained them using the psychometric pitch stimuli discussed earlier where they learned to discriminate between “high” and “low” versions of the same song excerpt. We found that subjects are able to learn this task, reaching excellent performance (mean $89.1 \pm 1.67\%$ correct, last 3 blocks) suggesting they are able to attend to absolute pitch features in

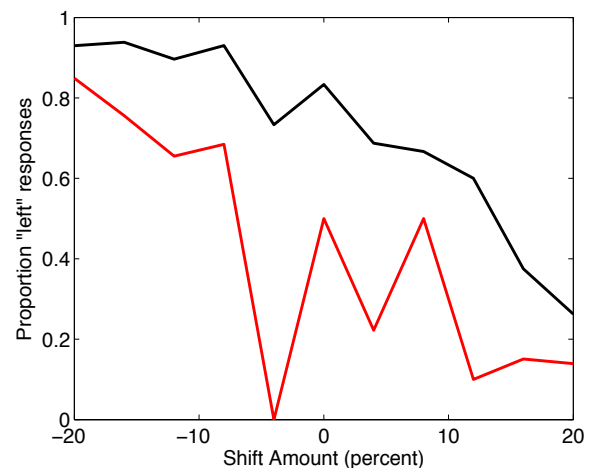


Figure 4: Responses to probe stimuli after two subjects were trained to respond “left” to low-pitch versions (-24%, -28%) of a song and “right” to high-pitch versions (+24%, +28%) of the same song. Each line shows the proportion of times that the subject responded “low” to test stimuli shifted to cover the range between $\pm 20\%$.

natural song stimuli. These data are consistent with previous studies observing absolute pitch perception in songbirds using tone sequence stimuli (Cynx, 1995).

Testing subjects using probe stimuli presented at intermediate pitches (figure 4) produced responses consistent with using song pitch directly in recognition. Song shifted by -10%, for example, elicited a majority of responses associated with the “low pitch” training stimulus

Recognition of transposed melodies was inconsistent. After learning to discriminate between two melodies (“Old Susanna” and “London Bridge”) by pecking for one melody and learning to withhold response for the other, two subjects were tested with versions transposed between 1-6 semitones. Figure 5 shows that transposition (even by as little as one semitone) caused performance for one subject to fall to chance, while the other subject’s performance remained above chance. The second subject, however, did show a significant performance decrease (from mean=2.137 to mean=.786, Mann-Whitney U-test, $z=-5.83$, $p<.00001$) indicating that transposition continued to have a dramatic effect on melody discrimination performance.

7. Discussion

Previous work found only elusive evidence for relative pitch perception in songbirds even after extensive training using tone sequence stimuli. In contrast, our data suggests that starlings (a species of songbird) are able to recognize species-specific songs that have been shifted to novel pitch ranges without explicit training or experience with pitch-shifted stimuli (although such training does improve performance).

Our results suggest that starlings are able to attend to pitch features and use them for recognition of conspecific song. When pitch is shifted to novel pitch ranges, both within and outside of the absolute frequency of the training range, subject’s discrimination performance remained high. During training, we exposed birds to two different stimulus sets, one of which contained shifted song and the other did not. There were performance differences between the two groups, with subjects exposed to shifted song during training exhibiting somewhat better generalization performance. This suggests that birds both hear and recognize that the song has been shifted, but their recognition performance is not dramatically impaired.

Perhaps our subject’s good performance can be attributed to the wealth of other features available as recognition cues in the case of conspecific song. Starling song varies in timbre, amplitude and pitch across time. Although starlings are able to use pitch in song recognition (as demonstrated by our explicit training using “high” and “low” pitch songs), they may be easily able to use other features when pitch is an unstable cue. This suggests that the processes underlying conspecific song recognition may not be entirely observable by investigating responses to tone sequence stimuli such as those used in previous experiments.

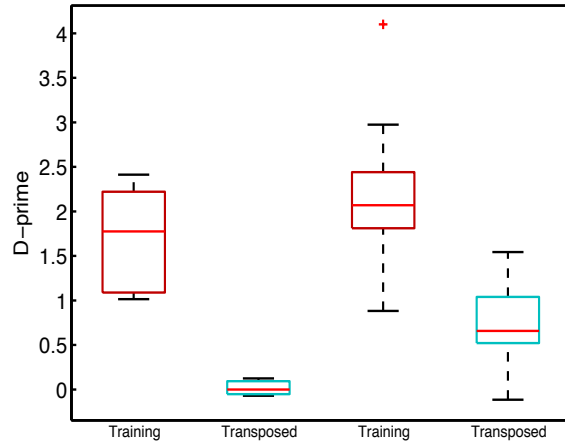


Figure 5: Discrimination between two human melodies. Performance on trained stimulus set and transposed melody set is shown for data collected during the test phase. One subject performed at chance when recognizing transposed melodies, while the other maintained above-chance discrimination performance.

Acknowledgments: ADP was supported by Neurosciences Research Foundation as part of its program on music and the brain at The Neurosciences Institute.

8. References

- Christie, P. J., Mennill, D. J., & Ratcliffe, L. M. (2004). Pitch shifts and song structure indicate male quality in the dawn chorus of black-capped chickadees. *Behavioral Ecology and Sociobiology*, 55, 341-348.
- Cynx, J., Hulse, S. H., & Polyzois, S. (1986). A psychophysical measure of pitch discrimination loss resulting from a frequency range constraint in European starlings (*Sturnus vulgaris*). *Journal of experimental psychology. Animal behavior processes*, 12, 394-402.
- Cynx, J. (1993). Auditory frequency generalization and a failure to find octave generalization in a songbird, the European starling (*Sturnus vulgaris*). *Journal of Comparative Psychology*, 107, 140-146.
- Cynx, J. (1995). Similarities in absolute and relative pitch perception in songbirds (starling and zebra finch) and a nonsongbird (pigeon). *Journal of Comparative Psychology*, 109, 261-267.
- Deutsch, D., Henthorn, T., Marvin, E., & Xu, H. (2006). Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period. *The Journal of the Acoustical Society of America*, 119, 719.
- 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.
- Hulse, S. H., Wright, A. A., Rivera, J. J., Shyan, M., & Neiworth, J. J. (2000). Music perception and octave generalization in rhesus monkeys. *Journal of experimental psychology. General*, 129, 291-307.
- Levitin, D. J. (1994). Absolute memory for musical pitch: evidence from the production of learned melodies. *Perception & Psychophysics*, 56, 414-23.
- Murphy, M. S., & Cook, R. G. (2008). Absolute and relational control of a sequential auditory discrimination by pigeons (*Columba livia*). *Behavioural processes*, 77, 210-22.
- Schellenberg, E., & Trehub, S. (2003). Good Pitch Memory Is Widespread. *Psychological Science*, 14, 262-266.
- Smith, N. A., & Schmuckler, M. A. (2008). Dial A440 for absolute pitch: absolute pitch memory by non-absolute pitch possessors. *The Journal of the Acoustical Society of America*, 123, EL77-84.
- Shackleton, S., Ratcliffe, L., & Weary, D. M. (1992). Relative Frequency Parameters and Song Recognition in Black-Capped Chickadees. *The Condor*, 94, 782-785.
- Weisman, R. G., Ratcliffe, L., Johnsrude, I., & Hurly, T. (1990). Absolute and Relative Pitch Production in the Song of the Black-Capped Chickadee. *The Condor*, 92, 118-124.
- Yin, P., Fritz, J. B., & Shamma, S. A. (2010). Do ferrets perceive relative pitch? *The Journal of the Acoustical Society of America*, 127, 1673-80.