Title: Vocal performance increases rapidly during the dawn chorus in Adelaide's warbler (*Setophaga adelaidae*)

Abbreviated title: Vocal warm-up during the dawn chorus

Juleyska Vazquez-Cardona [1], Tyler R. Bonnell [1], Peter C. Mower [1], Orlando J. Medina [2], Hester Jiskoot [3], and David M. Logue [1,4]

1. Department of Psychology, University of Lethbridge, Lethbridge, AB, Canada;
2. United States Fish and Wildlife Service, Cabo Rojo National Wildlife Refuge, PR, USA;
3. Department of Geography and Environment, University of Lethbridge, AB, Canada;
4. Departamento de Biología, Universidad de Puerto Rico, Mayagüez, PR, USA

Abstract

Many songbirds sing intensely during the early morning, resulting in a phenomenon known as the dawn chorus. We tested the hypothesis that male Adelaide's warblers (*Setophaga adelaidae*) warm up their voices during the dawn chorus. If warming up the voice is one of the functions of the dawn chorus, we predicted that vocal performance would increase more rapidly during the dawn chorus compared to the rest of the morning and that high song rates during the dawn chorus period contribute to the increase in vocal performance. The performance metrics *recovery time*, *voiced frequency modulation*, and *unvoiced frequency modulation* were low when birds first began singing, increased rapidly during the dawn chorus, and then leveled off or gradually diminished after dawn. These changes are attributable to increasing performance within song types. Reduction in the duration of the silent gap between notes is the primary driver of improved performance during the dawn chorus. Simulations indicated that singing at a high rate during the dawn chorus period increases performance in two of the three performance measures (*recovery time* and *unvoiced* *frequency modulation*) relative to singing at a low rate during this period. These findings are consistent with the hypothesis that vocal warm-up is one benefit of participation in the dawn chorus.

Keywords

bird song, dawn chorus, practice, vocal performance, warming up

Many songbirds sing at elevated rates early in the morning. Several hypotheses attempt to explain the adaptive function of the "dawn chorus" (Staicer et al. 1996, Gil and Llusia 2020). For example, birds may sing vigorously at dawn because (1) sound propagates further at dawn, (2) foraging is not profitable at dawn, (3) song re-establishes territorial claims at dawn, (4) females prospect for mates at dawn, or (5) they have extra energy they did not spend the night before (Staicer et al. 1996, Gil and Llusia 2020). It is possible that the dawn chorus serves multiple functions, within and among species. In the present study, we test Schraft et al.'s (2017) hypothesis that participation in the dawn chorus plays an important role in "warming up" the vocal apparatus.

We use the term "warm-up" to describe improvement in performance due to recent practice (Schraft et al. 2017, Déaux et al. 2020, Dinh et al. 2020). Various physiological process(es) might cause such a change in performance. For example, physical warming might underlie behavioral warm-up by increasing the speed of nerve conduction or metabolic reactions (Tan and Knight 2018, Bishop 2003a, Bishop 2003b). A prediction of that hypothesis is that performance will be lower at low temperature.

We define performance as an adaptive approach to a behavioral limit (Podos 1997). Studies show that warm-up improves performance in human (*Homo sapiens*) singers and athletes (Amir et al. 2005, Fradkin et al. 2010). Recent evidence suggests that singing birds' vocal performance also improves after warming up (Schraft et al. 2017, Dinh et al. 2020). The need to warm up the voice may be an important constraint on vocal signaling in birds because vocal performance influences responses from both male and female signal receivers (Podos and Sung 2020). According to the warm-up hypothesis, ancestral birds that warmed up their voices in the morning by singing at high rates had a competitive advantage over those that did not, sparking an evolutionary arms race that drove elaboration of the dawn chorus (Schraft et al. 2017).

The first study to link the dawn chorus to improvements in vocal performance showed that Frequency Excursion (*FEX*) – which measures the speed of frequency modulation in both notes and silent gaps within songs (Podos et al. 2016) – increases with the cumulative number of songs that a male Adelaide's warbler (*Setophaga adelaidae*) sings over the course of the morning (Schraft et al. 2017). That study relied on a linear model, so it did not provide a detailed description of changes in vocal performance over the course of the morning. If the dawn chorus functions as a warm-up period, we expect that performance will be low when birds first start singing, increase rapidly during the dawn chorus, and level off after the end of the dawn chorus. In the present study, we replaced the wholistic performance metric FEX with three fine-grained performance metrics (*voiced frequency modulation, unvoiced frequency modulation*, and *recovery time)*, each of which attempts to describe the performance of a particular class of vocal gestures (Goller 2022).

Constrained relationships between the acoustic properties of vocal signals can reveal the axes of vocal performance. Singing requires rapid, precise changes in the syrinx, respiratory system, and upper vocal tract (Nowicki et al. 1992, Catchpole and Slater 2003, Podos and Nowicki 2004). Constraints on the speed, precision, or coordination of these physiological changes set limits on song structure (Hoese et al. 2000, Podos and Nowicki 2004, Plummer and Goller 2008). Logue et al. (2020) described performance constraints in Adelaide's warbler at the note level, where notes are uninterrupted sound units.

The following example illustrates how acoustic trade-offs can provide evidence of note-level performance constraints. If there is a constraint on the speed that an Adelaide's warbler can modulate the fundamental frequency of a note, then at the limit of performance, notes with large frequency bandwidths will be longer in duration than notes with small frequency bandwidths (Logue et al. 2020). A scatterplot of duration versus frequency bandwidth would reveal a roughly triangular distribution in which the hypotenuse represents the constrained edge of the distribution

(Podos 1997). Quantile regression can be used to estimate the performance limit (the solid line in Fig. 1), making it possible to calculate the orthogonal distance ("deviation score") between a sound and the performance limit in acoustic space (Fig 1; Podos 2001, Wilson et al. 2014). Deviation scores are inverse measures of performance, so low scores indicate high performance and *vice versa* (Podos 2001).

Logue et al. (2020) used note-level acoustic constraints to calculate three axes of vocal performance in Adelaide's warblers' songs: *recovery time*, voiced frequency modulation (*voiced FM*), and unvoiced frequency modulation (*unvoiced FM*). *Recovery time* is the deviation score from a plot of *note duration* versus the duration of the subsequent silent gap (*gap duration*). At the limit of performance, longer notes require longer subsequent gaps, presumably because birds require long mini-breaths to replenish the air that they exhaled singing a long note (Hartley and Suthers 1989, Suthers and Zollinger 2004, Cardoso et al. 2007). *Recovery time* is strongly correlated with the song-level performance metric "percent of sound" (a.k.a., "acoustic density;" Cardoso et al. 2009, Logue et al. 2020). The performance axis *voiced FM* describes the speed with which birds can change frequency while voicing a note. It is the deviation score from a comparison of frequency bandwidth (*note BW*) and duration (*note duration*; Logue et al. 2020). Similarly, the metric *unvoiced FM* describes the speed with which birds can change frequency while they are not voicing a note. It is the deviation from a plot of the frequency bandwidth of silent gaps (*gap BW*) versus the *gap duration* (Geberzahn and Aubin, 2014). *Voiced* and *unvoiced FM* are constrained because birds require more time to make large frequency jumps than small ones, indicating a limit to the speed of frequency modulation. In the present study, we use these three metrics and the simple acoustic variables that underlie them to characterize changes in vocal performance during morning singing.

Our study tests the hypothesis that the dawn chorus functions to warm up the voice. Relative to the previous study of vocal warm-up in Adelaide's warbler (Schraft et al. 2017), we used a larger dataset, more granular measures of performance (Logue et al. 2020), and a modeling approach that reveals fine-scale changes in performance over time. If participation in the dawn chorus functions to warm up the voice, we predict that (1) vocal performance will increase more rapidly during the dawn chorus compared to later in the morning, and (2) singing at a high rate during the dawn chorus period will lead to higher vocal performance after dawn.

Methods

Study species

Adelaide's warbler is a tropical, year-round territorial, socially monogamous New World warbler (family: Parulidae), endemic to Puerto Rico and nearby islands (Toms 2020). Males sing repertoires of 22.6 ± 2.6 song types, all of which are frequency modulated trills (Fig. 2; Staicer 1992). Each song comprises 23.8 ± 4.7 notes (unpublished analysis from Logue et al. 2020).

Recording and annotation

We studied a population of Adelaide's warblers at the Cabo Rojo National Wildlife Refuge (U.S. Fish and Wildlife Service; 17.98° N, 67.17° W) in western Puerto Rico from March 3 – June 19, 2012, and April 13 – May 6, 2017. These dates are within the population's breeding season (Staicer 1992). Birds were captured with acoustic lures and mist nets and fitted with unique combinations of plastic colored leg bands and U.S. Fish and Wildlife Service metal leg bands for identification. Recordings were made with Marantz PMD-661 digital recorders and Sennheisser ME67 "shotgun" microphones (file format = wav, sampling rate = 44.1KHz, bit depth = 16 bits). Observers continuously recorded mated males from 45 minutes before sunrise until approximately 2.5 hours after sunrise (n = 9499 songs from 9 males in 2012 and n = 16381 songs from 18 males in 2017). The recordist visually confirmed the identity of the focal male by the end of each recording session.

Trained observers annotated the field recordings. Annotators visualized the recordings from 2012 in Syrinx PC v.2.6 (Burt 1995-2006) and entered data into a spreadsheet. Data included all songs from the focal birds, and the song type (*ST*) to which they belonged. Annotators scored *ST* by comparing each song spectrogram to spectrograms of the known repertoire of the focal bird. After all scoring was complete, one observer (author D.M.L.) reviewed all song types. It is straightforward to assign songs to *STs* within an individual's repertoire, and all our analyses treat *ST* as a within-individual variable. Data from the 2012 recordings were used previously in Schraft et al. (2017), Hedley et al. (2018), Kaluthota et al. (2019), and Logue et al. (2020). The recordings from 2017, which are unique to this study, were annotated in Raven Pro 1.6.1 (Center for Conservation Bioacoustics 2019). The initial round of *ST* scoring followed the 2012 protocol. After scoring was complete, one observer (author P.C.M.) reviewed all song types. Annotators scored the time of song delivery in five second increments (e.g., 05:36:00, 05:36:05, etc.) for the 2012 recordings, but they scored exact times for the 2017 recordings. We standardized temporal precision across recording years by binning the 2017 dataset into five-second intervals.

This study adheres to the guidelines from the Institutional Animal Care and Use Committee at the University of Puerto Rico, Mayaguez (September 17, 2010) and the Animal Welfare Committee at the University of Lethbridge (protocol #1605). It also follows the ASAB/ABS Guidelines for the use of animals in research. Field work was conducted with permission from the U.S. Fish and Wildlife Service (permits 2012-01, 41521-2016-11) and the Departamento de Recursos Naturales y

Ambientales (permit 2016-IC-068-1). Bird handling was conducted under D.M.L.'s master bird banding license (no. 2399).

Acoustic analysis

We acoustically analyzed a subset of songs from the recordings. For the 2012 recordings, we analyzed all songs with sufficiently high signal-to-noise ratios (n = 2879 songs). For the 2017 recordings, we attempted to randomly select ten songs from each hour of recording. For both years, our threshold for selection was that no other loud sounds overlapped the song, and that the notes of the song were sufficiently loud relative to the background noise that Luscinia could detect them. For hours that did not include ten song recordings with sufficiently high signal-to-noise ratios, we sampled as many as were available (n = 888 songs sampled).

We conducted acoustic analysis in Luscinia v.2.14(Lachlan 2007; Settings: max. freq. = 10 kHz, frame length = 5 ms, time step = 1 ms, dynamic range = 35 dB, dynamic equalization = 100 ms, de-reverberation = 100%, de-reverberation range = 100 ms, high pass threshold = 1.0 kHz, noise removal = 10 dB). Luscinia applies de-reverberation and noise removal algorithms to the sound prior to note identification (personal comment R. Lachlan). We chose conservative de-reverberation and noise removal parameters that appeared to reduce reverberation and noise without affecting the notes. Trained scorers used a stylus and a touchscreen monitor to highlight all visible notes on a spectrogram. Luscinia's algorithm then searched within the highlighted area for pixels that exceed an amplitude threshold, which is defined by the dynamic range parameter (35 dB in the present analysis). The algorithm then adds neighboring pixels to the growing note until it encounters either the edge of the highlighted area, or the amplitude falls below the cutoff.

We extracted the following metrics at the note level: maximum peak frequency (highest peak frequency in the note), minimum peak frequency (lowest peak frequency in the note), note start time, note end time, peak frequency at the beginning of the note, and peak frequency at the end of the note. Peak frequency is the frequency with the highest amplitude in the sample window. In Adelaide's warbler songs, the peak frequency is the fundamental frequency.

Note start and end times were used to calculate the duration of the notes (*note duration*) and the silent gaps after each note (*gap duration*). We calculated the frequency bandwidth of the notes (*note BW*) as the ratio of the maximum and minimum peak frequency of the note (Cardoso 2013, Logue et al. 2020). Similarly, we calculated *gap BW* as the ratio of the peak frequency at the end of one note and the beginning of the next note, with the larger value in the numerator (Logue et al. 2020). We excluded the final note of each song from the analysis because it lacks a *gap BW* and *gap duration*.

Variables

We calculated deviation scores for three performance metrics. We first generated three scatterplots with all data from 2012 and 2017 (*note duration* vs. *gap duration*, *note BW* vs. *note duration*, and *gap BW* vs. *gap duration*) and ran mixed quantile regressions with *Bird ID* as a random term (Logue et al. 2020). Quantile regression estimates a quantile (specified by the parameter *tau*) of y conditional on x. The goal of this analysis is to produce a regression line that parallels the constrained edge of the distribution, and the value of *tau* that best achieves this goal depends on the shape of the data cloud. Previous work on this population set *tau* = 0.10 (Logue et al. 2020), but we set *tau* = 0.05 to better fit the data. The resulting quantile regression lines that represent the performance limits are described in the Electronic Supplementary Materials (Table S1, Fig. S1). Quantile regression can produce imprecise estimates of performance limits (Cardoso 2019, Logue et al. 2020). Our data were not amenable to analysis with double quantile regression (Cardoso 2019; see "On Double Quantile Regression" in Electronic Supplementary Materials).

We calculated the deviation score for a given note as the orthogonal distance between the note and the performance limit, so lower deviation scores indicate higher performance (Fig. 1; Podos 2001). We averaged the deviation scores over the notes within a song because we were interested in how performance changes at the level of the whole song. For the sake of clarity, we refer to songs that closely approach the performance limit (i.e., songs with low deviation scores) as "high performance" songs. We also multiplied the Y-axes by -1 on the figures so that higher values indicate higher performance. We measured Pearson's correlation for each deviation score and the acoustic variables that contribute to it. We tested whether simple acoustic variables might be better metrics of performance than deviation scores, by examining correlations among acoustic variables and testing their skewness, which can indicate constraint (Cardoso 2017, Logue et al. 2020).

We used time relative to sunrise (*Time*) to test predictions of the hypothesis that male Adelaide's warblers warm up their voices during the dawn chorus. We calculated time relative to sunrise as the difference between the time of the song and the time of sunrise. Sunrise times were obtained from *timeanddate* (Thorsen 1995-2022). We included the cumulative number of songs that a bird had sung over the course of the morning (*Order*) in our models to account for variation in song rate (Schraft et al. 2017). We used air temperature *(Ta)* and relative humidity (*RH*) as covariates because these variables affect thermoregulation in endotherms and may affect warm-up (McKechnie and Wolf 2019, Levesque and Marshall 2021). Our weather data came from the *University of Utah's* MesoWest weather station at our field site (17.97 ° N, 67.16 ° W; The University of Utah). The weather station reports *Ta* and *RH* data once per hour to an accuracy of +/- 0.6 °C and 0-80% - +/- 2.00% at 25°C. We weight-smoothed these parameters to interpolate between hourly records. We used the National Oceanic and Atmospheric Administration (NOAA) weather station data in Lajas, Puerto Rico (18.10 ° N, 67.10° W) to validate the data from the MesoWest station (Vazquez-Cardona 2021). Finally, we used the ordinal number of the day in the year (*OD*) as a covariate.

Statistical analysis

Our statistical analysis follows the approach described by McElreath (McElreath 2020). All analyses were conducted in R x64 4.1.0. (R Core Team 2018). Descriptive data are summarized in Table S2.

Model construction occurred in two steps. First, we used networks of causal inference known as directed acyclic graphs (*DAG's*; Fig. S2) to choose predictor variables that isolate the direct effects of interest and avoid confounds (Westreich and Greenland 2013; McElreath 2020). Nodes in this network represent variables and their directed edges (arrows) represent causal influence (McElreath 2020). *DAG's* are useful because they make causal hypotheses explicit and reveal hidden confounds. The use of *DAG's* in model construction is unlike some popular model selection procedures (e.g., stepwise, information-based), because the main effects to include in the model are chosen *a priori* by the investigator, rather than by an algorithm. The goal of our analysis is to characterize the causal relationship between the dawn chorus (represented by *Time*) and vocal performance, while controlling for weather, the number of songs sung, and day of year. We applied the *adjustmentSets* function from the R package *DAGitty* (Textor et al. 2016), to confirm that our models can characterize the direct effect of *Time* on performance without confounds.

We then used information-based model selection to choose from a set of biologically plausible interaction terms. We began by considering all possible two and three-way interactions and identifying those that we deemed biologically plausible. We determined there were four biologically plausible two-way interactions (*Ta* * *RH*, *Time* * *Order*, *Ta* * *Time*, and *Ta* * *Order*) and one plausible three-way interaction (*Ta* * *Time* * *Order*). We considered these interactions plausible because (1) *Ta* and *RH* may interact to affect thermoregulation with consequences for vocal performance, (2) *Time* and *Order* may interact if song rate (songs / time) influences performance, (3) *Ta* and *Time* may interact if the effect of temperature on performance varies over the course of the morning (e.g., if high temperature influences performance more strongly early in the day), (4) *Ta* and *Order* may interact if the effect of temperature depends on how many songs the bird has sung, and (5) *Ta*, *Time*, and *Order*, may interact if the effect of temperature on performance depends on song rate (the interaction of Time and Order). For each dependent variable, we ran models with all combinations of the five biologically plausible interactions. We interpreted and visualized results for the converging models with the lowest Watanabe-Akaike information criterion (*WAIC*), and also considered any other models with WAIC ≤ 2.0 (Table S3).

We used the *brms* package to build hierarchical Bayesian models with Gaussian distributions (Bürkner 2018). All independent variables were fitted as splines to permit non-linear relationships. Examination of variance inflation factors revealed no evidence of strong multicollinearity (Kutner et al. 2005; Table S4). We standardized all predictor variables: *Ta*, *RH*, *Time*, *Order*, and *OD* prior to analysis. We also standardized the dependent variables (performance metrics and simple acoustic traits) to facilitate comparison among them. We included Y*ear* and *ST* nested within individual *(Bird ID)* as random effects to control for the non-independence of performance within year, song type, and individual. The inclusion of *ST* as a random variable means that any observed trend in performance should be interpreted as variation within, rather than among, song types. Since each song type in a male's repertoire belongs to a single song category, the random term *ST* also controls for variation attributable to song categories (Staicer 1992, Kaluthota et al. 2019). Our Markov Chain Monte Carlo (MCMC) conditioning engine used 4 chains and 3,500 iterations to estimate posterior distributions. We used priors centered on 0, which set the models' initial state to assume the independent variables have no effect on the dependent variable (McElreath 2020). We validated model fits with posterior predictive checks and R-hat values (see Fig. S3 and R-hat in Electronic Supplementary Materials; Gelman and Hill 2006, Gelman and Shirley 2011).

For all three dependent variables, we interpreted model variant #16, which included the interaction terms *Ta* * *Time*, *Ta* * *RH*, and *Time* * *Order* and the main effect *OD* (Table S3). Nine out of the 60 models (15%) did not converge. Of the models that did converge, model variant #16 had the lowest WAIC for *recovery time* and *voiced FM*. An otherwise-identical model (#17) that combined *Ta*, *Time*, and *Order* in a three-way interaction, had a slightly lower WAIC value (ΔWAIC = -0.8) for *unvoiced FM*, but the results were qualitatively identical to those of model #16. Similarly, model #12 for *voiced FM* (ΔWAIC = 1.9) had qualitatively identical results to model #16. For the sake of simplicity, we interpreted model variant #16 for all three dependent variables and the simple acoustic traits that comprise them.

We visualized the results in several ways. Plots of the conditional effects of *Time* on performance allowed us to visualize variation in performance over time in a way that accounts for the effects of covariates and random terms. To check the robustness of our findings, we used the *geom_smooth* function from the *ggplot2* package (Wickham 2016) to plot generalized additive models with cubic splines that estimate mean performance over *Time* without accounting for covariates or random terms. Certain conditional effects can be difficult to interpret in models that incorporate splines and interactions (McElreath 2020). For these effects, model predictions often paint a clearer picture of the consequences of varying the independent variables. We therefore entered simulated data into the conditioned models and visualized the output to characterize the effects of song rate on vocal performance. Specifically, we compared a simulated bird that sings at the average post-dawn rate (avg. based on full dataset = 0.6 songs / min.) throughout the whole morning to a simulated bird that sings at the average dawn chorus rate (4.8 songs / min.) before dawn, then switches to the average post-dawn rate 7 min. after dawn, when the dawn chorus typically ends (Kaluthota et al. 2019).

Results

The conditional effects of time on performance showed that male Adelaide's warblers' vocal performance was at its lowest when they first began to sing and increased during the dawn chorus (blue lines in Fig. 3). *Recovery time* and *unvoiced FM* increased rapidly during pre-dawn singing and then gradually decreased after sunrise (Fig. 3a, c). The conditional effect of *Time* on *Voiced FM* increased during pre-dawn singing, held steady after sunrise, and began to increase again at 80 minutes after sunrise (Fig. 3b). The conditional effects of *Time* on *recovery time* and *voiced FM* include strong slopes near the end of the sampling period (Fig. 3a, b). We have low confidence in the veracity of these terminal slopes because data are sparse and credible intervals are wide near the end of the sampling period, and because the terminal slopes are not present in the estimates of performance based on time alone. A comparison of the conditional deviation scores from the beginning of the dawn chorus (41 min before sunrise) to those at the end of the dawn chorus (7 min after sunrise), reveals a large change in *recovery time* (59.5% reduction in deviation score) and *unvoiced FM* (56.6%), and a moderate change in *voiced FM* (29.4%). Like the conditional effects, the curves based on performance data alone show that performance begins low, increases rapidly during the dawn chorus, and levels out after dawn (orange lines in Fig. 3). In theory, an increase in singing amplitude alone could produce spurious evidence of improving performance. We tested this hypothesis and concluded that it cannot explain the patterns in our data ("On Amplitude" in Supplementary Materials).

The simulations predicted that singing at an elevated rate would cause a 37.5% improvement in *recovery time*, a 4.0% reduction in *voiced FM*, and a 20.2% improvement in *unvoiced FM* over the course of the dawn chorus period (Fig. 4).

There was evidence that *Ta* had a small effect on performance. *Recovery time* and *unvoiced FM* peaked at intermediate *Ta's* (Fig. 5a, c). Conversely, *voiced FM* was at its lowest at intermediate *Ta's* and increased at higher temperatures, but this trend should be viewed with caution because data are sparse at high *Ta* (Fig. 5b). *Recovery time* and *unvoiced FM* increased in an approximately linear manner with *Order* (Fig. 6a, c). *Voiced FM,* however, tended to decrease slightly as birds sang more songs (Fig. 6b). The effect of *RH* and *OD* on vocal performance were weak, so we do not discuss them further (Figs. S4 and S5).

The conditional effects of *Time* on the simple acoustic traits showed that both *gap duration* and, to a lesser degree, *note duration* decreased during the dawn chorus and started to increase after sunrise (Fig. S6a-b). *Gap BW* and especially *note BW* narrowed during the dawn chorus and then start to broaden after sunrise (Fig. S6c-d). *Order* did not have strong independent relationships with any simple acoustic traits except for a negative relationship with *gap duration* (Fig. S7). The estimated effects of *Ta, RH,* and *OD* on the simple acoustic traits were weak (Figs. S8, S9, S10). Song rate simulations using the simple acoustic traits predicted that singing with an elevated rate during the dawn chorus period resulted in songs with shorter *gap duration*, and narrower *gap BW* than birds that do not participate in the dawn chorus (Fig. S11a, c). The simulation predicts only a slight difference in *note BW* and no difference in *note duration* between birds that do and do not participate in the dawn chorus (Fig. S11b, d).

Recovery time was highly correlated with *gap duration* (r = 0.94) and moderately correlated with *note duration* (r = 0.24). *Voiced FM* was highly correlated with *note duration* (r = 0.98), but very weakly correlated with *note BW* (r = 0.02). *Unvoiced FM* was highly correlated with *gap duration* (r = 0.96), and moderately correlated with *gap BW* (r = - 0.25). All measured variables were positively skewed. The skewness of deviation scores (*recovery time* = 3.44, *voiced FM* = 1.28, *unvoiced FM* = 2.48) exceeded that of the simple acoustic variables that comprised them (*note duration* = 1.11, *gap duration* = 2.44, *note BW* = 0.48, *gap BW* = 0.57; note that *gap duration* does not contribute to voiced FM).

Discussion

We modeled the effects of time of day and song rate on vocal performance to test the hypothesis that male Adelaide's warblers warm up their voices during the dawn chorus. Our key finding was that vocal performance was at its lowest when the birds begin singing, increased rapidly during the dawn chorus, then leveled off when the dawn chorus ended (Fig. 3). This finding was robust, occurring in all three performance metrics, with or without covariates and random variables. The trend appears in models with the random variable *ST*, indicating that it is driven by changes in performance within, rather than between, song types. It appears that the birds' voices "cool down" overnight when they do not sing, and warm up again during the dawn chorus, supporting a key prediction of the warm-up hypothesis.

If elaboration of the dawn chorus is an adaptation to warm up the voice, we expect that high song rates, which characterize the dawn chorus, should promote increased performance. Using linear models, Schraft et al. (2017) found that the cumulative number of songs sung positively influenced performance, but time of day negatively influenced performance. Those findings suggest that song rate (songs / time) may be a key driver of warm-up in vocal performance. Our models predicted that singing at a high rate during the dawn chorus period and then switching to a low song rate resulted in higher *recovery time* and *unvoiced FM* after dawn than did singing at a constant low rate throughout the morning (Fig. 4a, c). We also found that both *Time* and *Order* positively influence *recovery time* and *unvoiced FM* performance (Fig. 6a, c). Collectively, the song rate simulations and conditional effects of *Time* and *Order* are consistent with the idea that song rate plays an important role in vocal warm up.

It may be useful to think of performance as being influenced by "vocal performance potential" (*VPP*), which encompasses all the physiological substrates that promote vocal performance. Unlike "performance" itself, *VPP* can change even when the animal is not signaling. *VPP* may exceed observed performance if an animal chooses to signal with low performance when it could signal with high performance. We hypothesize that *VPP* can be improved (up to some point) by singing at a high rate. When birds do not sing, or sing at low rates, *VPP* tends to decrease gradually, until it reaches some lower limit. The finding that ruffed grouse (*Bonasa umbellus*) drumming performance decreases after animals take a break from displaying is consistent with the idea that display rate influences performance potential (Déaux et al. 2020). Similarly, human soccer players that re-warm up during half time exhibit higher sprint performance than players that rest (Mohr et al. 2004). Other variables, such as anatomy (Podos 2001), temperature (Fig. 5), hormone levels (Pasch et al. 2011) , and fatigue (Brumm and Slater 2006, Pasch et al. 2011), may also affect *VPP*.

Fatigue is particularly interesting if it trades off with warming up to shape the duration or intensity of the dawn chorus. Most current evidence, however, does not support this hypothesis. Schraft et al. (2017) did not find evidence that singing the same song type repeatedly induces fatigue in Adelaide's warblers (Lambrechts and Dhondt 1988). Similarly, *recovery time* and *unvoiced FM* tend to improve with the cumulative number of songs produced (Fig. 6a, c). The best existing evidence that fatigue affects performance in Adelaide's warblers is the slight decrease in *voiced FM* over song order (Fig. 6b). It would be useful to conduct a study that directly measures or manipulates song rate to examine its effects on fatigue.

We saw decreases over the course of the dawn chorus in all four of the simple acoustic variables on which our performance metrics were based (Fig. S6). During the dawn chorus, songs became faster overall (shorter *note duration* and *gap duration*) and their bandwidth narrowed (lower *note BW* and *gap BW*). While the songs got faster overall as males warmed up their voices, the duration of the silent gaps decreased faster than that of the notes, resulting in an increase in sound density (a.k.a., percent of sound; Cardoso et al. 2009, Logue et al. 2020) at the level of the whole song. Similarly, although both *note BW* and *gap BW* decreased, *note duration* and (especially) *gap duration* decreased more, resulting in an increase in *voiced* and *unvoiced FM* speed. We also saw a substantial decrease in *gap duration* with *order*, suggesting that song rate drives the reduction in *gap duration* (Fig. S7a). We conclude that *gap duration* is the only simple acoustic variable in our study that improves with warm-up. Improvement in *gap duration* compensates for decreases in the performance of other variables, driving improvements in *recovery time* and *unvoiced FM* during the dawn chorus.

One acoustic variable (*gap duration*) is the main driver of improvement in performance during the dawn chorus, raising the question of whether to focus on that variable alone or the composite variables (deviation scores) derived from trade-off plots (Cardoso 2017). The trade-offs observed in this study and in Logue et al. 2020 show that the processes that generate songs are limited. Songs that require birds to approach those limits more closely require higher performance whether those songs make greater demands along one or both acoustic axes of the trade-off graph. Thus, we can conclude that performance improves during the dawn chorus whether we focus on simple acoustic variables or deviation scores.

This study investigates the sender's side of the communication interaction, so it is reasonable to focus on the kind of variable whose production is most constrained. Consider the deviation score *recovery time* as an example. *Recovery time* is more constrained than the simple acoustic variable *gap duration* because *recovery time* accounts for the mediating effect of *note duration* on *gap duration*. Skewness analysis supports this conclusion. All deviation scores in our study were characterized by higher skewness than their constituent acoustic variables (Logue et al. 2020). We conclude that deviation scores are more constrained than the simple acoustic variables from which they are calculated, so it is appropriate to focus on deviation scores when considering constraints on signal production.

We observed slightly higher vocal performance at intermediate temperatures in two of the three performance metrics (Fig. 5a, c). If these trends are repeatable, they are evidence that some kinds of vocal performance peak at intermediate temperatures. Low performance at low temperatures is consistent with the idea that physical warming contributes to vocal warm-up. Follow-up studies that examine the relationships among *Ta*, body temperature, and vocal performance would be useful. Low performance at high temperatures is a common finding (Schulte et al. 2011; Brandt et al. 2018; Levesque and Marshall 2021). If high ambient temperature negatively affects vocal performance, increases in average temperatures caused by global warming might lower performance and select for males that sing during cooler periods (e.g., earlier in the morning).

In Adelaide's warbler, the dawn chorus is characterized not only by high song rates, but also by frequent song type switches and specific song types from the singer's repertoire ("Category B" songs; Staicer 1992, Kaluthota et al. 2019)). Frequent song type switching may warm up the voice better than monotonous singing if different song types require different motor patterns. Similarly, Category B songs may promote warm up if, for example, they require more varied motor patterns than do other songs (Staicer 1996). Previous findings that Category B songs have lower performance than Category A songs may be attributable to the fact that Category B songs tend to be sung before birds are fully warmed up (Beebee 2004, Price and Crawford 2013). Future studies should investigate the effects of song type switching and the use of Category B songs on vocal warm-up to further characterize the relationship between the structure of the dawn chorus and its function(s).

Voiced FM behaved differently than the other deviation scores in the *Order* effects, *Ta* effects, and song rate analysis. One explanation for these differences is that long *note durations* correspond to low *voiced FM* performance but high *recovery time* performance. Singers may sacrifice *voiced FM* to allow greater improvements in *recovery time*. The relative importance of different performance variables to both signalers and signal receivers is an important area for future study. A second possible explanation for the divergent behavior of *voiced FM* is that our measure of *voiced FM* is blind to patterns of within-note frequency modulation that do not affect *note BW* (Logue et al. 2020, Goller 2022). An analytic method that traces the contour of each note, like *FEX*, while measuring *voiced FM* separately from *unvoiced FM* would overcome this limitation. A third possible explanation is that our *voiced FM* metric ignores an important constraint. The songbird syrinx includes two sound sources that specialize in different frequency ranges (Suthers and Zollinger 2008). Our measure of *voiced FM* does not consider how *FM* may be constrained differently depending on whether the bird is using the right side of the syrinx, the left side, or both sides (Goller 2022).

Future studies with Adelaide's warblers should test whether performance functions in mate attraction and territory defense. The warm-up hypothesis for the function of the dawn chorus would be on firmer footing if it were clear that vocal performance influenced receiver responses in this species, as it appears to in some other songbirds (de Kort et al. 2009, Byers et al. 2010, Phillips and Derryberry 2017). A female preference for high performance songs, combined with our finding that Adelaide's warblers' vocal performance increases during the dawn chorus, could explain why the dawn chorus is restricted to the breeding season in this species (Staicer 1992). Receiver-side studies could also clarify which acoustic parameters, if any, are most salient to receivers (Phillips and Derryberry 2017).

We conclude that male Adelaide's warblers' vocal performance increases rapidly during the dawn chorus because vigorous (high rate) singing temporarily increases performance potential. Our results are the strongest support to date for the hypothesis that increasing vocal performance via warming up has contributed to the elaboration of the dawn chorus. A convincing demonstration of that hypothesis would require evidence that high vocal performance is adaptive in this taxon, and that evolutionary elaboration of the dawn chorus coincides with increased performance after dawn. An alternative interpretation of our data is that participation in the dawn chorus may warm up the voice incidentally without any adaptive benefit. We find this non-adaptive explanation less likely because Adelaide's warblers often sing near their performance limits (Logue et al. 2020) and vocal performance has been shown to have important fitness-related consequences in other species (Podos and Sung 2020). The hypothesis that warming up the voice is one function of the dawn chorus is compatible with other hypothesized functions of the dawn chorus. It seems plausible that different evolutionary paths have led to elaborate dawn choruses in different taxa (Staicer et al. 1996, Gil and Llusia 2020).

More broadly, this study highlights a possible role for "warming up" in the evolution of signaling behavior. It is well-established that selection for immediate communication functions shapes signaling behavior. While Adelaide's warblers may (and likely do) communicate during the dawn chorus, our study suggests that singing behavior during the dawn chorus has also been shaped by selection favoring rapid vocal warm-up. It may be the case that signaling behavior in other species is also influenced by selection to warm up quickly and efficiently.

Funding

This work was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (grant number RGPIN-2015-06553) to author [DML].

Acknowledgements

We thank Daniel Pereira, Josiris Rodríguez, Eva M. Rivera-Meléndez, Dawson Lehmann, Lambert Heatlie, Nolan Gooding, Shelby Prevost, Cole Pawlenchuck, Paloma Sanchez-Jaureguí, Fabio L. Tarazona, Jorge Illanas, Krystal Medina, Alfredo Lamela, Bailey Parker, Tony Shlakoff, Casandra Logue, Joshua Baez, Ruben Irizarry, and Alicia Garcia who helped to record birdsongs, process song recordings, and score the data. We also thank Liam R. Mitchell who helped with data processing and modeling. Dr. Robert Lachlan provided helpful advice and information about Luscinia. Dr. Gonçalo Cardoso provided useful information about double quantile regression. We would also like to thank Dr. Hannes A. Schraft, Dr. Andrew Iwaniuk, and members of the Behavioral and Evolutionary Ecology Research Seminar at the University of Lethbridge who provided insight into the conceptualization of this project and data analysis. Finally, we are grateful to Dr. Diego Gil and two anonymous reviewers for their helpful comments on a previous draft of this manuscript.

Analyses reported in this article can be reproduced using the data provided by Author (XXXX)

References

Amir O, Amir N, Michaeli O. 2005. Evaluating the influence of warmup on singing voice quality using acoustic measures. J Voice. 19(2): 252-260.

Beebee, MD. 2004. Variation in vocal performance in the songs of a wood-warbler: evidence for the function of distinct singing modes. Ethol. 110(7): 531-542.

Bishop, D. 2003a. Warm-up II: performance changes following active warm up on exercise performance. Sports Med. 33(1): 483-498.

Bishop, D. 2003b. Warm up I. Sports Med. 33(6): 439-454.

Brandt EE, Kelley JP, Elias DO. 2018. Temperature alters multimodal signaling and mating success in an ectotherm. Behav Ecol Sociobiol. 72(12): 191-205.

Brumm H, Slater PJ. 2006. Ambient noise, motor fatigue, and serial redundancy in chaffinch song. Behav Ecol Sociobiol. 60(4): 475-481.

Bürkner P-C. 2018. Advanced Bayesian multilevel modeling with the R package brms. The R Journal. 10(1): 395-411.

Burt J. 1995-2006. Syrinx. Version 2.6. Seattle, WA: University of Washington.

Byers J, Hebets E, Podos J. 2010. Female mate choice based upon male motor performance. Anim Behav. 79(4): 771-778.

Cardoso GC. 2013. Using frequency ratios to study vocal communication. Anim Behav. 85(6): 1529-1532.

Cardoso GC. 2017. Advancing the inference of performance in birdsong. Anim Behav. 125: e29-e32.

Cardoso GC. 2019. Double quantile regression accurately assesses distance to boundary trade-offs. Methods Ecol Evol. 10(8): 1322-1331.

Cardoso GC, Atwell JW, Ketterson ED, Price TD. 2007. Inferring performance in the songs of dark-eyed juncos (Junco hyemalis). Behav Ecol. 18(6): 1051-1057.

Cardoso GC, Atwell JW, Ketterson ED, Price TD. 2009. Song types, song performance, and the use of repertoires in dark-eyed juncos (Junco hyemalis). *Behav Ecol, 20(4): 901-907.*

Catchpole CK, Slater PJ. 2003. Bird song: biological themes and variations. Cambridge, England: Cambridge university press.

Center for Conservation Bioacoustics. 2019. Raven Pro: Interactive Sound Analysis Software. Version 1.6.1. Ithaca, NY: The Cornell Lab of Ornithology. Available from: https://ravensoundsoftware.com/software/raven-pro/.

de Kort SR, Eldermire ER, Cramer ER, Vehrencamp SL. 2009. The deterrent effect of bird song in territory defense. Behav Ecol. 20(1): 200-206.

Déaux EC, O'Neil NP, Jensen AM, Charrier I, Iwaniuk AN. 2020. Courtship display speed varies daily and with body size in the Ruffed Grouse (Bonasa umbellus). Ethol. 126(5): 528-539.

Dinh JP, Peters S, Nowicki S. 2020. Song performance improves with continued singing across the morning in a songbird. Anim Behav. 167: 127-137.

Fradkin AJ, Zazryn TR, Smoliga JM. 2010. Effects of warming-up on physical performance: a systematic review with meta-analysis. J Strength Cond Res. 24(1): 140-148.

Geberzahn N, Aubin T. 2014. Assessing vocal performance in complex birdsong: a novel approach. BMC Biol. 12(1): 1-9.

Gelman A, Hill J. 2006. Data analysis using regression and multilevel/hierarchical models. Cambridge, England: Cambridge university press.

Gelman A, Shirley K. 2011. Inference from simulations and monitoring convergence. Handbook of markov chain monte carlo. London: Chapman & Hall/CRC. p. 163-74.

Gil D, Llusia D. 2020. The bird dawn chorus revisited. In: Aubin, T, Mathevon, N, editors. Coding strategies in vertebrate acoustic communication. Cham, Switzerland: Springer. p. 45-90.

Goller F. 2022. Vocal athletics–from birdsong production mechanisms to sexy songs. Anim Behav. 184: 173-184.

Hartley RS, Suthers RA. 1989. Airflow and pressure during canary song: direct evidence for mini-breaths. J Comp Physiol A. 165(1): 15-26.

Hedley RW, Logue DM, Benedict L, Mennill DJ. 2018. Assessing the similarity of song-type transitions among birds: evidence for interspecies variation. Anim Behav. 140: 161-170.

Hoese WJ, Podos J, Boetticher NC, Nowicki S. 2000. Vocal tract function in birdsong production: experimental manipulation of beak movements. J Exp Biol. 203(12): 1845-1855.

Kaluthota CD, Medina OJ, Logue DM. 2019. Quantifying song categories in Adelaide's Warbler (Setophaga adelaidae). J Ornithol. 160(2): 305-315.

Kutner MH, Nachtsheim CJ, Neter J, Li W. 2005. Applied linear statistical models. New York, USA: McGraw-Hill New York.

Lachlan R. 2007. Luscinia: a bioacoustics analysis computer program. Version 2.14. Available from: https://rflachlan.github.io/Luscinia/.

Lambrechts M, Dhondt A. 1988. The anti-exhaustion hypothesis: A new hypothesis to explain song performance and song switching in the great tit. Anim Behav. 36(2): 327-334.

Levesque DL, Marshall KE. 2021. Do endotherms have thermal performance curves? J Exp Biol. 224(3): jeb141309.

Logue DM, Sheppard JA, Walton B, Brinkman BE, Medina OJ. 2020. An analysis of avian vocal performance at the note and song levels. Bioacoustics. 29(6): 709-730.

McElreath R. 2020. Statistical rethinking: a Bayesian course with examples in R and Stan. New York, USA: Chapman and Hall/CRC.

McKechnie AE, Wolf BO. 2019. The physiology of heat tolerance in small endotherms. Physiol. 34(5): 302-313.

Mohr M, Krustrup P, Nybo L, Nielsen JJ, Bangsbo J. 2004. Muscle temperature and sprint performance during soccer matches–beneficial effect of re-warm-up at half-time. Scand J Med Sci Sports. 14(3): 156-162.

Nowicki S, Westneat M, Hoese W. 1992. Birdsong: motor function and the evolution of communication. Semin Neurosc. 4: 385-90.

Pasch B, George AS, Campbell P, Phelps SM. 2011. Androgen-dependent male vocal performance influences female preference in Neotropical singing mice. Anim Behav. 82(2): 177-183.

Phillips JN, Derryberry EP. 2017. Vocal performance is a salient signal for male–male competition in White-crowned Sparrows. Auk. 134(3): 564-574.

Plummer EM, Goller F. 2008. Singing with reduced air sac volume causes uniform decrease in airflow and sound amplitude in the zebra finch. J Exp Biol. 211(1): 66-78.

Podos J. 1997. A performance constraint on the evolution of trilled vocalizations in a songbird family (Passeriformes: Emberizidae). Evol. 51(2): 537-551.

Podos J. 2001. Correlated evolution of morphology and vocal signal structure in Darwin's finches. Nature. 409(6817): 185-188.

Podos J, Moseley DL, Goodwin SE, McClure J, Taft BN, Strauss AVH, Rega-Brodsky C, Lahti DC. 2016. A fine-scale, broadly applicable index of vocal performance: frequency excursion. Anim Behav. 116: 203–212.

Podos J, Nowicki S. 2004. Beaks, adaptation, and vocal evolution in Darwin's finches. Bioscience. 54(6): 501-510.

Podos J, Sung H-C. 2020. Vocal performance in songbirds: from mechanisms to evolution. The neuroethology of birdsong. Cham, Switzerland: Springer. p. 245-268.

Price JJ, Crawford CL. 2013. Use and characteristics of two singing modes in pine warblers. Wilson J Ornithol. 125(3): 552-561.

R Core Team. 2018. R: A language and environment for statistical computing. Version 4.1.0. Vienna, Austria: R Foundation for Statistical Computing. Available from: https://www.R-427 project.org/.

Schraft HA, Medina OJ, McClure J, Pereira DA, Logue DM. 2017. Within-day improvement in a behavioural display: wild birds ‘warm up’. Anim Behav. 124: 167-174.

Schulte PM, Healy TM, Fangue NA. 2011. Thermal performance curves, phenotypic plasticity, and the time scales of temperature exposure. Integr Comp Biol. 51(5): 691-702.

Staicer CA. 1996. Acoustical features of song categories of the Adelaide's Warbler (Dendroica adelaidae). Auk. 113(4): 771-783.

Staicer CA, Spector D, Horn A. 1996. The dawn chorus and other diel patterns in acoustic signaling. In: Kroodsma D, Miller E, editors. Ecology and evolution of acoustic communication in birds**.** Ithaca, NY: Cornell University Press. p. 426-453

Staicer CA. The role of male song in the socioecology of the tropical resident Adelaide's warbler (Dendroica adelaidae). [Dissertation]. Amherst, MA: University of Massachusetts; 1992

Suthers RA, Zollinger SA. 2004. Producing song: the vocal apparatus. Ann NY Acad Sci. 1016(1): 109-129.

Suthers RA, Zollinger SA. 2008. From brain to song: the vocal organ and vocal tract. In: Zeigler HP, Marler P, editors. Neuroscience of birdsong. Cambridge, England: Cambridge University Press. p. 78-98.

Tan CL, Knight ZA. 2018. Regulation of body temperature by the nervous system. Neuron. 98(1): 31–48.

Textor J, Van der Zander B, Gilthorpe MK, Liskiewicz M, Ellison GTH. 2016. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. Int J Epidemiol. 45(6): 1887-1894.

The University of Utah. [Internet]. Meso West, Weather Conditions for CRRP4. Utah, USA: University of Utah; [updated 2022 June 5; cited 2022 June 5]. Available from: https://mesowest.utah.edu/cgi-bin/droman/meso_base_dyn.cgi?stn=CRRP4.

Thorsen S. [Internet]. 1995-2022. Timeanddate. Stavanger, Norway: timeanddate.com; [updated 2022 June 5; cited 2022 June 5]. Available from: https://www.timeanddate.com/sun/@4567986.

Toms JD. [Internet]. 2020. Adelaide's warbler (Setophaga adelaidae), version 1.0. Ithaca, NY (USA); birdsoftheworld.org; [updated 2020 Mar 4; cited 2022 June 5]. Available from: https://birdsoftheworld.org/bow/species/adewar1/1.0/introduction.

Vazquez-Cardona J. Is the dawn chorus an adaptation to warm-up the voice in Adelaide's warbler (Setophaga adelaidae)? [M.Sc. Master's thesis]. Lethbridge, AB: University of Lethbridge; 2021

Westreich D, Greenland S. 2013. The table 2 fallacy: Presenting and interpreting confounder and modifier coefficients. Am J Epidemiol. 177(4): 292-298.

Wickham H. 2016. *ggplot2*: Elegant graphics for data analysis. Springer-Verlag New York.

Wilson DR, Bitton PP, Podos J, Mennill DJ. 2014. Uneven sampling and the analysis of vocal performance constraints. Am Nat. 183(2): 214–228.

Figure legends

Figure 1. Scatterplot of a hypothetical constrained relationship between two traits. The black line is the 10th quantile regression line used to estimate a performance limit. The deviation score is the orthogonal distance between each data point and the performance limit. Figure reproduced from Logue et al. (Logue et al. 2020) with the authors' permission.

Figure 2. Sound spectrograms of four song types from male Adelaide's warblers (Hann window, 512 points / sample).

Figure 3. Plots showing the effect of time relative to sunrise on the acoustic performance variables (a) *recovery time*, (b) *voiced frequency modulation*, and (c) *unvoiced frequency modulation* in male Adelaide's warbler song. The signs of the performance variables are standardized and multiplied by -1, so high values indicate high performance. Each graph shows the songs (semi-transparent black dots), time of sunrise (vertical gray line), estimated mean performance based on vocal performance alone (orange line), and conditional effect of time on performance after accounting for covariates and random terms (curved blue line; shading = 95% CI). The dawn chorus is visible as a dense cloud of points that ends shortly after dawn.

Figure 4. Plots of simulated data comparing the expected distribution of (a) *recovery time*, (b) *voiced frequency modulation*, and (c) *unvoiced frequency modulation* over the course of the morning when birds do (blue) or do not (orange) sing at an elevated rate during the dawn chorus period. We standardized the performance variables and multiplied them by -1, so high values indicate high performance. Graphs show the time of sunrise (vertical gray line).

Figure 5. Conditional effects of air temperature on the acoustic performance variables (a) *recovery time*, (b) *voiced frequency modulation*, and (c) *unvoiced frequency modulation* in male Adelaide's warbler song. We multiplied the performance variables by -1 so that high values would indicate

high performance. Each graph shows the songs (semi-transparent black dots), the conditional effects of air temperature on performance from a generalized linear mixed model (curved blue line), and its 95% credible interval (turquoise).

Figure 6. Effects of the cumulative number of songs sung over the course of the morning on the acoustic performance variables (a) *recovery time*, (b) *voiced frequency modulation*, and (c) *unvoiced frequency modulation* in male Adelaide's warbler song. See Fig. 5 caption for details.