Behavioral Ecology
doi:10.1093/beheco/arm078
Advance Access publication 23 August 2007

Inferring performance in the songs of dark-eyed juncos (*Junco hyemalis*)

Gonçalo C. Cardoso,[a] Jonathan W. Atwell,[b] Ellen D. Ketterson,[b] and Trevor D. Price[a]
[a]Department of Ecology and Evolution, University of Chicago, 1101 East 57th Street, Chicago, IL 60637, USA and [b]Department of Biology, Indiana University, 1001 East 3rd Street, Bloomington, IN 47405, USA

Within bird species, songs differ in their attractiveness to females or effectiveness in male–male interactions. Some songs are more difficult to sing than others, and receivers may use a singer's performance of difficult songs as a means for evaluating the quality of the singer. The concept of song performance aims at quantifying how physiologically demanding are different songs. Using variation between song types of dark-eyed juncos, *Junco hyemalis*, we show that some song traits trade off with costly aspects of song output—short intervals between syllables or loud sound amplitude—suggesting that those traits are difficult to sing. First, after controlling for other traits, long syllables require longer intervals for recovery. This supports the idea that a measure of "respiratory performance" could be based on the relative lengths of syllables and intervals. Second, some syllable traits trade off strongly with sound amplitude, suggesting that these traits may be difficult to sing at high amplitudes. The ratio of frequency bandwidth and trill rate has been used to infer performance in other bird species, but we found no evidence that frequency bandwidth trades off with any aspect of song output in the junco. The negative association of bandwidth with trill rate may instead be a passive consequence of syllable length, with longer syllables randomly accumulating frequency modulation. We conclude that bird receivers may best evaluate how well a song is performed if they integrate multiple cues and discuss how researchers may similarly devise measures of song performance. *Key words:* bird song, dark-eyed junco, *Junco hyemalis*, song performance. *[Behav Ecol 18:1051–1057 (2007)]*

Bird songs are usually complex acoustic signals that function in attracting and stimulating females and in mediating competition between males (Catchpole and Slater 1995). In both contexts, receivers should try to evaluate sender quality or motivation, and song may encode this information through different types of cost (reviewed in Vehrencamp 2000; Gil and Gahr 2002; Searcy and Nowicki 2005, Chapter 3). For example, time spent singing may be limited by energetic needs (e.g., Thomas 1999a, 1999b) or the size of repertoires may reveal neuronal developmental costs (Nowicki et al. 2002). Songs themselves may differ in how difficult they are to sing. We use the term song performance to refer to how physiologically demanding different songs are to sing (e.g., Forstmeier et al. 2002; Ballentine et al. 2004). High-performance songs, that is, those that are more difficult to sing, may be salient to receivers if they give cues to singer quality (Podos et al. 2004; Searcy and Nowicki 2005, Chapter 3). For example, females of some species prefer certain song syllables or certain renditions of the same song type that are thought to be more physiologically costly (Vallet and Kreutzer 1995; Searcy 1996; Vallet et al. 1998; Ballentine et al. 2004).

The most influential work on how to evaluate song performance is Podos' (1997) demonstration that, across the Emberizidae, there is a trade-off between trill rate and frequency bandwidth, which suggests that songs with simultaneously fast trill rates and wide bandwidths are costly to sing. In this paper, we apply this trade-off reasoning to investigate performance costs for a variety of song characteristics. Songs differ in traits other than trill rate and bandwidth, for example, the presence of 2 voices, frequency inflections, multiple elements, or harmonics, and these may determine the cost of singing. For example, it is possible that some songs with slow trill rates and narrow bandwidth are elaborate in some other costly trait (e.g., contain more elements, more frequency inflections, longer harmonics) and are therefore still high-performance songs. This is tacitly recognized when some researchers restrict the use of the rate × bandwidth interaction to within–song type comparisons (Ballentine et al. 2004; Illes et al. 2006), this way controlling for variation in the many other traits that differ between song types. Ideally, one would want to evaluate the costs of many syllable traits in order to compute a measure of performance that integrates all of them.

In this paper, we use natural variation among song types of dark-eyed juncos, *Junco hyemalis*, an Emberizid species with simple trilled songs (Konishi 1964), to test how several syllable traits are related to aspects of song output that are a priori likely to indicate performance costs. We measure 2 aspects of song output that are physiologically costly—short intersyllable intervals and high sound amplitude—and ask if other song traits trade off with them, that is, if song traits are associated with longer intervals between syllables or lower sound amplitude.

First, in an extension of the reasoning used by Podos (1997), we consider trill rate. Rather than focus on trill rate as a single trait, we partition it into the length of syllables and the length of intervals between syllables because relations of other traits to each of these 2 components lead to different interpretations. In several species, intervals between syllables are required for taking minibreaths where the birds recover the volume of air expired during phonation (Suthers et al. 1999). Longer syllables tend to require the inspiration of a larger volume of air afterward (Hartley and Suthers 1989) and hence are followed by longer intervals (Mota and Cardoso 2001). Therefore, if a song trait is positively related to the length of intervals between syllables, this may indicate a respiratory cost for that trait. Based on this rationale, Forstmeier et al. (2002) successfully measured performance in the dusky

---

Address correspondence to G.C. Cardoso. E-mail: cardoso@uchicago.edu.

Received 23 February 2007; revised 18 July 2007; accepted 20 July 2007.

Table 1
Definitions of syllable traits

Trait	Definition
Frequency bandwidth	Highest minus lowest sound frequency of the syllable, as visible in the spectrogram (kHz)
Peak frequency	Sound frequency with the highest cumulative amplitude across the syllable (kHz)
Number of frequency inflections	Number of times that a rising frequency modulation is followed by a descending one or vice versa
Number of elements	Number of discontinuous and temporally nonoverlapping traces on the spectrogram
Length of harmonics	Duration of octaves or other harmonics visible in the spectrogram (seconds)
Length of 2 voices	Duration of overlapping independently modulated sounds (seconds)
Length of "rattles"	Duration of harsh sounds with very rapidly repeated (>50 Hz) subelements (seconds)
Length of gaps	Duration of within-syllable pauses between elements (seconds)
Length of syllables	Duration of the syllable, as visible in the spectrogram (seconds)
Length of intervals	Duration of the intervals between syllables, as visible in the spectrogram (seconds)
Amplitude	Maximum amplitude of sound in the syllable (volts), after filtering out low-frequency background noise (<2 kHz)

Note that syllable amplitude is a relative measure and is not comparable between recordings.

warbler (*Phylloscopus fuscatus*). Conversely, positive associations of song traits with syllable length may not be good indications of performance costs because they may be a passive consequence of syllable length, whereby syllable traits (frequency inflections, harmonics, etc.) accumulate more or less randomly through time. Granting this, we nevertheless analyze syllable length in order to understand why different song traits may relate to trill rate.

Second, as another measure of a likely performance cost, we ask if song traits trade off with sound amplitude. Louder bird sounds require stronger contraction of respiratory muscles and the ability to cope with higher air pressures, which should generally be more challenging for the vocal organs (Suthers et al. 1999). Therefore, if song traits are negatively related to sound amplitude, this could indicate that those traits are difficult to sing under those demanding conditions. Amplitude is an important aspect of song output that is independent from the temporal dimension, and, therefore, it may show costs that are not detected analyzing interval lengths. We conclude with a discussion of how inferred costs may be used to evaluate song performance.

METHODS

Recordings and measurements

We recorded free-living male dark-eyed juncos from February 2006 to July 2006 in 2 southern California populations: at the campus of the University of California at San Diego (32°52'N, 117°14'W) and in the nearby Laguna mountains (32°51'N, 116°26'W). Details of the study sites are given in Yeh and Price (2004). Most males were individually marked with a combination of colored leg bands, and the identity of unbanded males (13 out of 97 recorded males) could be confidently assigned based on location and song type. On average, each male was recorded 6.0 times (± 4.83 standard deviation [SD]), mostly on different days. Most recordings are of spontaneous song (468 out of 582 recordings), and others were obtained using playbacks. We used an Audio-Technica AT835b directional microphone and a M-Audio Microtrack 24/96 solid-state recorder. Recordings were downsampled to 22050-Hz files and songs measured on spectrograms with bandwidth of 56 Hz, time resolution of 1.45 ms, and lower cutoff filter at 2 kHz, using the software Avisoft SASLAB v. 4.34 (Avisoft Bioacoustics, Berlin, Germany).

Dark-eyed junco songs typically consist of a trill with a single syllable; a syllable being defined as the unit that repeats within the trill (Konishi 1964). Males in the studied populations sing 2–8 song types (Newman MM, Yeh PJ, Price TD, unpublished data). We assigned syllable types by visual inspection of spectrograms, based on the shape of elements within syllables (Newman MM, Yeh PJ, Price TD, unpublished data). Seventeen out of the 193 syllable types that we recorded were sometimes sung in multisyllable songs—that is, songs containing 2 or more different trilled syllables (Konishi 1964). In these cases, we treated each syllable type separately. From each recording, we chose a sequence of 5 songs with the best sound quality (i.e., the least noise and reverberation) to measure. In recordings with more than one song type, we chose 5 songs of each type adjacent in the recording. In short or very noisy recordings, the number of songs measured was limited by availability. On average, we measured 4.45 songs per recording. We excluded 10% of the syllables (range from 1 to 6 syllables) at the beginning and at the end of each song and measured all those in between because songs often increase in amplitude at the beginning and decrease at the end, sometimes with some elements apparently missing in the very first and last syllables (personal observation).

Song traits are defined in Table 1 and some are illustrated in Figure 1. The following traits were measured for every syllable: length, length of the next interval, frequency bandwidth, peak frequency, and amplitude. Amplitude cannot be compared between recordings because they differ in distance from the bird to the microphone and other conditions that affect its measurement. Therefore, we averaged amplitude measurements for all syllables of the same type within each recording, and later we restricted comparisons of syllable amplitude only within recordings (see below). For the other traits, measurements were first averaged for all syllables in each recording, then for all recordings of the same syllable type of each male, and then, in the case of syllable types sung by multiple males, for all males sharing that syllable type. Trill rate can be readily computed from the values of syllable and interval lengths, as the inverse of the sum of syllable and interval lengths, which is equivalent to the number of syllables per second. In addition, for each syllable type and male we chose a good quality song from the best quality recording available and measured the length of harmonics, of 2 voices, of rattled sounds, and of gaps between elements. Again, measurements were averaged for all syllables of the same type and, in the case of shared types, averaged across birds. Finally, for each syllable type we noted the number of frequency inflections and the number of elements. Nearly adjacent elements may appear a few milliseconds overlapped in the spectrograms and were counted as separate elements if not coincident in frequency.

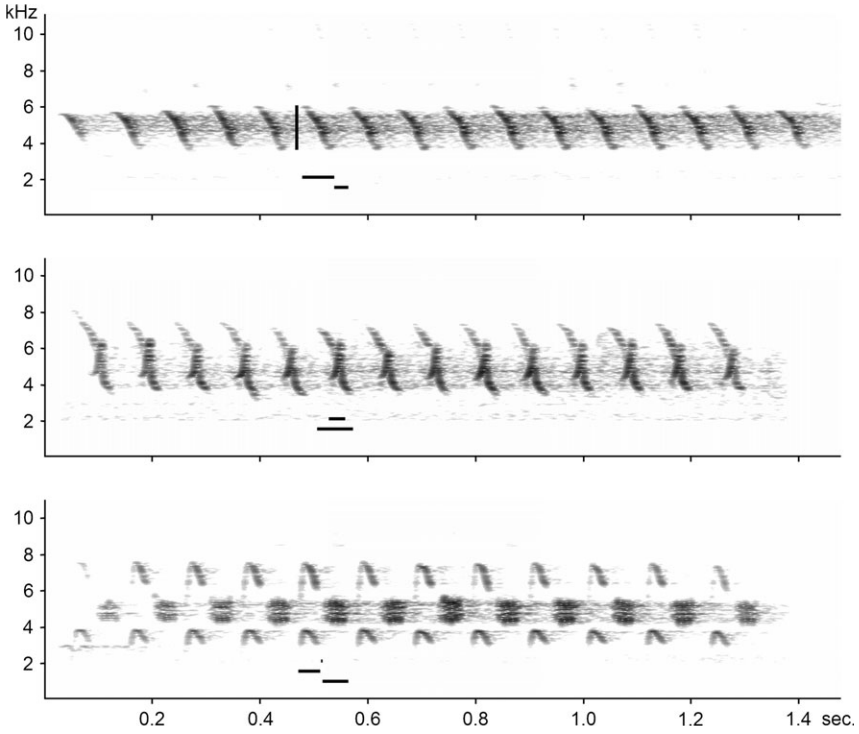

Figure 1
Spectrograms of 3 junco songs, illustrating some of the syllable measurements. Bars in the spectrograms illustrate, from top to bottom, frequency bandwidth, length of syllables, and length of intervals in the first spectrogram, length of 2 voices and length of syllables in the second, and length of gaps, length of harmonics, and length of "rattles" in the last.

Analyses

The sample points in these analyses are the different syllable types. We excluded 5 of the 193 syllable types because the lengths of the intervals between syllables were outliers (Figure 2). Three of those were sung much slower than all other syllable types, the fourth was atypical for junco song, almost a continuous rattle, and the fifth was often sung either in very short songs or songs with broken syntax (as defined in Podos 1996). The remaining 188 syllable types were sung by an average of 1.63 males each (maximum 7). On average, each syllable type was recorded on 3.14 (± 2.64 SD) separate occasions.

The main analyses consisted of 3 multiple linear regressions relating song traits with the length of intervals between syllables, length of syllables, or sound amplitude. In all 3 cases, the residuals from the regression approximated a normal distribution, were homoscedastic, and did not have structure (i.e., did not increase, decrease, or vary nonlinearly across the range of predicted values, Supplementary Figure 1). We also screened all partial plots (i.e., scatter plots of residuals of the dependent variable on the residuals of each independent variable, when both were regressed separately on the rest of the independent variables) and found no nonlinear relationships. This indicates that the linear models we used were appropriate for our dataset. Finally, multicollinearity was not a problem in these regressions because, in all cases, independent variables had substantial tolerance (i.e., amount of variance not explained by the other independent variables in the model; average tolerance was 0.72, minimum was 0.40). After running each of the 3 regressions, we proceeded to remove nonsignificant syllable traits in a stepwise manner. Multiple regressions were run with the software SPSS 13.0.

The first linear regression is of interval length on syllable traits. We did not use amplitude in this regression because it is not comparable between recordings. Multiple regression assesses the effect of each trait while controlling for every other trait. In particular, we expected syllable length to have an effect on the length of intervals, so that controlling for it should be necessary to investigate possible respiratory costs of the other traits. We also ran a multiple regression of syllable length on the other syllable traits because syllable length is a component of trill rate.

In 34 recordings, males sang more than one syllable type, and amplitudes between these types can be compared. From these recordings, we had a total of 30 different pairs of syllable types from 20 birds. Eight of those syllable type pairs were from multisyllable songs, 21 pairs were from recordings where the male changed or alternated between song types, and one

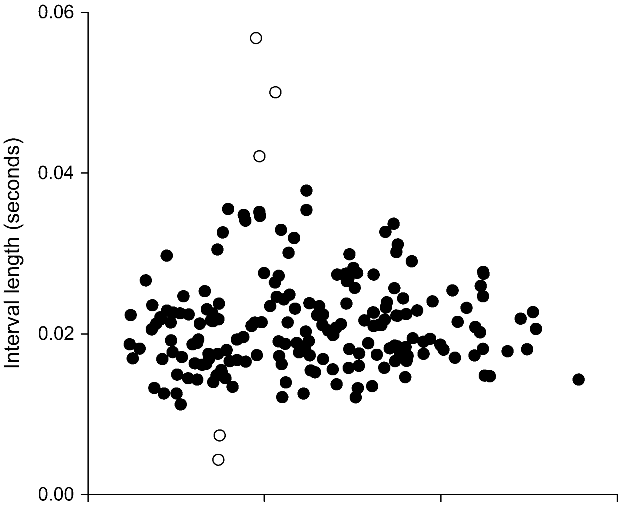

Figure 2
Relation between the length of syllables and the length of intervals between syllables. Outliers of interval length are shown in open circles.

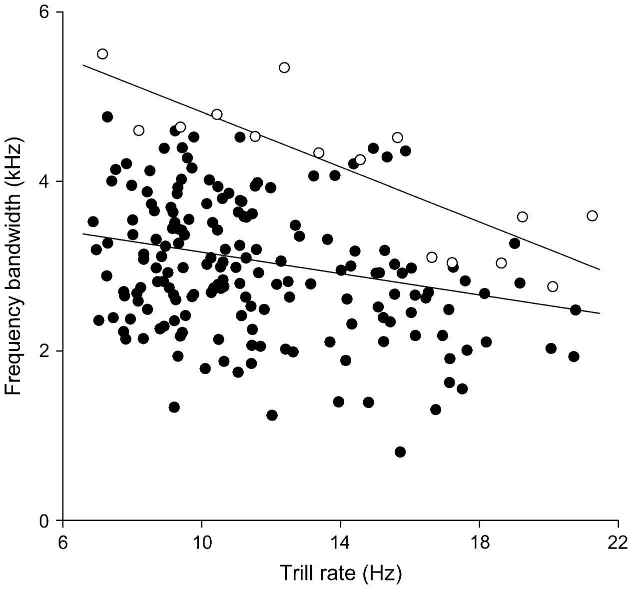

Figure 3
Relation between trill rate and frequency bandwidth. Open circles are the syllable types used for the upper bound relation and were chosen as the syllable types with the largest bandwidth in each 1-Hz rate class (e.g., Ballentine et al. 2004), starting at the lowest trill rate. Regression lines for all data and for the upper bound points are shown.

pair was found in both types of recordings. We computed the differences of each trait for each syllable type pair and averaged the values for different recordings of the same syllable type pair. We then regressed amplitude differences on differences of syllable traits. Because here we were working with differences, we forced the multiple regression through the origin.

RESULTS

Syllables were on average more than 3 times longer than intervals (averages were 0.07 and 0.02 s, respectively, $N = 188$ syllable types). Both syllable and interval lengths varied substantially among song types (coefficients of variation are 0.33 and 0.26, respectively; Figure 2). As in other songbird species (Podos 1997), frequency bandwidth was negatively correlated with trill rate ($r = -0.25$, $P < 0.01$, $N = 188$), and there was a significant upper bound limit to this relation ($r = -0.84$, $P < 0.01$, $N = 15$; Figure 3).

Effect of traits on interval length

Contrary to simple expectations based on respiratory demands, syllable length was not significantly positively correlated with interval length ($r = 0.05$, $P = 0.47$, $N = 188$). However, once other traits are controlled for in a multiple regression, syllable length was significantly positively associated with interval length (Table 2). This suggests that other traits interact with syllable length in determining respiratory demands. Stepwise removal of nonsignificant traits from the full model shows that the greatest decrease in the effect of syllable length is when number of elements is removed (Supplementary Table 1), indicating that the relation between syllable length and intervals stands among syllables with the same number of elements but not among syllables with different number of elements.

Contrary to syllable length, most other syllable traits had negative partial regression coefficients, and that of the length of rattles was significant (Table 2). Therefore, after controlling for syllable length, more elaborate syllable traits do not seem to imply greater respiratory recovery times. The full multiple regression model explained only 9% of the variation in interval length ($R^2 = 0.09$, $F_{9,178} = 1.96$, $P = 0.047$), which implies that for any combination of syllable traits, song types vary substantially in the length of intervals.

The negative relationships between elaboration of syllable traits and interval length could arise if high quality birds sing elaborated syllable traits with short intervals, whereas low-quality birds sing less elaborated traits with longer intervals. We checked this possibility using averages per song type within each male (not across males) and adding male identity as a random factor to the regression model in order to control for differences between the respiratory abilities of different males. The trend remained identical: the direction of slopes did not change in any case, and syllable length was still the only trait with a significant positive effect. Therefore, the associations

Table 2
Standardized partial regression coefficients and significance values from the multiple regressions of aspects of song output on syllable traits

	Regression of interval length[a]		Regression of syllable length[b]		Regression of amplitude differences[c]	
Traits	β	*P*	β	*P*	β	*P*
Frequency bandwidth	−0.130	0.136	**0.193**	**0.002**	0.215	0.277
Peak frequency	0.006	0.931	−0.041	0.460	0.045	0.811
Number of frequency inflections	−0.072	0.371	**0.167**	**0.005**	−0.334	0.082
Number of elements	−0.158	0.101	**0.480**	**<0.001**	**−0.711**	**0.005**
Length of harmonics	0.062	0.449	−0.001	0.989	−0.123	0.562
Length of 2 voices	−0.085	0.310	−0.042	0.502	**0.384**	**0.034**
Length of “rattles”	**−0.249**	**0.002**	**0.146**	**0.014**	**−0.530**	**0.013**
Length of gaps	−0.008	0.925	**0.146**	**0.014**	0.145	0.456
Length of syllables	**0.264**	**0.009**	—	—	0.291	0.259

Significant effects are marked in bold.

[a] Sample size is 188 syllable types; degrees of freedom (df) for the whole model are 9, 178.

[b] Sample size is 188 syllable types; df = 8, 179.

[c] Sample size is 30 pairs of syllable types; df = 9, 21.

of syllable traits with interval length are not likely to be due to differences in respiratory ability among males but more likely reflect the respiratory demands of the different song traits.

Effect of traits on syllable length

In the regression of syllable length on syllable traits ($R^2 = 0.49$, $F_{8,179} = 21.82$, $P < 0.01$), there were several significant positive effects (Table 2) and no significant negative effects; that is, longer syllables generally had more elaborated traits. These results remain unchanged after stepwise removal of nonsignificant traits from the model (Supplementary Table 2). The associations of these syllable traits with trill rate (not shown) were similar to those with syllable length (with inverted sign) because syllables are much longer than intervals and, thus, trill rate is determined mostly by syllable length.

Effect of traits on sound amplitude

The regression of amplitude differences on syllable trait differences explained about half the variation in amplitude ($R^2 = 0.48$, $F_{9,21} = 2.13$, $P = 0.07$), and despite the relatively small sample size, 3 effects were significant: the number of elements and the length of rattles are negatively associated with amplitude and the length of 2 voices is positively associated (Table 2). These effects become progressively weaker during stepwise removal of variables from the model and 2 of them (length of 2 voices and length of rattles) became nonsignificant (Supplementary Appendix Table 3). Thus, as was the case in the analysis of interval length, to detect effects on amplitude it seems important to control for other syllable traits.

To check that the trade-offs of number of elements and the length of rattles with amplitude are not indirect consequences of the respiratory demands of those traits, we added interval length as an independent variable to the full multiple regression. The results remained identical: interval length did not influence amplitude significantly (standardized partial regression coefficient, $\beta = 0.21$), the standardized regression coefficients of the other traits remained identical (maximum change of β is 0.12), and their significance never changed qualitatively. Therefore, song trait trade-offs with sound amplitude are largely independent from their respiratory demands.

DISCUSSION

We asked if particular song traits are associated with either longer intervals between syllables or lower absolute amplitude. Either of these associations would suggest that such traits are costly to sing. We also tested whether traits are related to syllable length in order to understand how song traits influence trill rate because trill rate is a composite of syllable length and interval length and is often used to infer song performance. We first interpret results for each of these tests and then discuss the implications for inferring song performance.

Long syllables should require the inspiration of a larger volume of air afterward for recovery (Hartley and Suthers 1989) and therefore be followed by longer intervals, but, in dark-eyed junco songs, syllable length is associated with longer intervals only when the other song traits are included in a multiple regression. This may be because syllable traits influence the volume of air that is expired during phonation, so that the effect of syllable length is only apparent after controlling for them. For example, rattles are negatively associated with interval length. Rattles are atonal and rapidly patterned sounds produced by a vibratory mechanism (Brackenbury 1978) that differs from the whistled mechanism of most elements (Suthers et al. 1999). It is likely that this vibratory mechanism obstructs airflow more than regular whistles, thus influencing the respiratory demands. The same may apply to the number of elements within syllables because different elements are generally produced by momentarily occluding airflow at the syrinx (Suthers et al. 1999), and this would explain why number of elements must be controlled for in order for the effect of syllable length to be apparent. Apart from syllable length, no other traits are positively associated with interval length, and, therefore, those other traits do not seem to have respiratory costs.

Number of elements strongly trade off with amplitude, and rattles also have a negative effect. It has been suggested that multielement syllables are difficult to sing because of the complex motor coordination that they require (Vallet et al. 1998; Suthers et al. 2004). Our results support this and for the first time identify a trade-off for singing multielement syllables, that is, reduced amplitude. The effect of rattles on amplitude is also easy to interpret as it is likely that their vibratory mechanism (Brackenbury 1978) dissipates more energy mechanically and therefore is less efficient than that of regular whistles in producing loud sounds. One trait—the length of 2 voices—has a positive effect on amplitude, which is probably due to the simultaneous use of both sides of the syrinx summing to produce a louder sound.

Finally, several syllable traits (number of elements, frequency bandwidth, and others) are positively and significantly correlated with syllable length and, as a consequence, negatively related to trill rate. It is not straightforward to interpret these as costs. For example, Podos (1997) suggested that motor constraints might limit the rate at which birds sing syllables with large bandwidths. This argument is based on plots of rate × bandwidth in bird song, which are often triangular with a negative upper bound (Podos 1997, 2001; Drăgănoiu et al. 2002; Ballentine et al. 2004; Beebee 2004; Illes et al. 2006, Figure 3). However, a simpler explanation for these triangular plots is that syllables accumulate frequency modulation through time more or less randomly, in which case short syllables (which correspond to faster trill rates) have narrower bandwidths and also narrower variance than long syllables (which correspond to slow trill rates). Similarly, the other correlations of syllable traits with syllable length may simply be due to a random walk of those traits over time.

How to infer song performance?

Receivers are expected to pay attention to song traits that are costly to sing. We a priori identified 2 aspects of song output that are likely to indicate costs: amplitude and interval length. If receivers were to use only these aspects to evaluate performance, then males should sing only loud, short, and simple syllables, which is not the case. In fact, syllables vary in many other traits, and this may provide additional information to receivers. Some syllable traits trade off with these costly aspects of song output, suggesting that they also have performance costs, so that the interaction between syllable traits and output should be used to evaluate song performance more accurately. Songs lying along a trade-off line between syllable traits and a costly output can be seen as of equal cost or of equal performance as elaboration of the song traits is offset by decrease in output. Given this situation the best way to infer performance may be to look at where songs lie in relation to these trade-off lines.

In assessing song performance, bird receivers may integrate many song traits but researchers have so far focused on only a few. First, after Podos' (1997) discovery of a negative association between trill rate and bandwidth, several papers used the distance between individual songs and the species rate × bandwidth upper limit as a measure of performance, reasoning that the closer a song is to this limit the more costly it is to sing. Second, Forstmeier et al. (2002) measured performance as the proportion of time during which sound amplitude exceeds a certain threshold, which mostly captures the relative length of syllables and intervals in song. This was interpreted as an index of respiratory performance because singing long syllables with small pauses for inspiration should hamper respiratory gas exchange (Hartley and Suthers 1989; Podos 1996). In the light of our results, we consider whether these measures are likely to provide good information to receivers about song performance.

After controlling for other traits, frequency bandwidth does not trade off either with interval length or amplitude, suggesting that, in dark-eyed juncos, it is not performance limited. There is a negative relationship of bandwidth with trill rate, but this arises because both are correlated with syllable length, which, as explained above, may not reflect a cost. Various researchers have used distance to the rate × bandwidth limit as a measure of song performance, but most results are difficult to interpret because they could be driven by the song traits individually (rate and/or bandwidth) instead of the interaction between them, which is how performance is inferred. For example, Huber and Podos (2006) and Ballentine (2006) showed that songs of larger billed birds are more distant from the rate × bandwidth limit, but in both cases the result is entirely due to differences in bandwidth. Other studies (Podos 2001; Ballentine et al. 2004; Beebee 2004; Christensen et al. 2006; Cramer and Price 2007) do not report on whether results are better predicted by the distance to the rate × bandwidth limit than they would be by rate or bandwidth alone. The study of Drăgănoiu et al. (2002) is the only study to test the behavioral effects of rate and bandwidth separately. Rate was manipulated by altering the length of intervals between syllables, and different syllable types provided 3 different bandwidths. These syllable types probably differ systematically in length: the syllable type with the largest bandwidth shown in Drăgănoiu et al. (2002) is about twice as long as the one with narrower bandwidth so that, at a same rate, differences in interval length are even larger. Thus, the female preferences that were found both in the rate and frequency experiments may be due to simple differences in the length of intervals. Illes et al. (2006) and Cardoso et al. (2007) manipulated the length of intervals between syllables alone and found that this is enough for male and female birds to respond differently.

The ratio of syllable length to the total length of syllables plus intervals is a simple measure of respiratory performance, and some playback experiments showed that at least for within–song type comparisons, it is biologically relevant (Drăgănoiu et al. 2002; Illes et al. 2006; Cardoso et al. 2007). In the multiple regression analysis of junco song, syllable length is the only trait significantly associated with longer intervals. However, it appears that other traits influence interval length, so that a receiver that pays attention to them should obtain an improved assessment of respiratory performance. We can easily compute a measure of respiratory performance that takes this into account as the residual from the multiple regression equation of interval length on syllable traits (Table 2). This regression explains a small proportion of the variation in interval length, suggesting that song types differ substantially in performance. This might also be because interval length varies due to factors other than respiratory demands, such as motivation (Cardoso GC, Atwell JW, Ketterson ED, Price TD, unpublished data). However, interval length is highly consistent within song types (Cardoso GC, Atwell JW, Ketterson ED, Price TD, unpublished data), and, as shown here, syllable traits influence interval lengths predictably, so that to a great extent residuals of this regression should reflect performance. Whether this measure is preferable to simpler ones depends on how strongly syllable traits influence respiratory demands in different species. Forstmeier et al. (2002) found that gains of extrapair paternity in the dusky warbler correlate with a measure of song performance that mostly captures the relative length of syllables and intervals within song. This suggests that, for dusky warblers, controlling for other syllable traits is not important. Our results indicate that in the dark-eyed junco other syllable traits are related with interval length, so that controlling for them may be an improvement.

We suggest that bird receivers evaluate performance by taking into account the costs of many song traits, and not only the respiratory costs addressed with the measures of performance in the previous paragraph. Amplitude trade-offs seem to be the best currency to infer performance in such an inclusive way because more traits trade off with amplitude than with interval length and this also explains a much higher proportion of variation of amplitude than of interval length. Curiously, dark-eyed juncos have a distinct class of songs that they always sing at very low amplitude and usually in close proximity to conspecifics, called short-range song (Titus 1998). Short-range songs are more phonologically diverse than regular song and contain many traits not observed in the latter (e.g., wider frequency bandwidths and clusters of different syllables; spectrograms in Titus 1998). This is consistent with the idea that trade-offs with sound amplitude constrain song phonology in this species, although we cannot exclude the possibility that low amplitude and more elaborated songs are both due to the particular situations during short-range singing (e.g., different social circumstances, less sound degradation, or greater risk of predation).

For the above reasons, integrating syllable traits with amplitude should be an important part of how birds assess song performance. As before, a measure of amplitude-related performance can be the residual from the multiple regression equation of sound amplitude on syllable traits (Table 2). This is feasible in laboratory conditions, where absolute amplitude can be measured easily, but in field studies, such as ours, it is virtually impossible to measure absolute amplitude accurately enough to compare between individuals. An alternative approach for field data is to plug measurements of the non-amplitude song traits into the multiple regression equation, and the result indicates how difficult songs are to sing at high amplitude; the songs that are predicted to be more difficult to sing at high amplitude are best considered to be high-performance songs.

We conclude that different song traits appear to incur costs in different dimensions of song output either requiring longer intersyllable intervals or diminishing sound amplitude. Trade-offs with interval length and amplitude are independent from each other, suggesting that receivers may obtain different information from each. We suggested measures of performance based on each of these 2 trade-offs. The validity of the measures of performance suggested here, or others, must be settled empirically, based on the birds' response to them. Researchers must also establish if birds truly respond to the interaction of traits used to infer performance or simply to individual song traits.

SUPPLEMENTARY MATERIAL

Supplementary Figure 1 and Tables 1–3 can be found at http://www.beheco.oxfordjournals.org/

FUNDING

Fundação para a Ciência e a Tecnologia (Postdoctoral grant SFRH/BPD/21509/2005 to G.C.C.); National Science

Foundation (graduate research fellowship to J.W.A.); Indiana University Faculty Research Support Program.

We would like to thank Russell Lande and Hopi Hoekstra for logistic support at University of California at San Diego. The editor and 3 anonymous reviewers provided constructive comments that improved the article.

REFERENCES

Ballentine B. 2006. Morphological adaptation influences the evolution of a mating signal. Evolution. 60:1936–1944.

Ballentine B, Hyman J, Nowicki S. 2004. Vocal performance influences female response to male bird song: an experimental test. Behav Ecol. 15:163–168.

Beebee MD. 2004. Variation in vocal performance in the songs of a wood warbler: evidence for the function of distinct singing modes. Ethology. 110:531–542.

Brackenbury JH. 1978. A comparison of the origin and temporal arrangement of pulsed sounds in the songs of the grasshopper and sedge warblers, *Locustella naevia* and *Acrocephalus schoenobaenus*. J Zool. 184:187–206.

Cardoso GC, Mota PG, Depraz V. 2007. Female and male serins (*Serinus serinus*) respond differently to derived song traits. Behav Ecol Sociobiol. 61:1425–1436.

Catchpole CK, Slater PJB. 1995. Bird song. Biological themes and variations. Cambridge: Cambridge University Press.

Christensen R, Kleindorfer S, Robertson J. 2006. Song is a reliable signal of bill morphology in Darwin's small tree finch *Camarhynchus parvulus*, and vocal performance predicts male pairing success. J Avian Biol. 37:617–624.

Cramer ERA, Price JJ. 2007. Red-winged blackbirds *Ageliaus phoeniceus* respond differently to song types with different performance levels. J Avian Biol. 38:122–127.

Drăgănoiu TI, Nagle L, Kreutzer M. 2002. Directional female preference for an exaggerated trait in canary (*Serinus canaria*) song. Proc R Soc Lond B. 269:2525–2531.

Forstmeier W, Kempanaers B, Meyer A, Leisler B. 2002. A novel song parameter correlates with extra-pair paternity and reflects male longevity. Proc R Soc Lon B. 269:1479–1485.

Gil D, Gahr M. 2002. The honesty of bird song: multiple constraints for multiple traits. Trends Ecol Evol. 17:133–141.

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

Huber SK, Podos J. 2006. Beak morphology and song features covary in a population of Darwin's finches (*Geospiza fortis*). Biol J Linn Soc Lond. 88:489–498.

Illes AE, Hall ML, Vehrencamp SL. 2006. Vocal performance influences male receiver response in the banded wren. Proc R Soc Lond B. 273:1907–1912.

Konishi M. 1964. Song variation in a population of Oregon juncos. Condor. 66:423–436.

Mota PG, Cardoso GC. 2001. Song organisation and patterns of variation in the serin (*Serinus serinus*). Acta Ethol. 3:141–150.

Nowicki S, Searcy WA, Peters S. 2002. Brain development, song learning and mate choice in birds: a review and experimental test of the nutritional hypothesis. J Comp Physiol A. 188:1003–1014.

Podos J. 1996. Motor constraints on vocal development in a songbird. Anim Behav. 51:1061–1070.

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

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

Podos J, Huber SK, Taft B. 2004. Bird song: the interface of evolution and mechanism. Annu Rev Ecol Evol Syst. 35:55–87.

Searcy WA. 1996. Sound-pressure levels and song preferences in female red-winged blackbirds (*Agelaius phoeniceus*) (Aves, Emberizidae). Ethology. 102:187–196.

Searcy WA, Nowicki S. 2005. The evolution of animal communication. Princeton (NJ): Princeton University Press.

Suthers RA, Goller F, Pytte C. 1999. The neuromuscular control of bird song. Philos Trans R Soc Lond B. 354:927–939.

Suthers RA, Vallet E, Tanvez A, Kreutzer M. 2004. Bilateral song production in domestic canaries. J Neurobiol. 60:381–393.

Titus RC. 1998. Short-range and long-range songs: use of two acoustically distinct song classes by dark-eyed juncos. Auk. 115: 386–393.

Thomas RJ. 1999a. Two tests of a stochastic dynamic programming-model of daily singing routines in birds. Anim Behav. 57: 277–284.

Thomas RJ. 1999b. The effect of variability on the food supply on the daily singing routines of European robins: a test of the stochastic dynamic programming model. Anim Behav. 57:365–369.

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

Vallet E, Kreutzer M. 1995. Female canaries are sexually responsive to special song phrases. Anim Behav. 49:1603–1610.

Vehrencamp SL. 2000. Handicap, index, and conventional signal elements of bird song. In: Espmark Y, Amundsen T, Rosenqvist G, editors. Animal signals: signalling and signal design in animal communication. Trondheim, Norway: Tapir Academic Press. p. 277–300.

Yeh PJ, Price TD. 2004. Adaptive phenotypic plasticity and the successful colonization of a novel environment. Am Nat. 164:532–542.