

Vocal performance influences female response to male bird song: an experimental test

Barbara Ballentine, Jeremy Hyman, and Stephen Nowicki

Department of Biology, Duke University, Box 90338, Durham, NC 27708-0338, USA

Female songbirds are thought to assess males based on aspects of song, such as repertoire size or amount of singing, that could potentially provide information about male quality. A relatively unexplored aspect of song that also might serve as an assessment signal is a male's ability to perform physically challenging songs. Trilled songs, such as those produced by swamp sparrows (*Melospiza georgiana*), present males with a performance challenge because trills require rapid and precise coordination of vocal tract movements, resulting in a trade-off between trill rate and frequency bandwidth. This trade-off defines a constraint on song production observed as a triangular distribution in acoustic space of trill rate by frequency bandwidth, with an upper boundary that represents a performance limit. Given this background on song production constraints, we are able to identify a priori which songs are performed with a higher degree of proficiency and, thus, which songs should be more attractive to females. We determined the performance limit for a population of swamp sparrows and measured how well individual males performed songs relative to this limit ("vocal performance"). We then compared female solicitation responses to high-performance versus low-performance versions of the same song type produced by different males. Females displayed significantly more to high-performance songs than to low-performance songs, supporting the hypothesis that females use vocal performance to assess males. *Key words:* bird song, female choice, indicator mechanism, *Melospiza georgiana*, motor constraint, sexual selection, vocal performance. [*Behav Ecol* 15:163–168 (2004)]

It is now well established that female mate choice in many songbird species is influenced by male singing behavior (Catchpole and Slater, 1995; Searcy and Yasukawa, 1996). Why females should pay attention to male song when choosing mates is less well understood, however. Song preferences are expected to evolve if, by mating with males having certain song characteristics and not others, females gain either direct material benefits or indirect genetic benefits for their offspring (Andersson, 1994; Kirkpatrick and Ryan, 1991). The question is how a male's singing behavior can serve as an accurate indicator of his quality in terms of his ability to provide direct or indirect benefits (Gil and Gahr, 2002). For a trait to be an accurate indicator of male quality, there must be some cost involved in producing that trait which ensures its honesty (Grafen, 1990; Johnstone and Grafen, 1993). For example, it costs something simply to sing, both in time and energy (Eberhardt, 1994; Oberweger and Goller, 2001), and thus, the ability of a male to sing more may be an accurate indicator of his condition (Greig-Smith 1982) or of the quality of his territory (Radesäter and Jakobsson, 1989). Correspondingly, females of several species have been shown to prefer males that sing longer songs or that sing more often (Alatalo et al., 1990; Eens et al., 1991; Kempenaers et al., 1997). The size or complexity of a male's vocal repertoire is another song feature widely reported to influence female choice (Andersson, 1994; Searcy and Andersson, 1986; Searcy and Yasukawa, 1996). In this case, the ability to produce a larger or more complex repertoire may be limited by developmental costs associated with brain growth and song learning (Buchanan, 2000; Catchpole, 1996; Nowicki et al., 1998, 2000, 2002).

Another feature of song that could potentially serve as an indicator of male condition is the ability to sing physically challenging songs (Suthers and Goller, 1997; Vallet and Kreuzer, 1995; Drăgănoiu et al., 2002). For example, female canaries (*Serinus canaria*) respond more to songs including

syllable types that appear to require more complex motor patterns for their production (Vallet and Kreuzer, 1995; Vallet et al., 1998). If female songbirds attend to subtle variation in songs associated with individual differences in male vocal performance, this information may serve, in turn, as an indicator of male quality. In the present study, we experimentally test a key prediction of this hypothesis, by asking whether female swamp sparrows are more responsive to the same song types produced by different males that vary in the degree of proficiency with which they produce those songs. Swamp sparrows are particularly suited for this experiment because our knowledge of song production mechanisms allows us to identify a priori which songs are produced with greater vocal proficiency.

Swamp sparrow songs typically comprise a single multinote syllable repeated in a continuous trill (Figure 1). The production of such trilled songs is limited by a constraint stemming from the need to coordinate syringeal and respiratory activity during song production with vocal tract motions, notably beak movements that dynamically modify vocal tract resonance (Nowicki et al., 1992; Podos, 1996; Suthers and Goller, 1997). To achieve this coordination, songbirds must open and close their beaks in register with the acoustic frequency of the sound being produced (Hoese et al., 2000; Westneat et al., 1993). In species such as the swamp sparrow that produce trills composed of rapid frequency-modulated notes (Figure 1), there is a performance trade-off between how fast a bird can repeat syllables in a trill ("trill rate") and how broad a range of frequencies each repeated syllable can encompass ("frequency bandwidth") (Podos, 1997). Specifically, the role of beak movements in song production predicts that birds producing trills with slow repetition rates should be able to include syllables having either narrow or wide frequency bandwidths; as trill rate increases; however, repeating syllables will be constrained to increasingly narrow bandwidths. Thus, the distribution of songs in the space defined by trill rate and bandwidth is expected to be triangular (Figure 2) (Podos, 1997), with the

Address correspondence to B. Ballentine. E-mail: beb5@duke.edu.
Received 15 October 2002; revised 4 March 2003; accepted 17 March 2003.

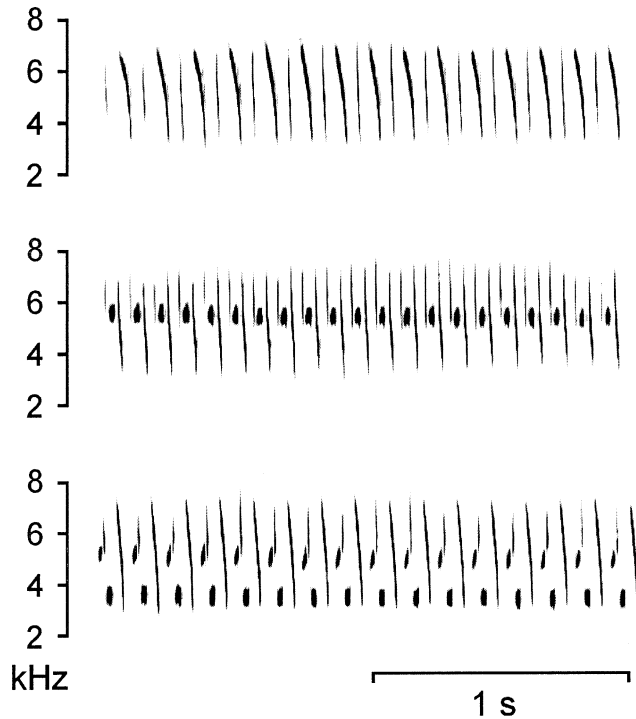


Figure 1

Sonograms of three song types from a single swamp sparrow, each comprising a single repeated syllable. The syllables in this example include two, three, and four distinct notes, respectively, although some songs may include syllables with as many as five notes. Trill rate is measured in terms of the number of syllables produced per second; frequency bandwidth is measured as the difference between the highest and lowest frequency, measured at a standard level of -36 dB relative to the highest amplitude frequency in the song.

diagonal representing an estimate of the performance limit for trill production and the orthogonal deviation from this line being a measure of vocal performance (Podos, 2001). To the extent that it is physically or physiologically more demanding to produce songs closer to a performance limit (Lambrechts, 1996), we expect variation among males in their ability to meet these demands. By assessing vocal performance, females may gain information about a male's phenotypic quality and, thus, information about the direct or indirect benefits she may receive from him.

To test the hypothesis that female swamp sparrows assess males based on vocal performance, we first measured a large sample of songs recorded from a wild population of swamp sparrows. We estimated a performance limit boundary for this population and then identified identical song types produced by different males that varied in how closely they approached this performance limit. Finally, we used the copulation solicitation assay (King and West, 1977; Searcy, 1992) to determine whether female response to song varies as a function of vocal performance.

METHODS

Male subjects and song recording

We captured 114 male swamp sparrows at Conneaut Marsh, Crawford County, Pennsylvania, USA, during May–June, 2001, marking them with unique combinations of colored leg bands to facilitate field identification. Because beak morphology may influence vocal performance (Podos, 2001), we measured

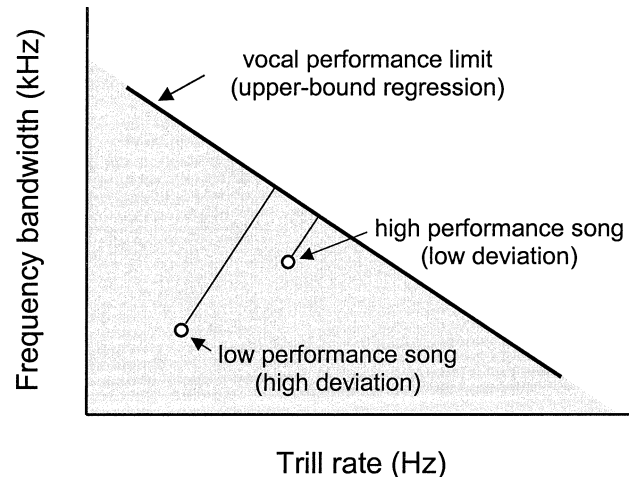


Figure 2

Schematic illustration of the predicted triangular distribution of songs in an acoustic space defined by bandwidth as a function of trill rate (Podos, 1997). The shaded area represents the region in which songs are normally produced, and the diagonal line represents the upper-bound regression line that operationally defines a motor performance limit. "Vocal deviation" (Podos, 1997) of a song is measured as the minimum orthogonal distance from the point representing an individual male's frequency bandwidth and trill rate to the upper-bound regression line. High-performance songs have small deviations from the upper-bound regression, whereas low-performance songs have large deviations.

bill length (exposed culmen), bill depth (depth at widest point), and bill width (width at the base of the bill) from each bird with vernier calipers (Tajima) to the nearest 0.05 mm. We noted the location where each male was captured on a map of the study site, as well as any other position at which it was observed during the study. We recorded songs from these males by using a Sony TCM-5000 EV or Sony TCD-5M recorder, a Shure SM-57 microphone, and a Sony Parabolic Reflector-330 or a Saul Mineroff SME Parabolic Reflector-1000.

We visually inspected sonograms of all recorded songs ($n = 3871$) to identify the song repertoires of each male (Kay Elemetrics DSP 5500 Sona-Graph; frequency range = 0–8 kHz, frequency resolution = 300 Hz, temporal resolution = 4.8 ms). We obtained recordings from 91 males that were of sufficient quality to be included in our analysis, with an average of 3.1 song types obtained from each male (range = 1–4) for a total of 280 song types in our sample. Many of the song types in our sample were produced by more than one male in the population. To determine which song types were shared by different males, two independent observers sorted the entire sample of 280 song types into categories without reference to male identity, yielding a total of 30 unique song types in the population. Interobserver agreement was 100%, which is not surprising given that swamp sparrow song types are easily distinguished on the basis of unique combinations of notes found in the repeating syllable (Marler and Pickert, 1984). Because we did not record birds exhaustively, it is possible we did not obtain the complete song type repertoire of every male in our sample. The average number of song types/male we recorded, however, is the typical repertoire size observed for this species (Mowbray, 1997).

Acoustic analyses

We digitized a mean of 6.3 exemplars per song type for each song type in each male's repertoire (range = 1–10). Songs

were digitized with 12-bit precision at 25 kpts/s (Data Translation DT2821-F), after high-pass filtering to eliminate noise below the frequency range of swamp sparrow song (Krohn-Hite 3500, 1-kHz corner frequency, 24 dB/octave) and low-pass filtering to prevent aliasing (Stanford Research Systems SR 640, 10-kHz corner frequency, 115 dB/octave). All acoustic analyses were performed by using Signal version 3.1 digital signal processing software (Engineering Design).

We calculated the trill rate for each song in our digitized sample from oscillograms, by determining the instantaneous trill rate (calculated as pulses/sec) for each syllable occurring in approximately the first 2 s of the song and then averaging across syllables. We measured the frequency bandwidth of songs from digital spectrograms (256 pt transform, frequency resolution = 98 Hz) by calculating the difference between minimum and maximum frequencies measured at -36 dB relative to the peak amplitude frequency in the song. This -36 -dB criterion was chosen because it captured variation in the frequency characteristics of songs in our sample while at the same time excluding extraneous background noise. Trill rate and frequency bandwidth data for each song type in each male's repertoire represent the mean of values obtained from the multiple exemplars we analyzed for each type.

We calculated an upper-bound regression from the plot of frequency bandwidth as a function of trill rate for the 280 song types ($n = 91$ males) in our sample, following the method of Podos (1997). This approach estimates the location of an upper boundary for data with a triangular distribution (Blackburn et al., 1992). Briefly, song types were binned by trill rate in 1-Hz increments, and then a regression was performed on the subset of data points representing the maximum frequency bandwidth within each trill-rate bin ($n = 10$ bins beginning at 4 Hz). This upper-bound regression represents an estimate of the performance limit boundary for the trade-off between the rate and magnitude of beak movement during trill production (Podos, 1997). We next calculated the "vocal deviation" (Podos, 2001) of all 280 song types in our distribution as the minimum orthogonal distance from each point to the upper-bound regression line (Figure 2). The vocal deviation thus is a measure of relative vocal performance, with lower deviations (points closer to the line) representing higher performance songs and vice versa.

Female subjects and experimental design

We used the solicitation display assay (King and West, 1977; Searcy, 1992) to compare female response to high-performance versus low-performance songs. We collected 10 female swamp sparrows during June 2001 from the same population where we had recorded males, noting the location where each bird was captured on a map of the study site. Females were maintained in captivity until testing began in February 2002. To bring females into breeding condition when tested, we accelerated the seasonal shortening and lengthening of day length with the shortest day (9-h light/15-h dark) occurring in early October 2001 and the longest day (15-h dark/9-h light) occurring at the beginning of February 2002. To help control for individual variation in hormonal state, females were given a subcutaneous implant of 17- β -estradiol in silastic tubing of 1.96-mm outer diameter, containing 7–9 mm of hormone 10 days before trials began (Searcy, 1992). During testing, females were housed singly in sound attenuation chambers (AC-1, Industrial Acoustics). Each chamber included a loudspeaker (Realistic 40-1298C) from which songs were played via computer (amplitude = 83 ± 1 dB SPL) and a window through which female behavior was videotaped.

We compared female response to different versions of the same song type as produced by different males, with one

version being high performance and the other version being low performance, measured relative to the upper-bound regression performance limit for this population. High-quality recordings were chosen as exemplars from the pool of recorded songs to maximize the difference in performance within stimulus pairs as determined by vocal deviation measurements. High-performance and lower-performance versions of song types were tested in separate trials, with two song types (both either high or low performance) heard in each trial to minimize habituation; thus, in each trial, a female heard nine presentations of the first song type and then nine presentations of the second, at a rate of six songs/min for a total length of 3 min. Females were tested with contrasting stimulus sets (i.e., high-performance song types versus low-performance song types) on the same day, with approximately 3 h between trials and with the order of presentation (i.e., high or low performance first) randomized. Two days later, each female was tested again with the same stimulus sets, also separated by 3 h, but with the order of presentation reversed. Individual female response to high- and low-performance stimulus sets was totaled over the 2 days of testing to minimize presentation order effects.

We used the total number of displays performed during the trial as our response measure, as in previous studies (see Nowicki et al., 2001). Exemplars used to create the 10 pairs of contrasting songs were obtained from 34 different males and included 14 of the song types found in our population (Table 1). No song type produced by the same male was used more than once in constructing stimulus sets, however, and each female was tested with a unique set of high- and low-performance song types to avoid pseudoreplication (McGregor et al., 1992), resulting in a sample size for statistical comparison of $n = 10$. Females were tested with songs that were recorded from at least five territories away from the territory in which females were captured to minimize the potential confound of familiarity with the male from which songs were recorded.

RESULTS

Frequency bandwidth varies positively with trill rate for all songs, but the relationship is highly scattered as expected for a triangular distribution ($r^2 = .031$, $p = .003$, $n = 280$) (Figure 3). The upper-bound regression for this population has a significant negative slope ($y = -99.3$, $x = 6550$, $r^2 = .561$, $p = .013$, $n = 10$). By using an ANCOVA (Sokal and Rohlf, 1981), we found no significant difference between our upper-bound regression line and the line reported by Podos (1997) ($F = 0.158$, $p = .696$, $n = 20$). There is considerable variation among males in their relative vocal performance abilities when measured as the average deviation from the performance limit regression line and averaging across all song types in an individual's repertoire (mean \pm SD = 6.4 ± 5.5 , range = -4.2 to 24.4 , CV = 0.87). Some of this between-male variation is due to variation in the average deviation of different song types (i.e., some song types consistently have low deviations and others high deviations regardless of which male sang them, suggesting that some song types are harder to produce than others) and the fact that males differed in the number of low-deviation and high-deviation song types they included in their repertoires. However, males that shared the same song type also differed considerably and consistently in how well they produced these song types (Table 1, overall mean CV = 0.48). We used a model II ANOVA to determine, for males singing the same song type, if between-male variation was significantly greater than within male variation in vocal performance (Sokal and Rohlf, 1981). By using the variance components from the ANOVA we calculated repeatability

Table 1
Summary of vocal deviation and trill rate, frequency bandwidth for males in the population for song types used in female tests

Song type	No. of males	Vocal deviation	Trill rate (Hz)	Bandwidth (Hz)	<i>F</i> ratio ^a	Repeatability ^b
E	30	14.0 ± 4.7	6.1 ± 0.7	4560 ± 430	<i>F</i> _{29,166} = 34.0	0.84
D	23	7.3 ± 5.4	9.5 ± 1.6	4880 ± 490	<i>F</i> _{22,105} = 26.6	0.80
I	21	8.0 ± 6.8	8.0 ± 0.8	4960 ± 650	<i>F</i> _{20,124} = 38.2	0.84
C	20	11.7 ± 4.1	6.5 ± 0.8	4750 ± 420	<i>F</i> _{19,101} = 13.7	0.65
M	19	4.6 ± 3.6	10.7 ± 0.7	5040 ± 360	<i>F</i> _{18,55} = 9.8	0.68
Y	20	12.5 ± 5.0	4.9 ± 0.4	4820 ± 500	<i>F</i> _{19,101} = 19.3	0.74
K	17	8.5 ± 4.7	11.2 ± 1.1	4590 ± 420	<i>F</i> _{16,67} = 16.3	0.75
H	15	6.6 ± 4.9	6.0 ± 0.6	5300 ± 450	<i>F</i> _{14,84} = 34.5	0.84
N	13	14.6 ± 5.3	6.5 ± 0.9	4460 ± 500	<i>F</i> _{12,75} = 15.5	0.66
J	12	10.3 ± 4.6	6.5 ± 0.7	4880 ± 470	<i>F</i> _{11,84} = 31.2	0.78
G	11	14.8 ± 4.1	5.2 ± 0.7	4570 ± 370	<i>F</i> _{10,70} = 22.0	0.73
F	11	15.2 ± 4.6	4.5 ± 0.2	4660 ± 470	<i>F</i> _{10,50} = 8.5	0.54
A	9	11.1 ± 3.8	8.6 ± 2.1	4600 ± 440	<i>F</i> _{8,54} = 19.4	0.73
U	7	5.8 ± 2.9	7.5 ± 1.4	5220 ± 270	<i>F</i> _{6,50} = 10.2	0.53

Values are mean ± SD. For each song type, *F* ratios are reported from model 2 ANOVAs and repeatability measures calculated from ANOVA components.

^a Results of model 2 ANOVA designed to estimate components of variation rather than for hypothesis testing (Sokal and Rohlf, 1981). For all reported *F* ratios, *p* < .0001.

^b Repeatability is calculated as the proportion of total variation in vocal deviation due to differences between individuals based on repeated measures of the same individual calculated from the variance components of the ANOVA (Lessells and Boag, 1987).

for males singing subsequent renditions of the same song type (Boake, 1989; Lessells and Boag, 1987). In all cases, variation between males was significantly greater than variation within males (Table 1). Repeatability measures were high for most song types in our population (range = 0.53–0.84) (Table 1). We found no relationship between vocal deviation and any measure of beak morphology (beak width, *r* = −.005; beak depth, *r* = .071; beak length, *r* = .048, *p* > .05 in each case).

The songs used in our two treatments (high versus low performance) differed significantly in their deviation from the vocal performance boundary (mean ± SD = 2.50 ± 4.05 versus 16.83 ± 3.50, *T*_{paired} = −11.97, *p* < .001, *n* = 20). Females responded to both high and low vocal performance songs, but gave significantly more copulation solicitation displays in response to high vocal performance than to low vocal performance songs (mean ± SD = 14.3 ± 8.6 versus 10.1 ± 10.3, Wilcoxon paired-comparison signed-ranks test *Z* = −2.194, *p* = .028). In nine of the 10 pair-wise comparisons, high-performance songs elicited stronger responses than did their corresponding low-performance songs. We conclude that songs produced closer to the population-wide vocal performance limit function better to elicit female sexual response than song produced farther from this limit.

DISCUSSION

The trade-off between frequency bandwidth and trill rate predicted by our hypothesis of motor constraints on the production of trilled songs predicts a triangular distribution of songs in the acoustic space defined by these two variables, with an upper boundary that operationally defines a vocal performance limit (Podos, 1996, 1997). The distribution of songs that we measured from our population of swamp sparrows (Figure 3) is similar to that found in a comparative study of 34 species in the family Emberizidae (which includes the swamp sparrow), with a significant negative upper-bound regression that is statistically indistinguishable from the line obtained for the family as a whole (Podos 1997). Our large sample also allowed us to compare vocal abilities among individuals in the population, many of which shared the same song types. We found considerable variation in male vocal

performance measured as the deviation from the upper-bound regression averaged across all songs in an individual's repertoire. Of more significance to the present analysis, we also found considerable variation in how well different males produce the same song type (Table 1). That is, among males sharing song types in the population, we found that some males sang high-performance versions of those song types (i.e., with a small deviation from the trill rate × frequency bandwidth upper-bound regression), whereas other males sang low-performance versions of the same types (with a large deviation). Furthermore, individual males consistently produced particular song types at the same level of vocal performance, as evidenced by high repeatability scores and significantly higher between-male versus within-male variation for each song type (Table 1). In her review of the use of repeatability in studies of the evolution of mating behavior, Boake (1989) reported a range of repeatabilities for traits known to be used in courtship and mate attraction signals from a low of 0.21 to a high of 0.85. Our measures of vocal performance fall squarely into the high end of this range, suggesting that this feature of song could provide a reliable source of information to females. Thus, our results support the existence of a performance limit in the production of swamp sparrow songs and further demonstrate that there is measurable and consistent variation among males in their ability to approach this performance limit.

The females in our solicitation assays displayed significantly more to high-performance versions of song types than they did to low-performance versions of those same types. Given that solicitation assay results have been shown to correspond well to mating preferences in several species (West et al., 1981; Catchpole, 1986; Catchpole et al., 1984), this result supports the hypothesis that female swamp sparrows may use vocal performance to assess males in the context of mate choice. Vallet et al. (1998) reached a similar conclusion based on their finding that female canaries display more to songs that include particular kinds of phrases (see also Vallet and Kreuzer, 1995). They suggested post hoc that these preferred phrase types might be more challenging to produce based on their acoustic complexity, although a specific mechanism responsible for limiting a male's ability to produce these phrases was not identified. Similarly, in a field study of dusky

warblers (*Phylloscopus fuscatus*), Forstmeier et al. (2002) found that a male's ability to maintain high amplitude during song output, used as a measure of "singing performance," predicted relative paternity losses and gains by males. Forstmeier et al.'s (2002) data are consistent with the hypothesis that females judge male quality based on some measure of song performance, but it is difficult to rule out the alternative hypotheses that females are cuing in on some other factor that happens to correlate with vocal proficiency or that observed effects on paternity stem from male-male competition and are not the result of female choice at all. By experimentally comparing female response to high- and low-performance versions of the same song types as produced by different males, with an a priori prediction about which songs should elicit more responses, we can conclude with certainty that females are attending to subtle differences in song reflecting male motor capabilities. Thus, our results provide a crucial piece of evidence in support of the general hypothesis that female birds assess male quality on the basis of vocal performance.

In the present study, the physical basis of the performance limit seems clear: there is a trade-off between how fast and how wide a bird can open and close its beak, which means it is more difficult to produce a broad frequency bandwidth at higher trill rates (Nowicki et al., 1992; Podos, 1996, 1997; Podos and Nowicki, 2003). Further, our experiment compared female response to the same song types as produced by different males with varying proficiency. The only consistent difference between our high-performance and low-performance songs was their relative deviation from a quantitatively determined performance limit for the population. Thus, our results demonstrate that females may pay attention not just to the inherent difficulty of the song a male is attempting to sing, as suggested by Vallet et al. (1998), but more specifically to how well a male can perform a song type it shares in common with other males in the population.

Constraints other than those imposed by beak movements are likely contribute to performance limits in the production of bird song more generally, with certain kinds of sounds more difficult to produce than others because of constraints on syringeal activity, respiratory activity, or the coordination of the two (Suthers and Goller, 1997). Thus, the trade-off between trill rate and frequency bandwidth probably is only an approximation of the relative physical challenge of producing a swamp sparrow syllable. Nonetheless, there is good evidence that this particular trade-off represents a real constraint on song production. When young male swamp sparrows tutored with songs that had been manipulated to create trill rates elevated above those normally produced by wild birds, they were able to memorize these accelerated models, but were unable to reproduce them accurately (Podos, 1996). The kinds of inaccuracies these birds introduced into their songs supported the idea that vocal production is constrained by physical limitations of the vocal apparatus: models were reproduced with slower trill rates, with notes omitted, or with a "broken syntax" in which birds produced songs as bursts of syllables separated by silent gaps (Podos, 1996). Similarly, canaries that have had song production made more challenging by temporarily adding weight to their beaks produce songs with an altered tonal quality caused by an inability to move the beak as quickly as in normal singing (Hoese et al., 2000).

Our current model for how motor constraints affect song production predicts that vocal performance should vary as a function of vocal tract morphology, especially beak morphology (Nowicki et al. 1992; Podos, 1997), with larger beaks expected to be less able to produce rapid movements required for fast broadband trills than smaller beaks.

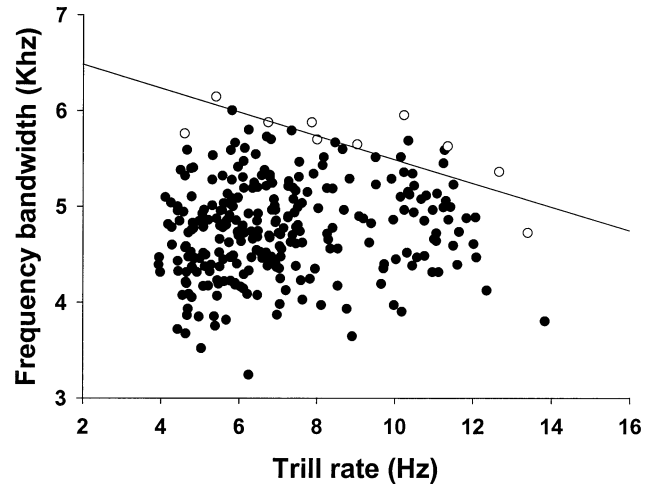


Figure 3
Trill rate by frequency bandwidth plot and upper-bound regression line for 280 songs recorded from 91 male swamp sparrows. Open circles represent the points that were used to calculate the upper-bound regression line.

Consistent with this hypothesis, Podos (2001) found that the deviation of a male's songs from the upper-bound regression between trill rate and frequency bandwidth is predicted by beak size, both across different species in which variation in beak morphology is expected to be pronounced, as well as among individuals in a single population of *Geospiza fortis*. We did not find a similar correspondence between beak morphology and vocal performance in our swamp sparrow population. Our failure to detect such a relationship is not surprising, however, given that Darwin's finches exhibit a particularly high degree of intraspecific morphological variation in beak characteristics driven by long-term fluctuations in food availability (Gibbs and Grant, 1987; Grant, 1999), and we have no reason to believe a similar pattern of selection is acting on our swamp sparrow population.

If male vocal performance is used by females to assess potential mates, then we expect those capabilities should correspond in some way to a male's ability to provide direct benefits to females, such as by holding a superior territory or providing better parental care, or indirect benefits in the form of genes that increase the vigor of her offspring. A male's ability to sing songs closer to a physical performance limit may be energetically more demanding (Lambrechts, 1996) and thus depend on his current condition, much the same way a lizard in better condition can achieve faster sprint speeds (Garland and Losos, 1994). Alternatively, variation in vocal capabilities may reflect longer-term phenotypic differences among males, such as in vocal tract morphology or in the peripheral or central neural mechanisms underlying song production, although we have yet to identify such correlates in swamp sparrows. Nonetheless, our finding that female swamp sparrows respond more to higher-performance songs coupled with our understanding of the physical basis of performance constraints and our ability to estimate song performance relative to a quantitatively-defined limit, suggests this system can provide an important new model for how song might serve as an honest indicator of male quality.

We thank Susan Peters for technical advice and assistance; Megan Insko for assistance; Jeff Podos, Bill Searcy, David Westneat, and two anonymous reviewers for helpful comments; the Pymatuning Laboratory of Ecology for logistical support; and the Pennsylvania State Game Commission for access to field sites. This research was

supported by grants from the McKinley Fund of the University of Pittsburgh, The American Ornithologist's Union, and Duke University Department of Biology to B.B., and from the National Science Foundation (IBN-9974743) to S.N.

REFERENCES

- Alatalo RV, Glynn C, Lundberg A, 1990. Singing rate and female attraction in the pied flycatcher: an experiment. *Anim Behav* 39: 601–603.
- Andersson MB, 1994. *Sexual selection*. Princeton, New Jersey: Princeton University Press.
- Blackburn TM, Lawton JH, Perry JN, 1992. A method of estimating the slope of upper bounds of plots of body size and abundance in natural animal assemblages. *Oikos* 65:107–112.
- Boake CR, 1989. Repeatability: its role in evolutionary studies of mating behavior. *Evol Ecol* 3:173–182.
- Buchanan KL, 2000. Stress and the evolution of condition-dependent signals. *Trends Ecol Evol* 15:156–160.
- Catchpole CK, 1986. Song repertoires and reproductive success in the great reed warbler *Acrocephalus arundinaceus*. *Behav Ecol Sociobiol* 19:439–445.
- Catchpole CK, 1996. Song and female choice: good genes and big brains? *Trends Ecol Evol* 11:358–360.
- Catchpole CK, Dittami J, Leisler B, 1984. Differential responses to male song in female songbirds implanted with oestradiol. *Nature* 312:563–564.
- Catchpole CK, Slater PJB, 1995. *Bird song*. Cambridge: University Press.
- Drăgănoiu TI, Nagle L, Kreutzer M, 2002. Directional female preference for an exaggerated male trait in canary (*Serinus canaria*) song. *Proc R Soc Lond B* 269:2525–2531.
- Eberhardt LS, 1994. Oxygen consumption during sing by male Carolina wrens (*Thryothorus ludovicianus*). *Auk* 111:124–130.
- Eens M, Pinxten R, Verheyen RF, 1991. Male song as a cue for mate choice in the European starling. *Behaviour* 116:21–238.
- Forstmeier W, Kempenaers B, Meyer A, Leisler B, 2002. A novel song parameter correlates with extra-pair paternity and reflects male longevity. *Proc R Soc Lond B* 269:1479–1485.
- Garland TJ, Losos JB, 1994. Ecological morphology of locomotor performance in squamate reptiles. In: *Ecological morphology: integrative organismal biology* (Wainwright PC, Reilly SM, eds). Chicago: University of Chicago Press; 240–302.
- Gibbs HL, Grant PR, 1987. Oscillating selection on Darwin's finches. *Nature* 327:511–513.
- Gil D, Gahr M, 2002. The honesty of bird song: multiple constraints for multiple traits. *Trends Ecol Evol* 17:133–141.
- Grafen A, 1990. Biological signals as handicaps. *J Theor Biol* 144: 517–546.
- Grant PR, 1999. *Ecology and evolution of Darwin's finches*. Princeton, New Jersey: Princeton University Press.
- Greig-Smith PW, 1982. Song rates and parental care by individual male stonechats *Saxicola torquata*. *Anim Behav* 30:245–252.
- Hoese WJ, Podos J, Boetticher NC, Nowicki S, 2000. Vocal tract function in birdsong production: experimental manipulation of beak movements. *J Exp Biol* 203:1845–1855.
- Johnstone RA, Grafen A, 1993. Dishonesty and the handicap principle. *Anim Behav* 46:759–764.
- Kempenaers B, Verheyen GR, Dhondt AA, 1997. Extrapair paternity in the blue tit (*Parus caeruleus*): female choice, male characteristics, and offspring quality. *Behav Ecol* 8:481–492.
- King AP, West MJ, 1977. Species identification in the North American cowbird: appropriate responses to abnormal song. *Science* 195: 1002–1004.
- Kirkpatrick M, Ryan MJ, 1991. The evolution of mating preferences and the paradox of the lek. *Nature* 350:33–38.
- Lambrechts MM, 1996. Organization of birdsong and constraints on performance. In: *Ecology and evolution of acoustic communication in birds* (Kroodsma DE, Miller EH, eds). Ithaca, New York: Cornell University Press; 305–320.
- Lessells CM, Boag PT, 1987. Unrepeatable repeatabilities: a common mistake. *Auk* 104:116–121.
- Marler P, Pickert R, 1984. Species-universal microstructure in the learned song of swamp sparrow (*Melospiza georgiana*). *Anim Behav* 32:673–689.
- McGregor PK, Catchpole CK, Dabelsteen T, Falls JB, Fusani L, Gerhardt HC, Gilbert F, Horn AG, Klump GM, Kroodsma DE, Lambrechts MM, McComb KE, Nelson DA, Pepperberg IM, Ratcliffe L, Searcy WA, Weary DM, 1992. Design of playback experiments: the Thornbridge Hall NATO ARW Consensus. In: *Playback and studies of animal communication* (McGregor PK, ed). New York: Plenum Press; 1–9.
- Mowbray TB, 1997. Swamp sparrow. In: *The birds of North America* (Poole A, Gill F, eds). Philadelphia: The Academy of Natural Sciences, The American Ornithologist's Union; 1–24.
- Nowicki S, Hasselquist D, Bensch S, Peters S, 2000. Nestling growth and song repertoire size in great reed warblers: evidence for song learning as an indicator mechanism in mate choice. *Proc R Soc Lond B* 267:2419–2424.
- Nowicki S, Peters S, Podos J, 1998. Song learning, early nutrition and sexual selection in songbirds. *Am Zool* 38:179–190.
- Nowicki S, Searcy WA, Hughes M, Podos J, 2001. Motor performance and the evolution of birds song: male and female response to song innovation in swamp sparrows. *Anim Behav* 62:1189–1195.
- Nowicki S, Searcy WA, Peters S, 2002. Brain development, song learning and mate choice in birds: a review and experimental test of the "nutritional stress hypothesis." *J Comp Physiol A* 188: 1003–1014.
- Nowicki S, Westneat M, Hoese W, 1992. Birdsong: motor function and the evolution of communication. *Sem Neurosci* 4:385–390.
- Oberweger K, Goller F, 2001. The metabolic cost of birdsong production. *J Exp Biol* 204:3379–3388.
- 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 signals in Darwin's finches. *Nature* 409:185–188.
- Podos J, Nowicki S, 2003. Performance limits on birdsong production. In: *Nature's musicians: the science of birdsong* (Marler P, Slabbekoorn H, eds) (in press).
- Radesäter T, Jakobsson S, 1989. Song rate correlations of replacement territorial willow warblers *Phylloscopus trochilus*. *Ornis Scand* 20: 71–73.
- Searcy WA, 1992. Measuring responses of female birds to male song. In: *Playback and studies of animal communication* (McGregor PK, ed). New York: Plenum Press; 175–189.
- Searcy WA, Andersson M, 1986. Sexual selection and the evolution of song. *Annu Rev Ecol Syst* 17:507–533.
- Searcy WA, Yasukawa K, 1996. Song and female choice. In: *Ecology and evolution of acoustic communication in birds* (Kroodsma DE, Miller EH, eds). Ithaca, New York: Cornell University Press; 454–473.
- Sokal RR, Rohlf FJ, 1981. *Biometry*, 2nd ed. San Francisco: W.H. Freeman.
- Suthers RA, Goller F, 1997. Motor correlates of vocal diversity in songbirds. In: *Current ornithology* (Nolan J, ed). New York: Plenum Press; 235–288.
- 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.
- Westneat MW, Long JH Jr, Hoese W, Nowicki S, 1993. Kinematics of birdsong: functional correlations of cranial movements and acoustic features in sparrows. *J Exp Biol* 182:147–171.
- West MJ, King AP, Eastzer DH, 1981. Validating the female bioassay of cowbird song: relating differences in song potency to mating success. *Anim Behav* 29:490–501.