

# A novel song parameter correlates with extra-pair paternity and reflects male longevity

Wolfgang Forstmeier<sup>1,2\*</sup>, Bart Kempenaers<sup>3</sup>, Axel Meyer<sup>4</sup>  
and Bernd Leisler<sup>1</sup>

<sup>1</sup>Max Planck Research Centre for Ornithology, Vogelwarte Radolfzell, Schloss Moegglingen, Schlossallee 2, 78315 Radolfzell, Germany

<sup>2</sup>Theodor-Boveri-Institute, Department of Animal Ecology and Tropical Biology, Biozentrum, Am Hubland, 97074 Würzburg, Germany

<sup>3</sup>Max Planck Research Centre for Ornithology, Postfach 1564, 82305 Seewiesen, Germany

<sup>4</sup>Department of Biology, University of Konstanz, 78457 Konstanz, Germany

Although elaborate bird song provides one of the prime examples of a trait that evolved under sexual selection, it is still unclear whether females judge the quality of males by attributes of their song and whether these song features honestly signal a male's genetic quality. We measured the ability of male dusky warblers *Phylloscopus fuscatus* to maintain a high sound amplitude during singing, which probably reflects an individual's physiological limitations. This new measure of singing performance was correlated with male longevity and with extra-pair paternity, indicating that females who copulated with better singers obtained 'good genes' for their offspring. Our findings are consistent with the idea that females assess male quality by subtle differences in their performance during the production of notes, rather than by the quantity or versatility of song. In addition, observations on territorial conflicts indicate that attractive males invest less in competition over territories because they can reproduce via extra-pair paternity.

**Keywords:** female choice; genetic benefits; male quality; repertoire size; singing performance; viability

## 1. INTRODUCTION

Song is one of the most conspicuous aspects of bird behaviour (Catchpole & Slater 1995). Since Darwin (1881), it has been assumed that song is a sexually selected trait and song has been interpreted as the acoustic analogue of the peacock's tail (Catchpole 1987). However, there is controversy about what song features could honestly reflect male genetic quality (Searcy & Yasukawa 1996; Gil & Gahr 2002). To date, most studies have assessed male singing ability by measuring either how much a male sings (song rate) or how many different syllable types he sings (repertoire size). Several studies have shown a female preference for males with a high song rate (Catchpole & Slater 1995; Searcy & Yasukawa 1996; Buchanan & Catchpole 2000) or a large repertoire (Catchpole & Slater 1995; Hasselquist *et al.* 1996; Searcy & Yasukawa 1996; Catchpole 2000). Song rate could honestly reflect male quality because it is costly to allocate time and energy to advertising (Gil & Gahr 2002). However, males living in territories with more food can spend more time singing because they need less time for foraging (Davies & Lundberg 1984; Gottlander 1987; Reid 1987). Thus, song rate could reflect territory quality, rather than the intrinsic quality of the male (Searcy & Yasukawa 1996). Repertoire size may reflect male quality, because song-learning processes may be sensitive to environmental stress. Males that are better able to cope with this stress will learn better than others and therefore gain larger song repertoires (von Schantz *et al.* 1999; Catchpole 2000; Nowicki *et al.* 2000).

Recent physiological work suggests a new approach to investigating song features in relation to male quality. Singing males may reach physiological limitations during sound production (Lambrechts 1996; Podos 1996, 1997, 2001; Suthers & Goller 1997) and males may differ in how far they can push these limits. This led us to the idea that females might judge male quality from quite subtle differences in the sounds they produce. To date, the relationship between female mating preferences and the quality of sound production has hardly been investigated. This seems unfortunate because a promising study by Vallet *et al.* (1998) indicated that some parts of the song might be more difficult to produce than others and that females prefer males that are able to sing these difficult parts ('sexy syllables').

In this paper, we assess the importance of fine-scaled variation in sound production in male dusky warblers *Phylloscopus fuscatus*, a species with remarkable singing and mating behaviour (Forstmeier & Balsby 2002). In contrast to the majority of passerines (Birkhead & Møller 1992), male dusky warblers do not guard their fertile female, but spend most of the daytime singing. Also of interest is that males sing at the highest rate when their social partner is fertile (Forstmeier & Balsby 2002), which is consistent with the hypothesis that males sing to ensure within-pair paternity and to gain extra-pair paternity (Møller 1991). Hence, this species is an excellent model in which to test whether female choice for extra-pair copulation partners depends on differences in male singing abilities.

The song of a male dusky warbler comprises a substantial repertoire of different syllable types and neighbouring males share a significantly higher proportion of syllable types than would be expected by chance (Forstmeier &

\* Author for correspondence (forstmeier@mail.uni-wuerzburg.de).

Balsby 2002). Moreover, the similarity of their repertoires increases with the time they spend as neighbours (Forstmeier & Balsby 2002), which indicates that males continuously learn new elements by copying other males. As found in many other species (Catchpole & Slater 1995), counter-singing males often match their syllable types. This led us to the hypothesis that males compete with each other over the quality of their song and that females judge males by comparing their performance. When males match their syllable types, they may compete with each other over (i) how many syllable types they can sing (repertoire size); (ii) how fast they can repeat them within strophes (syllable rate); or (iii) how well they can sing a given syllable type (differences in the quality of sound production). We assess the quality of singing by measuring the percentage of time within strophes during which sound amplitude exceeds a threshold level (see § 2). We suggest that this is a good measure of singing quality because it reveals a male's ability to maintain a high sound amplitude during rapid frequency modulations and to sing with short pauses between syllables, which is difficult to achieve (Suthers & Goller 1997; Podos 1996). Note that this measure reflects subtle differences in sound production but has nothing to do with song rate (i.e. the number of strophes sung per unit of time; see § 2).

## 2. MATERIAL AND METHODS

The dusky warbler is a single-brooded, long-distance migrant passerine with a cryptic, sexually monomorphic plumage lacking any obvious sexual ornaments. We studied a population of individually colour-ringed birds breeding near Magadan, in the Russian Far-East (59°51' N, 154°14' E), from 1997 to 1999. This population shows a variable mating system with 11% of males unmated, 70% monogamous, 17% polygynous and one case of polyandry ( $n = 95$  males). We can unambiguously define a female's social partner as the owner of the territory where she builds the nest. Several male and territory characteristics were measured to test for correlations with singing behaviour. Male age was determined as first year or older partly based on ringing data and partly on a discriminant analysis using morphological characteristics (tail length, wing length and wing shape) of 65 males of known age. The latter analysis correctly classified 91% of these 65 males (Forstmeier 2001). Food availability on territories and male contribution to feeding the offspring (corrected for hatching date) was measured as described elsewhere (Forstmeier *et al.* 2001).

### (a) Song analysis

We recorded singing males using a Nagra SN recorder and a Sennheiser MKH815T directional microphone in 1997, a Sony TCM-459V recorder and a Shure Prologue 16AM microphone in 1998, a Sony TC-D5 PROII recorder and a Sennheiser K6M67 microphone in 1999. We could not detect any influence of recording equipment on the parameters measured here. Moreover, all comparisons presented here are made within years (with the same equipment) and results were consistent among the years. Recordings were analysed with AVISOFT SASLAB PRO v. 3.4 using the following settings: sampling frequency 22 050 Hz, 16 bit, time resolution 5.8 ms, bandwidth 111 Hz. To remove background noise, all recordings were filtered using the FIR high-pass filter in AVISOFT (corner frequency: 1.5 kHz) prior to analysis.

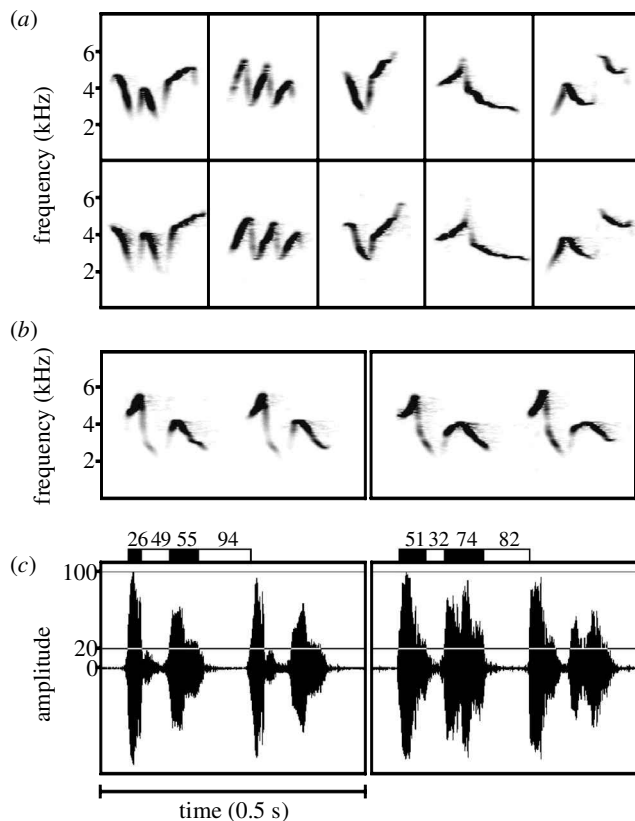


Figure 1. Measuring singing performance. (a) Spectrograms showing examples of five different syllable types (syllables taken out of the strophes) sung by two neighbouring males (top versus bottom). Note the diversity of syllable types and the precision of syllable-type matching. (b) Spectrograms showing the same syllable type sung by two males differing in performance. (c) Oscillograms corresponding with the above spectrograms. They show how to measure 'percentage peak performance', defined as the time when the amplitude exceeds a threshold value of 20% of the maximum value. The bars and numbers on top of the figure indicate the time intervals (in milliseconds) when song above (filled) or below (open) the threshold value is produced. The male on the left reaches peak performance 36% of the time, the male on the right 52%. From these measurements, 'population performance' would be calculated as 44% (average among all males) and 'residual performances' as -8% and +8%, respectively.

The dusky warbler has two distinct types of song: a stereotyped territorial song and a variable advertising song. Our analysis is restricted to the latter because this song is used most during the period when females are fertile (Forstmeier & Balsby 2002). This song type is structured in discrete strophes of 1–2 s, separated by pauses of *ca.* 5 s. A strophe is made up of identically repeated syllables (mean  $\pm$  s.d. =  $5.9 \pm 2.0$  syllables,  $n = 2920$ ), but often includes a switch of syllable type within the strophe (mean 0.5 switches, range 0–4,  $n = 2920$ ).

We recorded 60 strophes from each of 46 males during the egg-laying period. For a subset of 32 males where we had paternity data, we classified all syllables into 205 types. Syllable types were categorized by visual inspection of printed spectrograms (all done by W.F. to ensure consistency of classification; figure 1). We define repertoire size as the number of different syllable types sung by each male (mean: 27, range: 10–40,  $n = 32$ ). For each syllable type, we calculated syllable rate within

strophes as the number of syllables sung per second and the 'percentage peak performance' (abbreviated to 'performance') of a strophe as the proportion of time during which the sound amplitude exceeded 20% of the maximum value for that strophe (figure 1). To measure the length of the time intervals, we used the automatic gate function (signal/silence duration) tool of the AVISOFT software (settings: threshold 20%, delay 1 ms, time constant 1 ms). The threshold of 20% was chosen arbitrarily to avoid interference by background noise. Inspection of the oscillograms indicates that varying the threshold would not greatly affect the outcome of the analyses. For each male, multiple measures of the same syllable type were averaged (using all 60 strophes). For each syllable type, population means of 'syllable rate' and 'percentage peak performance' were calculated by averaging among male means (range: 2.5–9.8 syllables  $s^{-1}$  and 21.4–62.0%;  $n = 205$  syllable types). To compare males, we used residual 'syllable rate' and residual 'percentage peak performance' calculated for each syllable type as the deviation from the population mean and then averaged over all syllable types (range:  $-0.45$  to  $+0.38$  syllables  $s^{-1}$  and  $-4.5\%$  to  $+6.3\%$ ;  $n = 32$  males).

It is likely that our measure of percentage peak performance is affected by the circumstances during recording, such as atmospheric conditions, wind and the position of the bird relative to the microphone. This probably introduces random noise into the data for which we cannot control. Nevertheless, our measure of percentage peak performance was significantly repeatable ( $R = 0.36$ ;  $F_{14,22} = 2.4$ ;  $p = 0.035$ ; analysis based on 15 males recorded two or three times on different days, with at least 30 strophes per recording; following Lessells & Boag (1987)). Although the repeatability is not very high, it is comparable with those obtained for other sexually selected song traits (Forstmeier *et al.* 2003).

In 1998, we measured song rate by counting once per day (between 04.30 and 10.00) the number of songs sung by each of 15 males within a 10 min interval (see Forstmeier & Balsby (2002) for further details).

### (b) Paternity analysis

In 1998 and 1999, we studied paternity using the microsatellite markers Phtr1, Phtr2, Phtr3 (Fridolfsson *et al.* 1997), Pocc5 (Bensch *et al.* 1997), FhU2 (Primmer *et al.* 1996), HrU7 (Primmer *et al.* 1995) and Gf05 (Petren 1998). DNA was extracted from blood samples using the QIAamp Blood Kit. PCRs were done as described by Bensch *et al.* (1997), but using fluorescence-labelled primers, and products were scored on an ABI 377 sequencer. The seven markers allowed us to assign offspring to potential fathers with an average error probability (Kempnaers *et al.* 1999) of  $p = 0.016$  (range: 0.000 02–0.117,  $n = 195$ ; 17 cases with  $p > 0.05$ ). For the data used in figure 3, this average probability was  $p = 0.008$  (range: 0.000 02–0.095,  $n = 49$ ; one case with  $p > 0.04$ ; the latter was assigned to the same father as its nest sib). Phtr3 and Pocc5 are located on the Z-chromosome (Phtr3: 76% of 174 males, but 0% of 135 females were heterozygous; Pocc5: 54% of males and 0% of females heterozygous; birds were sexed genetically as described by Griffiths *et al.* (1998)). For HrU7, the observed rate of heterozygosity (60.2%,  $n = 93$  adults) was significantly lower than expected under Hardy–Weinberg equilibrium (72.6%; binomial exact test:  $p = 0.0033$ ). Moreover, 12 of 195 nestlings (from six nests) were lacking the paternal allele of the (seemingly homozygous) social father ( $n = 3$  males involved). The remaining six microsatellite loci confirmed that these nestlings were

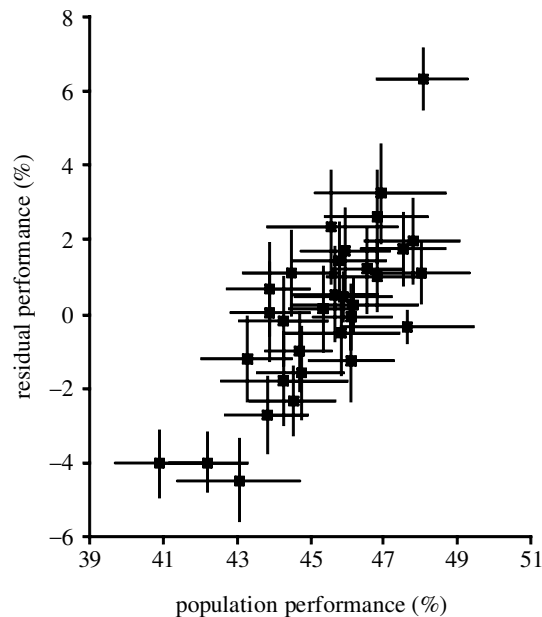


Figure 2. Residual performance (how well males sing) (mean  $\pm$  s.e.m.) versus population performance (what males sing) (mean  $\pm$  s.e.m.) for 32 males. Residual performance reflects how much individual males deviate from the population mean in the percentage peak performance of a given syllable type (averaged over all types), and thus indicates how well males sing relative to others. Population performance reflects which syllable types males include in their repertoire, and thus indicates what males sing. Individuals with high scores selected those syllable types that have an inherently higher performance (population-wide means). The graph shows that males singing high-performance syllable types also sing those types with a higher performance relative to other males. Note that total performance of individual males can be calculated as population performance plus residual performance.

sired by the social fathers (mean error probability:  $p = 0.008$ ; range: 0.0004–0.027). Therefore *ca.* 12% of the individuals must have carried a null-allele for HrU7. No other locus showed such patterns. For 99 nestlings (53 sons, 46 daughters) from 24 nests, DNA samples of the social mother (19 different females) were available, allowing 601 mother–child comparisons ( $99 \times 5$  loci plus  $53 \times 2$  loci). There were no mismatches in this sample, indicating that mutations are rare. After allowing for the presence of null-alleles at HrU7, 88 of 195 nestlings showed at least one mismatch ( $9 \times 1$ ,  $22 \times 2$ ,  $29 \times 3$ ,  $21 \times 4$ ,  $6 \times 5$ ,  $1 \times 6$  mismatches) with the social father. These were considered extra-pair young. We identified the genetic fathers of 74 and 100% of the extra-pair young from 1998 and 1999, respectively. In 1998, we were unable to collect blood samples from 26% of breeding males ( $n = 58$ ).

### 3. RESULTS

Extra-pair paternity was unusually frequent (Petrie & Kempnaers 1998): extra-pair males sired 45% of 195 nestlings and extra-pair fertilizations occurred in 27 out of 46 broods (59%). Overall, we found the genetic fathers of 82% of the extra-pair young. Twenty-seven out of 31 identified fathers (87%) were among the five nearest neighbours.

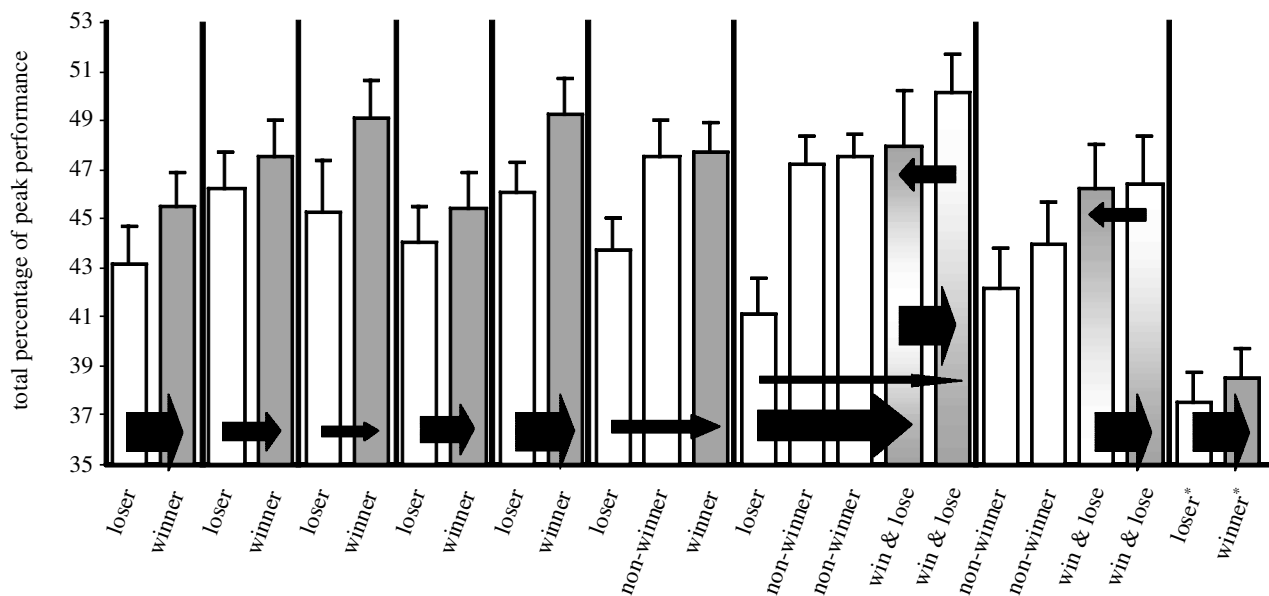


Figure 3. Within-brood comparisons of the singing performance of losers and winners of paternity. Vertical lines separate independent cases ( $n = 9$ ) from one another. We compared males that lost paternity in their own nest ('loser') with the extra-pair males that fathered offspring in these nests ('winner'). Males that did not gain any paternity in those nests despite singing closer to the nest than the 'winner' are referred to as 'non-winner'. In two cases, males exchanged paternity reciprocally (both called 'win&lose'). In all cases, 'winners' produced song of a higher percentage peak performance compared with both 'losers' and 'non-winners'. Arrow width reflects the percentage of nestlings lost (range: 17 to 100%). The last case (\*) does not represent extra-pair paternity, but a case where two males were mated socially to the same female (polyandry) and both males were feeding at the nest. In all other cases, extra-pair fathers never helped at the nest.

To test whether females preferred particular males as extra-pair fathers, we compared song characteristics of the social partner with those of the extra-pair fathers within broods (i.e. pairwise comparisons). Excluding two cases where males reciprocally exchanged paternity (see below), extra-pair males did not differ from cuckolded males in repertoire size (paired  $t_6 = -0.4$ ,  $n = 7$ ,  $p = 0.68$ ), nor in residual syllable rate ( $t_6 = 0.4$ ,  $n = 7$ ,  $p = 0.70$ ), but they produced songs of significantly higher residual performance ( $t_6 = 4.2$ ,  $n = 7$ ,  $p = 0.006$ ). Including the two cases of reciprocal paternity exchange and taking each female as an independent statistical unit, the conclusions did not change (residual performance: paired  $t_{10} = 2.7$ ,  $n = 11$ ,  $p = 0.023$ ). We are cautious about drawing strong conclusions based on such a limited sample size, but the hypothesis that singing performance correlates with extra-pair paternity can be tested further (see below).

If females prefer males that produce song with a higher residual performance (a measure of how well males perform relative to others), the use of different syllable types (i.e. what males sing) should also reflect male quality. This is because different syllable types vary greatly in performance at the population level (average 'population performance' ranging from 21.4% to 62.0%; see § 2). By averaging the 'population performance' of all syllable types included in a male's repertoire, we can quantify the degree to which a given male uses syllable types of inherently high or low performance. The residual performance of 32 males was strongly positively correlated with this measure of syllable use (figure 2; reduced major axis regression:  $r^2 = 0.594$ ,  $n = 32$ ,  $p < 0.0001$ ). This means that males who frequently sing high-performance syllable types also sing those types with a higher performance relative to other

males. We then ranked males according to their total performance (population performance plus residual performance; figure 2). If females prefer higher-ranked males as fathers of their offspring, we predict that extra-pair fathers should rank higher than the male they cuckolded. This is indeed the case (paired  $t_{10} = 2.6$ ,  $n = 11$ ,  $p = 0.026$ ; including the two cases of reciprocal paternity exchange; figure 3). Furthermore, if females can choose freely, one would expect that extra-pair fathers rank higher than all other available males (i.e. males breeding nearer to the female than the extra-pair male). This is also true for all five possible comparisons (paired  $t_4 = 2.7$ ,  $p = 0.055$ ; figure 3). Finally, in the one polyandrous case, the male fathering all of the offspring was ranked higher (figure 3), as would be expected. If females perform extra-pair copulations with males that are of higher quality than their social partner, we would not expect that neighbouring males reciprocally exchange paternity. However, we did find two such cases. This might still be consistent with female choice for singing ability if the two males involved in such exchange both show a high singing performance and are similar to each other. As shown in figure 3, this is indeed the case and the male that fathered most offspring was ranked the highest. Thus, all comparisons that could be made were in the predicted direction (multinomial test,  $n = 9$ ,  $p = 0.00002$ ; figure 3), supporting our above findings on a more limited dataset of pairwise comparisons. When we use each of the two variables 'residual performance' and 'population performance' separately to rank males, in both cases all but one comparison were in the predicted direction (multinomial test,  $n = 9$ ,  $p = 0.0008$ ). Thus, what males sing (syllable types) and how well they sing (residual performance) can explain paternity patterns equally well.

Table 1. Relationships between singing performance and other male and territory characteristics.

trait	<i>n</i>	statistics	<i>p</i>
repertoire size	32	$r = -0.24$	0.18
song rate	15	$r = -0.28$	0.32
mating status <sup>a</sup>	45	$F = -0.18$	0.83
male age <sup>b</sup>	46	$t = -1.59$	0.12
arrival date	37	$r = -0.26$	0.13
pairing date	36	$r = -0.06$	0.72
paternal care	30	$r = 0.27$	0.15
wing length	44	$r = 0.07$	0.65
tail length	43	$r = -0.07$	0.66
tarsus length	46	$r = -0.19$	0.20
body mass	45	$r = -0.07$	0.65
food abundance	46	$r = -0.10$	0.49

<sup>a</sup> Unmated ( $n = 6$ ), monogamous ( $n = 31$ ), polygynous ( $n = 8$ ).

<sup>b</sup> First year ( $n = 17$ ), older ( $n = 29$ ).

Furthermore, we found that nine out of ten males that did not lose paternity in their own nest (including five males shown in figure 3) had a total performance above the overall population mean (binomial test:  $p = 0.011$ ).

The parameter 'total performance' was not correlated with any other male trait investigated (table 1). There was only a non-significant tendency for older males to sing with a higher performance than first-year males. However, it is unlikely that age was a confounding factor because females did not prefer older males for fertilizations. First-year males were more successful in gaining paternity in five out of ten cases where competing males were of a different age class. Singing performance was also unrelated to parameters reflecting territory quality (food abundance) and social pairing success (mating status and pairing date; table 1). To investigate whether females obtained indirect benefits from extra-pair copulations with 'better' singing males (good-genes hypothesis), one would need to compare the survival of within- and extra-pair offspring. Unfortunately, most of the surviving young settle far from the natal area, so these data are hard to obtain. Instead, we analyse the probability of adult male winter survival in relation to singing performance. Adult males show strong site tenacity and often defend territories during late summer, before departure to the winter quarters (Forstmeier 2002). They thus seem to pre-occupy territories in anticipation of the next breeding season (Forstmeier 2002). Of 18 males defending territories in the late summer of 1997 or 1998, eight (44%) returned to exactly the same territory in the following breeding season. Total singing performance was positively correlated with the probability of winter survival (logistic regression:  $\chi^2 = 5.2$ ,  $n = 18$ ,  $p = 0.022$ ) and no other parameter explained the observed patterns of survival.

Given the high rate of extra-pair paternity in this species and the importance of singing performance, one might expect males to follow different mating strategies depending on their singing ability (Kempnaers *et al.* 2001). We observed nine long-lasting fights between males over the possession of a territory where a female had just settled. In all nine cases, the winners of these fights sang with a lower total performance than the losers (paired  $t_8 = -3.8$ ,  $p = 0.005$ ). Eight 'losers' settled in a

neighbouring territory and five of them remained unmated. However, the paternity data for three nests of 'winners' show that those nearby 'losers' had fathered 20%, 100% and 100% of the offspring, respectively.

#### 4. DISCUSSION

Recent studies on physiological limitations during singing indicate that very subtle differences in sound production can have the potential to reflect male quality honestly (Podos 1996, 1997, 2001; Suthers & Goller 1997). In the present study, we show that such fine-scaled variation in sound output correlates with both male success in obtaining extra-pair copulations and with male longevity. This strongly suggests that females might judge male quality by their singing performance and that females may obtain viability genes for their offspring by copulating with good singers (Kempnaers *et al.* 1992, 1997; Haselquist *et al.* 1996).

Moreover, we found that in long-lasting conflicts over the possession of territories low-quality singers won over high-quality singers. We suggest that males with good singing ability might invest less in territorial fights because they can reproduce via extra-pair paternity, whereas low-quality singers may partly compensate for their loss of paternity by increased investment in competition for high-quality territories. This contrasts with the widespread idea that male quality and territory quality should generally be positively correlated. Such a positive correlation was not found in the present study (table 1). As a consequence, we can reject the hypothesis that the female preference for males singing with high performance has evolved because the preference confers direct benefits resulting from high territory quality (see Searcy & Yasukawa 1996). Instead, it seems that females obtain only genetic benefits from their choice of copulation partners.

Our results show that percentage peak performance catches relevant aspects of male performance during sound production. However, it should be obvious that the syllables are not optimized to meet our arbitrary 20% threshold (see § 2). If females only pay attention to this criterion, males would do better by producing pure tones with a constant frequency and amplitude. However, dusky warbler song is characterized by rapid frequency modulations, where a high sound amplitude cannot be maintained (see figure 1). Only experimental work (e.g. using song playback) would allow us to tease apart which criteria females use to assess male singing ability.

This field study may stimulate new directions of research on acoustic communication because it indicates that females may listen to (i) *what* males sing (syllable types differing in 'performance') and (ii) *how* males sing (performance relative to other males). It therefore supports the idea of 'sexy syllables' (Vallet *et al.* 1998) and indicates that physiological limitations may reinforce the honesty of signals (Lambrechts & Dhondt 1986; Lambrechts 1996; Suthers & Goller 1997). Males may sing close to a limit set by respiratory or phonetic motor constraints, which could render signals uncheatable (Lambrechts 1996; Suthers & Goller 1997; Vallet *et al.* 1998). Our findings indicate that the speed at which the warblers' air sacs can be refilled by taking 'minibreaths' (Suthers & Goller 1997) between syllables may be limited

and that it may be difficult to maintain a high amplitude over the course of the most rapid frequency modulations (figure 1). As the vast majority of passerines sing series of notes interrupted by short pauses, our measure of singing performance can be studied in almost any other songbird species.

We thank Vladimir Arkhipov, Hans-Heiner Bergmann, Cornelia van Hulsten, Dries Kuijper, Lars Schmidt-Eisenlohr, Maarten van Steenis and Ingo Weiß for help in the field, Alexander Andreev, Olga Butorina, Iwan Fedutin, Indira Hardani and Olga Mochalova for logistic support, Thorsten Balsby and Raimund Specht for assistance with the song analyses and Topac Ertan for help with laboratory work. Staffan Bensch, Tim Birkhead, Donald Blomqvist, Clive Catchpole, Diego Gil, Franz Goller, Peter McGregor, Kate Lessells, Marc Naguib, Markus Neuhäuser and Tony Wilson provided helpful comments and suggestions. The constructive comments from two anonymous referees helped to improve our paper. We are grateful to Peter Berthold and Karl Eduard Linsenmair for support throughout this study. This project was funded by the Studienstiftung des deutschen Volkes and by the Stifterverband für die Deutsche Wissenschaft (Flora Immerschitt-Stiftung: T 181/08.005/98).

## REFERENCES

- Bensch, S., Price, T. & Kohn, J. 1997 Isolation and characterization of microsatellite loci in a *Phylloscopus* warbler. *Mol. Ecol.* **6**, 91–92.
- Birkhead, T. R. & Møller, A. P. 1992 *Sperm competition in birds*. London: Academic.
- Buchanan, K. L. & Catchpole, C. K. 2000 Song as an indicator of male parental effort in the sedge warbler. *Proc. R. Soc. Lond. B* **267**, 321–326. (DOI 10.1098/rspb.2000.1003.)
- Catchpole, C. K. 1987 Bird song, sexual selection and female choice. *Trends Ecol. Evol.* **2**, 94–97.
- Catchpole, C. K. 2000 Sexual selection and the evolution of song and brain structure in *Acrocephalus* warblers. *Adv. Stud. Behav.* **29**, 45–97.
- Catchpole, C. K. & Slater, P. J. B. 1995 *Bird song*. Cambridge University Press.
- Darwin, C. 1881 *The descent of man, and selection in relation to sex*. London: Murray.
- Davies, N. B. & Lundberg, A. 1984 Food distribution and a variable mating system in the Dunnock, *Prunella modularis*. *J. Anim. Ecol.* **53**, 895–912.
- Forstmeier, W. 2001 Individual reproductive strategies in the dusky warbler (*Phylloscopus fuscatus*): female and male perspectives. PhD thesis, Würzburg University, Germany.
- Forstmeier, W. 2002 Benefits of early arrival at breeding grounds vary between males. *J. Anim. Ecol.* **71**, 1–9.
- Forstmeier, W. & Balsby, T. J. S. 2002 Why mated dusky warblers sing so much: territory guarding and male quality announcement. *Behaviour* **139**, 89–111.
- Forstmeier, W., Leisler, B. & Kempenaers, B. 2001 Bill morphology reflects female independence from male parental help. *Proc. R. Soc. Lond. B* **268**, 1583–1588. (DOI 10.1098/rspb.2001.1692.)
- Forstmeier, W., Leisler, B. & Beier, J. 2003 Singing performance does not improve with age, but reflects male viability in a German great reed warbler population. *Anim. Behav.* (Submitted.)
- Fridolfsson, A.-K., Gyllenstein, U. B. & Jakobsson, S. 1997 Microsatellite markers for paternity testing in the Willow warbler *Phylloscopus trochilus*: high frequency of extra-pair young in an island population. *Hereditas* **126**, 127–132.
- Gil, D. & Gahr, M. 2002 The honesty of bird song: multiple constraints for multiple traits. *Trends Ecol. Evol.* **17**, 133–141.
- Gottlander, K. 1987 Variation in the song rate of male pied flycatchers (*Ficedula hypoleuca*): causes and consequences. *Anim. Behav.* **35**, 1037–1043.
- Griffiths, R., Double, M. C., Orr, K. & Dawson, R. J. G. 1998 A DNA test to sex most birds. *Mol. Ecol.* **7**, 1071–1075.
- Hasselquist, D., Bensch, S. & von Schantz, T. 1996 Correlation between male song repertoire, extra-pair paternity and offspring survival in the great reed warbler. *Nature* **381**, 229–232.
- Kempenaers, B., Verheyen, G. R., Vandenbroeck, M., Burke, T., Vanbroeckhoven, C. & Dhondt, A. A. 1992 Extra-pair paternity results from female preference for high-quality males in the blue tit. *Nature* **357**, 494–496.
- Kempenaers, B., Verheyen, G. R. & Dhondt, A. A. 1997 Extra-pair paternity in the blue tit (*Parus caeruleus*): female choice, male characteristics and offspring performance. *Behav. Ecol.* **8**, 481–492.
- Kempenaers, B., Congdon, B., Boag, P. & Robertson, R. J. 1999 Extra-pair paternity and egg hatchability in tree swallows: evidence for the genetic compatibility hypothesis? *Behav. Ecol.* **10**, 304–311.
- Kempenaers, B., Everding, S., Bishop, C., Boag, P. & Robertson, R. J. 2001 Extra-pair paternity and the reproductive role of male floaters in the tree swallow (*Tachycineta bicolor*). *Behav. Ecol. Sociobiol.* **49**, 251–259.
- Lambrechts, M. M. 1996 Organization of birdsong and constraints on performance. In *Ecology and evolution of acoustic communication in birds* (ed. D. E. Kroodsma & E. H. Miller), pp. 305–320. Ithaca, NY: Cornell University Press.
- Lambrechts, M. & Dhondt, A. A. 1986 Male quality, reproduction, and survival in the great tit (*Parus major*). *Behav. Ecol. Sociobiol.* **19**, 57–63.
- Lessells, C. M. & Boag, P. T. 1987 Unrepeatable repeatabilities: a common mistake. *Auk* **104**, 116–121.
- Møller, A. P. 1991 Why mated songbirds sing so much: mate guarding and male announcement of mate fertility status. *Am. Nat.* **138**, 994–1014.
- 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. (DOI 10.1098/rspb.2000.1300.)
- Petren, K. 1998 Microsatellite primers from *Geospiza fortis* and cross-species amplification in Darwin's finches. *Mol. Ecol.* **7**, 1782–1784.
- Petrie, M. & Kempenaers, B. 1998 Extra-pair paternity in birds: explaining variation between species and populations. *Trends Ecol. Evol.* **13**, 52–58.
- 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.
- Primmer, C. R., Møller, A. P. & Ellegren, H. 1995 Resolving genetic relationship with microsatellite markers: a parentage testing system for the swallow *Hirundo rustica*. *Mol. Ecol.* **4**, 493–498.
- Primmer, C. R., Møller, A. P. & Ellegren, H. 1996 A wide-range survey of cross-species microsatellite amplification in birds. *Mol. Ecol.* **5**, 365–378.
- Reid, M. L. 1987 Costliness and reliability in the singing vigour of Ipswich sparrows. *Anim. Behav.* **35**, 1735–1743.
- Searcy, W. A. & Yasukawa, K. 1996 Song and female choice. In *Ecology and evolution of acoustic communication in birds* (ed.

- D. E. Kroodsma & E. H. Miller), pp. 454–473. Ithaca, NY: Cornell University Press.
- Suthers, R. A. & Goller, F. 1997 Motor correlates of vocal diversity in songbirds. In *Current ornithology*, vol. 14 (ed. V. Nolan Jr, E. D. Ketterson & C. F. Thompson), pp. 235–288. New York: Plenum.
- Vallet, E., Beme, I. R. & Kreutzer, M. 1998 Two-note syllables in canary songs elicit high levels of sexual displays. *Anim. Behav.* **55**, 291–297.
- von Schantz, T., Bensch, S., Grahn, M., Hasselquist, D. & Wittzell, H. 1999 Good genes, oxidative stress and condition-dependent sexual signals. *Proc. R. Soc. Lond. B* **266**, 1–12. (DOI 10.1098/rspb.1999.0597.)

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.