



Song repertoire size correlates with measures of body size in Eurasian blackbirds

Nana Hesler^{a,*}, Roger Mundry^b, Thomas Sacher^c, Timothy Coppack^{c,d},
Franz Bairlein^c and Torben Dabelsteen^a

^a Institute of Biology, University of Copenhagen, Universitetsparken 15,
2100 Copenhagen Ø, Denmark

^b Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

^c Institute of Avian Research 'Vogelwarte Helgoland', Wilhelmshaven, Germany

^d Institute of Applied Ecology, Department of Ornithology, Neu Broderstorf, Germany

*Corresponding author's e-mail address: nana.hesler@wiff.dk

Accepted 22 February 2012

Abstract

In most oscine bird species males possess a repertoire of different song patterns. The size of these repertoires is assumed to serve as an honest signal of male quality. The Eurasian blackbird's (*Turdus merula*) song contains a large repertoire of different element types with a flexible song organisation. Here we investigated whether repertoire size in Eurasian blackbirds correlates with measures of body size, namely length of wing, 8th primary, beak and tarsus. So far, very few studies have investigated species with large repertoires and a flexible song organisation in this context. We found positive correlations, meaning that larger males had larger repertoires. Larger males may have better fighting abilities and, thus, advantages in territorial defence. Larger structural body size may also reflect better conditions during early development. Therefore, under the assumption that body size is linked to overall male quality, our results are in accordance with the hypothesis that repertoire size represents an honest signal in Eurasian blackbirds that has evolved in response to sexual selection.

Keywords

bird song, body measures, male quality, repertoire size, sexual selection, honest signal, *Turdus merula*, Eurasian blackbird.

1. Introduction

The song of oscine birds is known to have the dual function of mate attraction and territory announcement. Specific song features are assumed to

advertise different aspects of male quality, both in inter- and intra-sexual contexts (Catchpole & Slater, 2008). Females have been shown to prefer, e.g., high song rates (Alatalo et al., 1990a) and song output (Poesel et al., 2001), specific song or syllable types (Leitner et al., 2006), local dialects (Baker et al., 1987) or, in species where males possess a repertoire of different element or song types, large repertoire sizes (Searcy, 1992). While some of the features are more dependent on the current condition (e.g., song duration) or arousal (e.g., song rate) of the singer, repertoire size seems to be more stable and can, therefore, be expected to be linked to more long-term stable aspects of male quality (e.g., genetic quality, or features depending on conditions during early development; Borowiec et al., 2007). Some species, 'open ended learners', have the ability to enlarge their repertoire after song crystallisation as adult birds since the memorisation of different song patterns is not restricted to one sensitive period early in life. However, even in these species the largest increase was often found between the first and the second breeding season (e.g., starling (*Sturnus vulgaris*), Eens et al., 1992; Mountjoy & Lemon, 1995), while repertoire size in fully adult birds seems stable (Garamszegi et al., 2007; Kipper & Kiefer, 2010).

Studies in several species found correlations between repertoire size and different aspects of male quality, e.g., between repertoire size and body condition (starling, Mountjoy & Lemon, 1996; sedge warbler (*Acrocephalus schoenobaenus*), Buchanan et al., 1999; song sparrow (*Melospiza melodia*), Reid et al., 2005; Pfaff et al., 2007; common nightingale (*Luscinia megarhynchos*), Kipper et al., 2006), between repertoire size and body size (whitethroat (*Sylvia communis*) Balsby, 2000; common nightingale, Kipper et al., 2006), between repertoire size and age (e.g., common nightingale, Kiefer et al., 2006, 2009), but also between repertoire size and social status (starling, Spencer et al., 2004) or learning ability (zebra finch (*Taeniopygia guttata*), Boogert et al., 2008). According to these results, song repertoire size seems to be an indicator for different correlates of a male's general quality, and by choosing mates with large repertoires, females might obtain larger, healthier, or more experienced males.

How to measure a male's quality is not always obvious as the ultimate measure, his lifetime reproductive success, is very difficult to measure in field studies. Several studies found links between different body measures and quality. In many mammal or bird species, larger and/or heavier individuals had better survival or reproduction rates (Andersson, 1994; Festa-Bianchet et al., 1998). In passerines, key biometric measures such as tarsus,

wing and beak length can predict survival prospects or reproductive success and may, therefore, be related to male quality (great tit (*Parus major*) Garnett, 1981; blue tit (*Parus caeruleus*), Kempenaers et al., 1992, 1997; red-winged blackbird (*Agelaius phoeniceus*), Weatherhead & Boag, 1995). In great tits and willow tits (*Parus montanus*), tarsus length does not change after day 14 and is partly heritable, partly depending on nestling condition (Garnett, 1981; Thessing & Ekman, 1994). Therefore, adult tarsus length contains information about the condition early in life, which may correlate with other aspects of male quality.

In Eurasian blackbirds (*Turdus merula*, hereafter blackbird), males' culmen length (i.e., beak length) was positively associated with the intensity of the orange beak colouration (Bright et al., 2004). Male blackbird beak colour is based on carotenoids playing a role in immune response (Faivre et al., 2003a) and has been shown to correlate with different aspects of male quality (body weight, Faivre et al., 2001; cellular immunocompetence, Faivre et al., 2003b; number of fledglings and better parental care, Pr eault et al., 2005). Therefore, this association could hint at a link between beak length and male quality. Furthermore, beak, wing and tarsus length are thought to be representative of general body size (Rising & Somers, 1989). For male blackbirds, larger body size may offer direct advantages as they often engage in territorial contests, including escalated fighting (Snow, 1958; own observations).

How song repertoire size can be measured depends on song structure. While some species produce a continuous sequence of elements (the shortest song unit), other species' song is discontinuous and consists of distinct songs ('strophes') that are separated by silent inter-song intervals and comprise several elements. A set of song types (in discontinuous song, e.g., common nightingale) or element types (in continuous song, e.g., skylark (*Alauda arvensis*)) builds up an individual's repertoire. The blackbird's song is discontinuous (Dabelsteen, 1984) but has no fixed song types (Todt, 1968a,b) as 'motifs', fixed sequences of elements (Dabelsteen, 1984; Rasmussen & Dabelsteen, 2002), are combined flexibly within songs. Each blackbird song consists of two acoustically distinct parts, starting with the far ranging, low frequency whistle part and ending with the twitter part. The latter has a lower sound pressure level and broader frequency range and, therefore, a much shorter transmission range (Dabelsteen, 1984; Dabelsteen et al., 1993). Both whistle and twitter parts contain large repertoires of different element types,

but little is known about the function of their repertoire size (but see Dabelsteen, 2005; Hesler et al., 2011; Klump et al., 2012).

Here we investigate the relationship between element repertoire size and measures of body size as a correlate of quality in male blackbirds. So far, only very few studies in this context have been conducted on species with a large repertoire and flexible song organisation, as the blackbird. Both repertoire size (Rasmussen & Dabelsteen, 2002; Hesler, 2010; Hesler et al., 2011, 2012) and body size (Glutz von Blotzheim & Bauer, 1985, p. 845ff; Stephan, 1999, p. 25) vary to a large extent among individuals in blackbirds. According to the evidence and hypotheses from the literature, we expect that larger birds have larger repertoires, which would be in accordance with the hypothesis that repertoire size serves as an honest signal of male quality. By distinguishing between whistle and twitter repertoires, we also aim at investigating potential differences in the information content of these distinct parts of the blackbird song.

2. Methods

2.1. Study population and body measurements

We conducted this study on the island of Helgoland, Germany, in a blackbird population consisting of circa 80 breeding pairs at the time of the study. The population has been studied since 2003 (Sacher et al., 2006) with ringing permission given by the Institute of Avian Research 'Vogelwarte Helgoland'. Birds were caught either as nestlings or as older birds with spring loaded net traps and marked with a unique combination of three colour and one aluminum ring. The handling of nestlings took less than 5 min per bird and about 15 min per nest. No nest was abandoned by the parents after ringing the nestlings. While catching older birds, we observed loaded traps continuously and released birds immediately after they were caught. Handling (ringing and/or measuring) took at most 5 min. In this study, we used only fully adult males (at least in their 3rd calendar year/2nd breeding season, determined by a moult boulder in the greater coverts; Svensson, 1992; from now on referred to as 'adult'). We measured wing length (accuracy ± 0.5 mm), length of the 8th primary (counted from the body towards the tip of the wing (± 0.5 mm), length of the tibiotarsus (tarsus, ± 0.1 mm), and beak length from the tip of the beak to the start of the skull (± 0.1 mm; Svensson, 1992; Bairlein, 1994).

2.2. Song recordings and repertoire analysis

We recorded song of territorial males during dawn chorus in the breeding seasons 2006–2008, between 8 April and 4 June, starting circa 1 h before sunrise (between 04:00–06:20 h CEST). Repertoire size (song-standardised combined repertoire, see below) did not correlate with recording date (Spearman's $\rho = 0.400$, $N = 9$, $p = 0.291$). For this study, we used a sample of nine males; eight of them were paired when we recorded them; for one the breeding state was unknown. We recorded each male during one morning to record one full dawn chorus, i.e., during circa 30 min (20–50 min recording time).

We recorded song using a directional microphone (Sennheiser K6/ME67) and a Marantz PMD670 solid state recorder with a sample rate of 22 050 Hz (2006) or 44 100 Hz (2007, 2008). For sound analyses we used Avisoft-SASLab Pro 4.4–4.52. We down-sampled recordings in 44 100–22 050 Hz and filtered all recordings (Butterworth, high pass 1.5 kHz, low pass 10 kHz). We defined songs as a sequence of elements separated by inter-song pauses of ≥ 0.5 s. For repertoire analyses, we classified all elements in the used recordings by visual inspection according to their overall appearance on printed spectrograms, i.e., frequency and temporal characteristics (spectrogram settings: FFT 512, window Hamming, bandwidth 56 Hz, resolution 43 Hz, frame size 100%, overlap 75%, resolution 5.8 ms; graphical scale on prints: 1 s = 36 mm, 1 kHz = 6 mm). We defined an element as a single continuous trace in time without visible sound free pauses, i.e., pauses longer than circa 0.015 s (Figure 1). We classified elements as either whistle (low frequency, almost pure toned, high sound pressure level) or twitter (broader frequency range, more energy on harmonics, lower sound pressure level) elements (Figure 1). For a more detailed description of element classification, see Hesler (2010) or Hesler et al. (2011).

We determined repertoire sizes by counting the number of different element types in 100 consecutive songs ('song-standardised combined repertoire'). Because there was a large variation in the number of elements per song, we also measured repertoire size as number of different element types in the first 1000 analysed elements ('element-standardised combined repertoire'). Whistle element repertoires were measured separately as number of different whistle element types found in the first 670 analysed whistle elements ('element-standardised whistle repertoire') and twitter element repertoires as number of different twitter elements types in the first 270 analysed

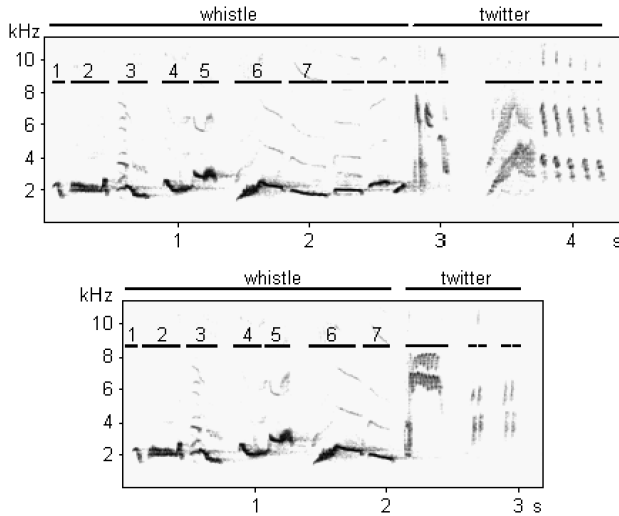


Figure 1. Spectrograms of two blackbird songs from the same individual with indication of whistle/twitter parts and of single elements by black bars. Element types which occur in both songs are labelled with numbers.

twitter elements ('element-standardised twitter repertoire'). The analysed numbers of songs and elements, respectively, were the maximum numbers available in the shortest recordings. Especially the twitter part can be very short or even missing (Dabelsteen, 1984). Therefore, we found only 270 twitter elements in one recording, even though it contained 100 songs.

All repertoire analyses were carried out by the same person (N.H.), but to check for observer objectivity we compared N.H.'s result of one recording containing 102 songs with the result of another person, B. Klump, after giving her instructions. Her result of a combined repertoire size of 253 element types differed by 0.02% from N.H.'s result of 257 element types.

2.3. Sample size of songs

A sample of 100 songs usually captures the vast majority of a blackbird's repertoire although additional song patterns may occur even after more than 300 songs (Hesler, 2010). For this study we were less interested in absolute repertoire sizes, but in reliably assessing inter-individual differences which can be large and stable even before the cumulative plateau of element type repertoires is reached (Mountjoy & Lemon, 1996; Gil & Slater, 2000). After visual checks of saturation curves of ten males for which we had analysed between 250 and 385 songs, i.e., until curves levelled out, we concluded that

analysing 100–150 songs should be sufficient to detect such inter-individual differences (unpublished data). Since the visual determination of repertoire sizes based on element types is very time consuming, we estimated the minimum number of songs which should be analysed. For this, we analysed in another study (Hesler, 2010) 140 songs of six adult blackbirds. For each bird, we then correlated the repertoire size found in all 140 songs to that found in 139 songs, in 138 songs, and so on, to see where correlations would drop considerably. We found that already after less than 40 analysed songs, repertoire sizes correlated highly and significantly with those found after 140 songs (resulting in correlation coefficients of over $\rho = 0.95$ and p values < 0.05 , exact Spearman rank correlation). Therefore, we reasoned that 100 songs should be sufficient to reliably determine inter-individual differences in repertoire sizes.

2.4. Variables and statistical analysis

Since the sample of birds that had been recorded and measured in the same year was too small, we had to use body measurement and song data from different years (body measures: years 2004–2007; song recordings: years 2006–2008). Song recordings were taken in the same year or later than body measurements (on average 1.44 years later; min 0, max 3 years). Contrary to body weight, the here used body measures are assumed to be stable over time, especially tarsus and beak and particularly in adult birds (Garnett, 1981; Thessing & Ekman, 1994; Gosler et al., 1998). In a Polish blackbird population, Wysocki & Kiriaka (2007) found that neither wing nor metatarsus length changed significantly between years in adult birds. Between the first and second breeding season however wing length increased (Wysocki & Kiriaka, 2007). Therefore, we only used birds which were measured and recorded at least in their second breeding season, assuming that the measured body parts have not increased between measuring and recording date.

For two birds we were missing one value each: for one bird wing length, and for another bird 8th primary length, due to worn out feather tips. Since both measurements correlated strongly in the remaining seven birds (exact Spearman correlation: $\rho = 0.964$, $N = 7$, $p = 0.002$), we estimated the two missing values with a reduced major axis regression (Quinn & Keough, 2002).

In order to reduce the number of body size variables (length of wing, 8th primary, beak, tarsus), we ran a Principal Components Analysis (PCA).

Since tarsus length correlated only weakly with the other variables (largest absolute correlation 0.536), we excluded it from the PCA. The PCA extracted one principal component (PC) with an Eigenvalue > 1 , explaining 86.1% of the total variance. Loadings of the variables on it were 0.96 (wing length), 0.97 (length of 8th primary) and 0.84 (beak length), meaning that these measurements were larger for birds with higher PC scores. The PCA was justified as suggested by large correlations between these variables, and by the Kaiser–Meyer–Olkin measure of sampling adequacy (0.65) as well as Bartlett’s test of sphericity ($\chi^2 = 25.5$, $df = 3$, $p < 0.001$; McGregor, 1992). In the following analyses we used the factor score derived from the PCA as well as tarsus length as measures of body size. Note that if tarsus length was included into the PCA it revealed two PCs onto one of which only tarsus length loaded strongly. Hence, our approach led to the same number of variables used and a direct inclusion of tarsus length rather than a PC correlating with it.

We checked for correlations between the PC score and tarsus length and the four different repertoire size measurements (song-standardised combined repertoire, element-standardised combined repertoire, element-standardised whistle repertoire, element-standardised twitter repertoire) using Spearman rank correlations and exact permutation tests.

Testing the same null hypothesis (i.e., no correlation between body measures and repertoire sizes) several times required an error level adjustment to correct for multiple testing. We achieved this using Fisher’s omnibus test with a p value determined based on a permutation procedure (Adams & Anthony, 1996; Manly, 1997). Fisher’s omnibus test combines a number of p values into a single χ^2 -distributed variable with degrees of freedom equalling twice the number of p values (Haccou & Meelis, 1994). If Fisher’s omnibus test reveals significance this indicates that the combined p values are on average smaller than expected by chance and that at least some of the individual p values are significant. We, therefore, considered inspection of the individual p values as a post-hoc procedure which does not require a correction because the overall null-hypothesis was already rejected by Fisher’s omnibus test as a global test. However, this procedure requires independent p values, which cannot be assumed in case of our four different repertoire measurements and two body measures from the same individual. Hence, the

theoretical χ^2 distribution is not valid (Potter & Griffiths, 2006). In order to achieve a valid test nevertheless, we permuted tarsus length and PC score simultaneously across subjects, calculated the eight correlations again, and finally combined the p values derived into a χ^2 value like in Fisher's omnibus test. We ran 10 000 permutations, into which we included the original data as one permutation, and determined the p value assigned to the χ^2 of the original data as the proportion of permutations revealing a χ^2 at least as large as that of the original data. By doing so we obtained a single overarching test, revealing the statistical significance for the relation between body size and repertoire size.

PCAs were calculated using SPSS 15.0.0 for Windows. Spearman rank correlations, permutation tests and fisher omnibus test were calculated in R (version 2.9.1; R Development Core Team, 2008) using self-written scripts. All non-parametric tests were exact (Mundry & Fischer, 1998) and two-tailed.

3. Results

Overall, there was a clear relation between body measures and repertoire sizes (Fisher's omnibus test combined with permutation: $\chi^2 = 41.8$, 10 000 permutations, $p = 0.011$, Figure 2). In detail, birds with longer tarsi had significantly larger combined repertoires, both song-standardised and element-standardised (Figure 2). Birds with higher PC scores (meaning that they had longer beaks, wings and 8th primaries) showed a tendency in the same direction, i.e., they tended to have larger element- and song-standardised combined repertoires (Figure 2). Considering whistle and twitter repertoires separately, there were no significant correlations, but a clear tendency for birds with higher PC scores to have larger whistle repertoires while birds with longer tarsi tended to have larger twitter repertoires (Figure 2).

One individual had very low values for both repertoire size and tarsus length. However, as we used non-parametric tests, our results should not depend too much on this one individual. Nevertheless, to estimate whether or to what extent our results are influenced by any particular individual we excluded each subject once and repeated the analysis. This revealed no drastically different correlation coefficients and p values (Table 1), and, hence, we conclude that our results are not too strongly influenced by any one individual.

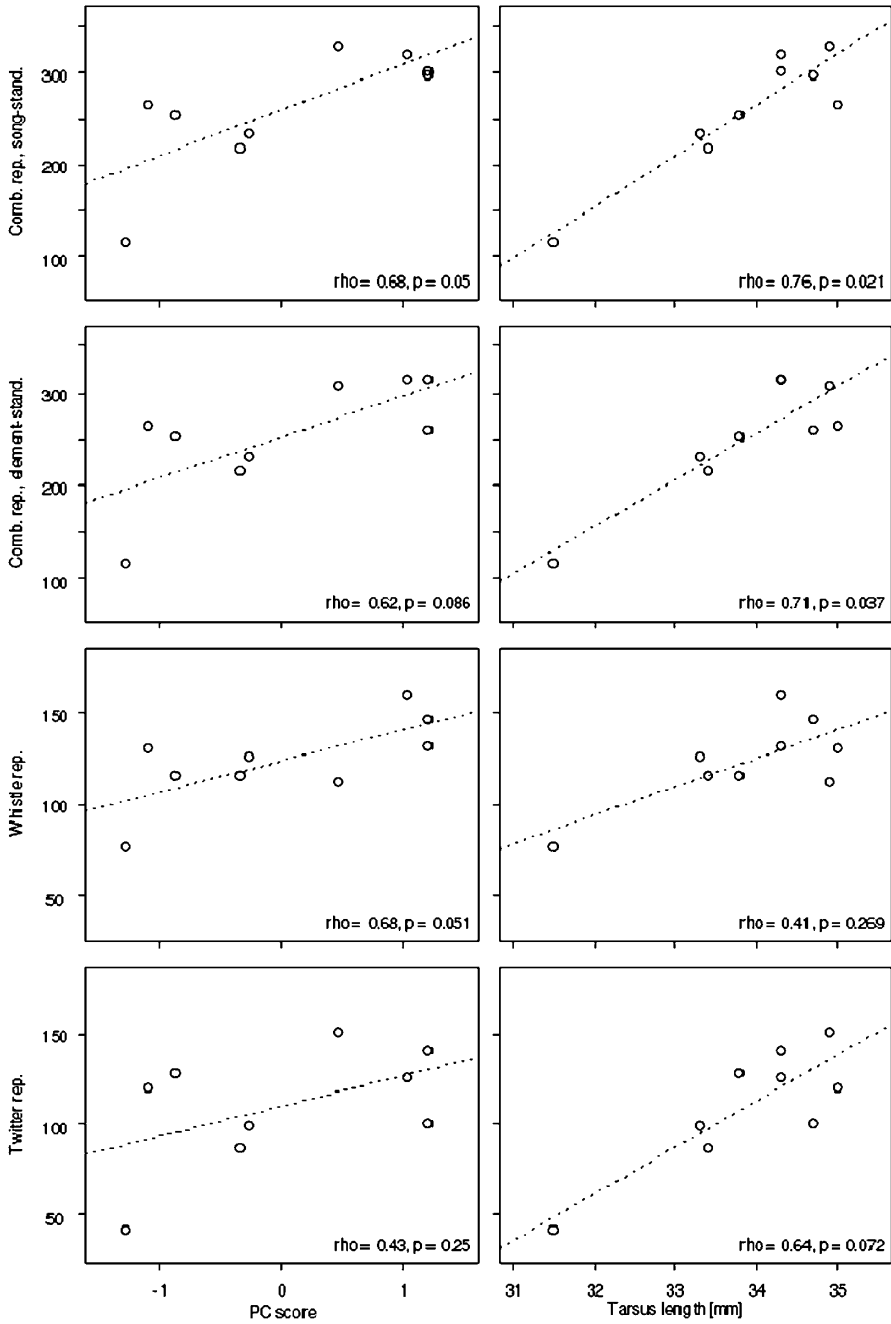


Table 1.
Results (ρ and p values) of the analysis ($N = 9$).

Body measure	Repertoire	ρ (range)	p (range)
PC score	comb., song-stand.	0.68 (0.67–0.69)	0.05 (0.07–0.08)
	comb., element-stand.	0.62 (0.55–0.76)	0.09 (0.04–0.17)
	Whistle	0.68 (0.59–0.78)	0.05 (0.03–0.13)
	Twitter	0.43 (0.36–0.64)	0.25 (0.10–0.39)
Tarsus length	comb., song-stand.	0.76 (0.66–0.90)	0.02 (0.01–0.08)
	comb., element-stand.	0.71 (0.59–0.81)	0.04 (0.02–0.14)
	Whistle	0.41 (0.36–0.73)	0.27 (0.04–0.38)
	Twitter	0.64 (0.50–0.79)	0.07 (0.02–0.21)

Values in parentheses depict the range of the results when excluding each subject once.

4. Discussion

We found that males with longer tarsi, beaks, wings and 8th primaries had larger element repertoires, even though the correlation between the PC score (representing beak, wing and 8th primary length), and the element repertoire size just failed to reach significance. Assuming that these measures represent a male's overall body size (Rising & Somers, 1989), it should be possible to obtain information about a male's body size from hearing it singing, which may be of functional significance both in intra- and in inter-sexual contexts. Larger males may have advantages in territorial contests, which often include extensive fights and, thus, obtain better territories. Larger body size may also reflect better conditions during early development (Horak, 1994; Nowicki et al., 1998). Similar results have been found in whitethroats (correlations between wing length and repertoire size, but not tarsus, Balsby, 2000) and nightingales (correlations between wing length and repertoire size, Kipper et al., 2006); both of them are species with very large repertoires. However, in other species, no such correlations were found (song sparrow, Searcy et al., 1985; great reed warbler (*Acrocephalus arundinaceus*), Catchpole, 1986).

Figure 2. Repertoire sizes and body measures. (Left) PC score: principal component representing length of wing, 8th primary and beak, (right) tarsus length. The different measures of repertoire size are: Combined rep. song-stand.: combined repertoire, song-standardised; Combined rep., element-stand.: combined repertoire, element-standardised; Whistle rep.: whistle repertoire, element-standardised; Twitter rep.: twitter repertoire, element-standardised. Each dot represents one subject. Results of exact Spearman rank correlations are indicated ($N = 9$). Trend lines are only included for visualisation.

Looking separately at the repertoire sizes in the two different categories of blackbird song, whistle and twitter elements, we did not find significant correlations with body measures, but clear trends in the same direction as for the overall repertoire. That these correlations were not significant may be due to the fact that the number of considered whistle and twitter elements was rather small compared to the overall number of combined elements. Even though we analysed the same number of songs for all birds, as one bird sang quite short twitter parts containing in total only 270 twitter elements, we could use only 270 twitter elements in all birds to assess twitter repertoires.

The found correlations between repertoire and body size are in accordance with the hypothesis that repertoire size signals male quality. In addition, blackbirds have not only a large repertoire compared to other species, but also a complex singing style with immediate variety, i.e., only few repetitions, both within songs (see Figure 1) and between songs (Hesler, 2010, article 5). As stated by Byers & Kroodsma (2009), if large repertoires advertise male quality, males should sing in a way that easily reveals a large repertoire. However, in many species repertoires are delivered with eventual variety, i.e., with many repetitions, restricting a fast assessment of large repertoires (Byers & Kroodsma, 2009). The singing style of blackbirds on the other hand reveals large repertoires efficiently. This allows, e.g., a female to assess repertoire size relatively fast and would, thus, make sense in the light of the hypothesis that large repertoires advertise male quality.

Another explanation for the existence of song repertoires is offered by the song sharing hypothesis stating that not repertoire size, but the degree of song sharing with neighbours is the selected trait (Beecher & Brenowitz, 2005; Lapierre et al., 2011). Shared songs can signal belonging to a community and allow for song matching which in turn allows addressing specific individuals and has been shown to be an aggressive signal (Beecher & Brenowitz, 2005). While repertoires in general are required, larger repertoires are not necessarily advantageous for song matching (Beecher & Brenowitz, 2005). Blackbirds indeed share parts of their repertoires (Rasmussen & Dabelsteen, 2002), and song matching interactions are common (Todt, 1981), but it is not known whether the degree of repertoire sharing relates to male quality and/or repertoire size. Also, the degree of sharing may differ between whistle and twitter repertoire. Whether blackbirds with larger repertoires possess more shared elements needs future investigations.

Since it is not obvious how larger repertoires may cause costs, the causes for the link between repertoire size and male quality are not clear (Oberweger & Goller, 2001; Ward et al., 2004). The ‘nutritional stress hypothesis’ (Nowicki et al., 1998), or, broader, ‘developmental stress hypothesis’ (Buchanan et al., 2003) suggests that this link may be found in the early development. The high vocal centre (HVC) and other song nuclei crucially involved in song learning develop during an early post-hatching phase in which birds are likely to experience nutritional and other kinds of stress (Nowicki et al., 1998, 2002). Recent studies showed that birds that experienced nutritional and other kinds of stress in fledgling age indeed had smaller adult song nuclei (Nowicki et al., 2002; Buchanan et al., 2004; Spencer et al., 2005) and developed smaller adult repertoires (Nowicki et al., 2002; Spencer et al., 2004, but see Brumm et al., 2009). Thus, adult repertoire size may inform about an individual’s ability to cope with developmental stress and, thus, about aspects of his genotypic quality (Nowicki et al., 1998). Since individuals that coped better with developmental stress may also develop better general phenotypic condition, adult repertoire size may also inform about phenotypic condition (Nowicki et al., 1998). At the same time, nutritional stress early in life can result in smaller adult body sizes (Nowicki et al., 1998) and explain the found correlations between body size and repertoire size.

Hence, according to the developmental stress hypothesis, females choosing larger males with larger repertoires could obtain indirect benefits by choosing males with better genes. There are most likely also direct benefits from choosing larger males as mates, as larger males are probably stronger in fights and may, therefore, obtain better territories (Andersson, 1994). Therefore, information about a male’s body size would be useful for male competitors and, thus, in intra-sexual contexts, too. However, in a former study we found no differential responses by territorial blackbird males towards playback of conspecific song with different repertoire sizes (Hesler et al., 2011). The tested context was a territory intrusion, i.e., we broadcast song in the test birds’ territories, and possibly the overall strong but not differential responses were due to a ceiling effect. In other intra-sexual contexts repertoire size may nevertheless be meaningful. For example, males may assess their neighbours’ repertoires during dawn chorus and base future decisions on repertoire sizes.

Another possible explanation for the found correlations might be that both body measures and repertoire size increase with age and, therefore, correlate.

Since we did not know the exact age of our study subjects, we could not test for this possibility. However, in a different blackbird population, both wing and metatarsus length remained stable between two successive years in adult birds while wing length did increase between the first and second breeding season (Wysocki & Kiriaka, 2007). Since we only used adult birds, it therefore seems unlikely that the observed differences in body measures are due to continued growth. Also in other species, particularly tarsus length is reported to be fixed already at the age of fledging (Garnett, 1981; Alatalo et al., 1990b), and beak length to be stable over time too, while wing length may be more variable (Rising & Somers, 1989).

As our body measures were taken either in the same year or later than the song recordings, the song repertoires cannot have increased after the body measures were taken. Whether repertoire size increases in adult blackbirds is unknown and needs to be investigated in further studies. In another study, we found that adult blackbirds tended to have larger repertoires than yearlings (Hesler et al., 2012). In other species, most differences in repertoire size have been found between yearlings and adult birds (O'Loughlen & Rothstein, 1993; Gil et al., 2001; Kipper et al., 2004; Kiefer et al., 2006, 2009; Garamszegi et al., 2007; Kipper & Kiefer, 2010). Age-related differences in repertoire size could either be due to different mortality rates of males with different repertoire sizes, or to actual changes in the repertoire size of the individuals. As adult blackbirds have been observed to incorporate new motifs (Messmer & Messmer, 1956; Thielcke, 1961), they may have the option to increase their repertoires also later in life. However, even the addition of new element types does not necessarily lead to a repertoire increase: in some other species a large turnover of elements between years has been observed, meaning that new elements are included, while old elements are dropped (Nicholson et al., 2007). In a very small sample of blackbirds, we also observed a substantial repertoire turnover between males' first and second breeding season with comparably small increases in repertoire size (Hesler et al., 2012). Considering the small differences in repertoire size which we found between yearling and adult males in a cross-sectional comparison (Hesler et al., 2012), a significant increase in repertoire size in adult males seems not very likely, but possible, requiring longitudinal studies on adult birds' repertoires.

That our study was carried out in a small island population raises the question whether our results are only valid for this population or the outcome of

universal patterns in the whole species. Island populations have been found to differ from mainland populations in different aspects (Berry, 1983). In several passerine species, individuals of island populations were found to have longer beaks and higher body weight than those of the respective mainland populations (e.g., Grant, 1965; Clegg & Owens, 2002; Boyer & Jetz, 2010). Proximately, founder effects can lead to less genotypic and phenotypic variance and to a rapid prevalence of certain traits in island populations. Also culturally transmitted traits, such as the song repertoire, can be affected by founder effects. Song complexity seems to be frequently lower on islands (Thielcke, 1973; Baker et al., 2006; Catchpole & Slater, 2008). As young birds acquire their song patterns from tutors that they hear singing, the repertoire of an isolated island population may be restricted to the repertoire of the founders even after the population increased in number of individuals. This may result in smaller individual repertoires too.

Our study population is still young (dating back to 1983) and seems to be genetically isolated (Sacher, 2010), thus possibly subjected to founder effects. The inter-individual trait variation we found was however at least sufficiently high to find significant correlations between repertoire size and body measures even in our small sample. Morphologically, individuals from our island population differed from migrating, mostly Scandinavian individuals caught on the island by having longer wings, beaks and partly tarsi (Sacher, 2010). We have, however, no comparison to ecologically more similar populations (i.e., from the close mainland and non-migratory). Concerning repertoire sizes, there are no studies on blackbird song allowing comparisons. Thus, body measures as well as repertoire sizes may differ between our study population and other blackbird populations. For the interpretation of our results, the more important question is whether the direction of selective pressures is different. If larger body size is advantageous for male blackbirds on Helgoland, larger repertoires may signal higher male quality. In blue tits, Doutrelant et al. (2000) found stronger positive correlations between tarsus length and repertoire size in two island populations than in one mainland population. The authors suggest that this was due to different levels of inter-specific competition for acoustic space, as the relative great tit densities were lower on the island and blue tit repertoire sizes increased with decreasing great tit densities. It would be very interesting to investigate, whether the correlations between repertoire size and body size that we found in our study population also exist in other blackbird populations. However, the fact that

our finding is in line with other investigations of correlations between repertoire size and morphological measures in other species (Lampe & Espmark, 1994; Mountjoy & Lemon, 1996; Balsby, 2000; Kipper et al., 2006) may suggest that our findings are not specific to the island population we studied, but a more general trend.

In conclusion, our results are consistent with the hypothesis that repertoire size serves as an honest signal of male quality in the blackbird. Further research is needed to investigate the assumed relationship between body size and male fitness and whether other blackbird populations show the same correlations between repertoire size and body size.

Acknowledgements

We thank O. Hüppop with the team of the Inselstation Helgoland and C. Horn for assistance during the field work. We also thank L. Holman and J. Hesler for language improvement and two anonymous reviewers for their helpful comments. The study was funded by a combined PhD fellowship from NaFöG (Promotionsstipendium nach dem Nachwuchsförderungsgesetz des Landes Berlin, Germany), SNAK (Graduate School on Sense organs, Nerve systems, Behaviour and Communication, funded by the Danish Research Agency) and the University of Copenhagen to N.H. and a framework grant from the Danish Research Council for Nature and Universe (No. 272-07-0477) to T.D. This study meets the guidelines of the Association for the Study of Animal Behaviour and complies with laws in Germany and Denmark.

References

- Adams, D.C. & Anthony, C.D. (1996). Using randomization techniques to analyse behavioural data. — *Anim. Behav.* 51: 733-738.
- Alatalo, R.V., Glynn, C. & Lundberg, A. (1990a). Singing rate and female attraction in the pied flycatcher—an experiment. — *Anim. Behav.* 39: 601-603.
- Alatalo, R.V., Gustafsson, L. & Lundberg, A. (1990b). Phenotypic selection on heritable size traits—environmental variance and genetic response. — *Am. Nat.* 135: 464-471.
- Andersson, M. (1994). *Sexual selection*. — Princeton University Press, Princeton, NJ.
- Bairlein, F. (1994). *Manual of field methods*. — European–African Songbird Migration Network, Wilhelmshaven.
- Baker, M.C., Baker, M.S.A. & Tilghman, L.M. (2006). Differing effects of isolation on evolution of bird songs: examples from an island-mainland comparison of three species. — *Biol. J. Linn. Soc.* 89: 331-342.

- Baker, M.C., Bjerke, T.K., Lampe, H.U. & Espmark, Y.O. (1987). Sexual response of female yellowhammers to differences in regional song dialects and repertoire sizes. — *Anim. Behav.* 35: 395-401.
- Balsby, T.J.S. (2000). Song activity and variability in relation to male quality and female choice in Whitethroats *Sylvia communis*. — *J. Avian Biol.* 31: 56-62.
- Beecher, M.D. & Brenowitz, E.A. (2005). Functional aspects of song learning in songbirds. — *Trends Ecol. Evol.* 20: 143-149.
- Berry, R.J. (1983). Diversity and differentiation: the importance of island biology for general theory. — *Oikos* 41: 523-529.
- Boogert, N.J., Giraldeau, L.A. & Lefebvre, L. (2008). Song complexity correlates with learning ability in zebra finch males. — *Anim. Behav.* 76: 1735-1741.
- Borowiec, M., Losak, K., Osiejuk, T.S. & Halupka, K. (2007). Song repertoire is not affected by stress in an adult male songbird, the Whitethroat *Sylvia communis*. — *Behaviour* 144: 879-887.
- Boyer, A.G. & Jetz, W. (2010). Biogeography of body size in Pacific island birds. — *Ecography* 33: 369-379.
- Bright, A., Waas, J.R., King, C.M. & Cuning, P.D. (2004). Bill colour and correlates of male quality in blackbirds: an analysis using canonical ordination. — *Behav. Process.* 65: 123-132.
- Brumm, H., Zollinger, S.A. & Slater, P.J.B. (2009). Developmental stress affects song learning but not song complexity and vocal amplitude in zebra finches. — *Behav. Ecol. Sociobiol.* 63: 1387-1395.
- Buchanan, K.L., Catchpole, C.K., Lewis, J.W. & Lodge, A. (1999). Song as an indicator of parasitism in the sedge warbler. — *Anim. Behav.* 57: 307-314.
- Buchanan, K.L., Leitner, S., Spencer, K.A., Goldsmith, A.R. & Catchpole, C.K. (2004). Developmental stress selectively affects brain nuclei HVC in the zebra finch. — *Proc. Roy. Soc. Lond. B: Biol.* 271: 2381-2386.
- Buchanan, K.L., Spencer, K.A., Goldsmith, A.R. & Catchpole, C.K. (2003). Song as an honest signal of past developmental stress in the European starling (*Sturnus vulgaris*). — *Proc. Roy. Soc. Lond. B: Biol.* 270: 1149-1156.
- Byers, B.E. & Kroodsma, D.E. (2009). Female mate choice and songbird song repertoires. — *Anim. Behav.* 77: 13-22.
- Catchpole, C.K. (1986). Song repertoires and reproductive success in the great reed warbler *Acrocephalus arundinaceus*. — *Behav. Ecol. Sociobiol.* 19: 439-445.
- Catchpole, C.K. & Slater, P.J.B. (2008). Bird song-biological themes and variations. — Cambridge University Press, Cambridge.
- Clegg, S.M. & Owens, P.F. (2002). The 'island rule' in birds: medium body size and its ecological explanation. — *Proc. Roy. Soc. Lond. B: Biol.* 269: 1359-1365.
- Dabelsteen, T. (1984). An analysis of the full song of the Blackbird *Turdus merula* with respect to message coding and adaptations for acoustic communication. — *Ornis. Scand.* 15: 227-239.

- Dabelsteen, T. (2005). Public, private or anonymous? Facilitating and countering eavesdropping. — In: Animal communication networks (McGregor, P.K., ed.). Cambridge University Press, Cambridge, p. 38-62.
- Dabelsteen, T., Larsen, O.N. & Pedersen, S.B. (1993). Habitat-induced degradation of sound signals: quantifying the effects of communication sounds and bird location on blur ratio, excess attenuation, and signal-to-noise ratio in blackbird song. — *J. Acoust. Soc. Am.* 93: 2206-2220.
- Doutrelant, C., Blondel, J., Perret, P. & Lambrechts, M.M. (2000). Blue Tit song repertoire size, male quality and interspecific competition. — *J. Avian Biol.* 31: 360-366.
- Eens, M., Pinxten, R. & Verheyen, R.F. (1992). Song learning in captive European starlings, *Sturnus vulgaris*. — *Anim. Behav.* 44: 1131-1143.
- Faivre, B., Gregoire, A., Pr eault, M., Cezilly, F. & Sorci, G. (2003a). Immune activation rapidly mirrored in a secondary sexual trait. — *Science* 300: 103.
- Faivre, B., Pr eault, M., Th ery, M., Gaillard, M. & Cezilly, F. (2003b). Bill colour and immunocompetence in the European blackbird. — *Anim. Behav.* 65: 1125-1130.
- Faivre, B., Pr eault, M., Th ery, M., Secondi, J. & C ezilly, F. (2001). Breeding strategy and morphological characters in an urban population of blackbirds, *Turdus merula*. — *Anim. Behav.* 61: 969-974.
- Festa-Bianchet, M., Gaillard, J.M. & Jorgenson, J.T. (1998). Mass- and density-dependent reproductive success and reproductive costs in a capital breeder. — *Am. Nat.* 152: 367-379.
- Garamszegi, L., J anos, T., Hegyi, G., Sz oll osi, E., Rosivall, B. & Eens, M. (2007). Age-dependent expression of song in the collared flycatcher, *Ficedula albicollis*. — *Ethology* 113: 246-256.
- Garnett, M.C. (1981). Body size, its heritability and influence on juvenile survival among great tits, *Parus major*. — *Ibis* 123: 31-41.
- Gil, D. & Slater, P.J.B. (2000). Song organisation and singing patterns of the willow warbler, *Phylloscopus trochilus*. — *Behaviour* 137: 759-782.
- Gil, D., Cobb, J.L.S. & Slater, P.J.B. (2001). Song characteristics are age dependent in the willow warbler, *Phylloscopus trochilus*. — *Anim. Behav.* 62: 689-694.
- Glutz von Blotzheim, U. & Bauer, K. (1985). Handbuch der V ogel Mitteleuropas. — Akademische Verlagsgesellschaft, Wiesbaden.
- Gosler, A.G., Greenwood, J.J.D., Baker, J.K. & Davidson, N.C. (1998). The field determination of body size and condition in passerines: a report to the British Ringing Committee. — *Bird Study* 45: 92-103.
- Grant, P.R. (1965). The adaptive significance of some size trends in island birds. — *Evolution* 19: 355-367.
- Haccou, P. & Meelis, E. (1994). Statistical analyses of behavioural data. — Oxford University Press, Oxford.
- Hesler, N. (2010). Song complexity in common blackbirds-an honest signal of male quality? — Dissertation, University of Copenhagen, Copenhagen.
- Hesler, N., Mundry, R. & Dabelsteen, T. (2011). Does song repertoire size in common blackbirds play a role in an intra-sexual context? — *J. Ornithol.* 152: 591-601.

- Hesler, N., Mundry, R. & Dabelsteen, T. (2012). Are there age-related differences in the song repertoire size of Eurasian blackbirds? — Acta Ethol. (doi:10.1007/s10211-012-0127-z).
- Horak, P. (1994). Effect of nestling history on adult size and reproduction in the great tit. — Ornis Fennica 71: 47-54.
- Kempenaers, B., Verheyen, G.R. & Dhondt, A.A. (1997). Extrapair paternity in the blue tit (*Parus caeruleus*): female choice, male characteristics, and offspring quality. — Behav. Ecol. 8: 481-492.
- 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.
- Kiefer, S., Sommer, C., Scharff, C., Kipper, S. & Mundry, R. (2009). Tuning towards tomorrow? Common nightingales *Luscinia megarhynchos* change and increase their song repertoires from the first to the second breeding season. — J. Avian Biol. 40: 231-236.
- Kiefer, S., Spiess, A., Kipper, S., Mundry, R., Sommer, C., Hultsch, H. & Todt, D. (2006). First-year common nightingales (*Luscinia megarhynchos*) have smaller song-type repertoire sizes than older males. — Ethology 112: 1217-1224.
- Kipper, S. & Kiefer, S. (2010). Age-related changes in birds' singing styles: on fresh tunes and fading voices? — Adv. Stud. Behav. 41: 77-118.
- Kipper, S., Mundry, R., Hultsch, H. & Todt, D. (2004). Long-term persistence of song performance rules in nightingales (*Luscinia megarhynchos*): a longitudinal field study on repertoire size and composition. — Behaviour 141: 371-390.
- Kipper, S., Mundry, R., Sommer, C., Hultsch, H. & Todt, D. (2006). Song repertoire size is correlated with body measures and arrival date in common nightingales, *Luscinia megarhynchos*. — Anim. Behav. 71: 211-217.
- Klump, B., Hesler, N. & Dabelsteen, T. (2012). Quiet song in common blackbirds-repertoire size and use in an aggressive context. Conference abstract. — Bioacoustics 21: 40-41.
- Lampe, H.M. & Espmark, Y.O. (1994). Song structure reflects male quality in pied flycatchers, *Ficedula hypoleuca*. — Anim. Behav. 47: 869-876.
- Lapierre, J.M., Mennill, D.J. & MacDougall-Shackleton, E.A. (2011). Spatial and age-related variation in use of locally common song elements in dawn singing of song sparrows *Melospiza melodia*; old males sing the hits. — Behav. Ecol. Sociobiol. 65: 2149-2160.
- Leitner, S., Marshall, R.C., Leisler, B. & Catchpole, C.K. (2006). Male song quality, egg size and offspring sex in raptive canaries (*Serinus canaria*). — Ethology 112: 554-563.
- Manly, B.F.J. (1997). Randomization, bootstrap and Monte Carlo methods in biology. — Chapman & Hall, New York, NY.
- McGregor, P.K. (1992). Quantifying responses to playback: one, many, or composite multivariate measures? — In: Playback and studies of animal communication (McGregor, P.K., ed.). Plenum Press, New York, NY, p. 79-96.
- Messmer, E. & Messmer, I. (1956). Die Entwicklung der Lautäußerungen und einiger Verhaltensweisen der Amsel (*Turdus merula merula* L.) unter natürlichen Bedingungen und nach Einzelaufzucht in schalldichten Räumen. — Z. Tierpsychol. 13: 341-441.
- Mountjoy, D.J. & Lemon, R.E. (1995). Extended song learning in wild European starlings. — Anim. Behav. 49: 357-366.

- Mountjoy, D.J. & Lemon, R.E. (1996). Female choice for complex song in the European starling: a field experiment. — *Behav. Ecol. Sociobiol.* 38: 65-71.
- Mundry, R. & Fischer, J. (1998). Use of statistical programs for nonparametric tests of small samples often leads to incorrect *P* values: examples from *Animal Behaviour*. — *Anim. Behav.* 56: 256-259.
- Nicholson, J.S., Buchanan, K.L., Marshall, R.C. & Catchpole, C.K. (2007). Song sharing and repertoire size in the sedge warbler, *Acrocephalus schoenobaenus*: changes within and between years. — *Anim. Behav.* 74: 1585-1592.
- Nowicki, S., Peters, S. & Podos, J. (1998). Song learning, early nutrition and sexual selection in songbirds. — *Am. Zool.* 38: 179-190.
- Nowicki, S., Searcy, W.A. & 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.* 188: 1003-1014.
- Oberweger, K. & Goller, F. (2001). The metabolic cost of birdsong production. — *J. Exp. Biol.* 204: 3379-3388.
- O’Loughlen, A.L. & Rothstein, S.I. (1993). An extreme example of delayed vocal development-song learning in a population of wild brown-headed cowbirds. — *Anim. Behav.* 46: 293-304.
- Pfaff, J.A., Zanette, L., MacDougall-Shackleton, S.A. & MacDougall-Shackleton, E.A. (2007). Song repertoire size varies with HVC volume and is indicative of male quality in song sparrows (*Melospiza melodia*). — *Proc. Roy. Soc. Lond. B: Biol.* 274: 2035-2040.
- Poesel, A., Foerster, K. & Kempnaers, B. (2001). The dawn song of the blue tit *Parus caeruleus* and its role in sexual selection. — *Ethology* 107: 521-531.
- Potter, D.M. & Griffiths, D.J. (2006). Omnibus permutation tests of the overall null hypothesis in datasets with many covariates. — *J. Biopharm. Stat.* 16: 327-341.
- Préault, M., Chastel, O., Cezilly, F. & Faivre, B. (2005). Male bill colour and age are associated with parental abilities and breeding performance in blackbirds. — *Behav. Ecol. Sociobiol.* 58: 497-505.
- Quinn, G.P. & Keough, M.J. (2002). *Experimental designs and data analysis for biologists*. — Cambridge University Press, Cambridge.
- R Development Core Team (2008). *R: a language and environment for statistical computing*. [2.9.1]. — R Foundation for Statistical Computing, Vienna.
- Rasmussen, R. & Dabelsteen, T. (2002). Song repertoires and repertoire sharing in a local group of blackbirds. — *Bioacoustics* 13: 63-76.
- Reid, J.M., Arcese, P., Cassidy, A.L.E.V., Hiebert, S.M., Smith, J.N.M., Stoddard, P.K., Marr, A.B. & Keller, L.F. (2005). Fitness correlates of song repertoire size in free-living song sparrows (*Melospiza melodia*). — *Am. Nat.* 165: 299-310.
- Rising, J.D. & Somers, K.M. (1989). The measurement of overall body size in birds. — *Auk* 106: 666-674.
- Sacher, T. (2010). *Genetic differentiation and migration behaviour of an island population of the Common Blackbird*. — Dissertation, University of Oldenburg, Oldenburg.

- Sacher, T., Engler, J., Gorschewski, A., Gottschling, M., Hesler, N., Bairlein, F. & Coppack, T. (2006). Die Helgoländer Amselpopulation: ein Modell für Populationsgenetik und Zugbiologie. — Ornithol. Jahresber. Helgoland 16: 76-84.
- Searcy, W.A. (1992). Song repertoire and mate choice in birds. — Am. Zool. 32: 71-80.
- Searcy, W.A., McArthur, P.D. & Yasukawa, K. (1985). Song repertoire size and male quality in song sparrows. — Condor 87: 222-228.
- Snow, D.W. (1958). A study of blackbirds. — George Allen and Unwin, London.
- Spencer, K.A., Buchanan, K.L., Goldsmith, A.R. & Catchpole, C.K. (2004). Developmental stress, social rank and song complexity in the European starling (*Sturnus vulgaris*). — Proc. Roy. Soc. Lond. B: Biol. 271: S121-S123.
- Spencer, K.A., Buchanan, K.L., Leitner, S., Goldsmith, A.R. & Catchpole, C.K. (2005). Parasites affect song complexity and neural development in a songbird. — Proc. Roy. Soc. Lond. B: Biol. 272: 2037-2043.
- Stephan, B. (1999). Die Amsel: *Turdus merula*. — Wertarp, Hohenwarsleben.
- Svensson, L. (1992). Identification guide to European passerines. — Svensson, Stockholm.
- Thessing, A. & Ekman, J. (1994). Selection on the genetic and environmental components of tarsal growth in juvenile willow tits (*Parus montanus*). — J. Evol. Biol. 7: 713-726.
- Thielcke, G. (1961). Ergebnisse der Vogelstimmen-Analyse. — J. Ornithol. 102: 285-300.
- Thielcke, G. (1973). On the origin of divergence of learned signals (songs) in isolated populations. — Ibis 115: 511-516.
- Todt, D. (1968a). Wiederholung akustischer Muster im Gesang der Amsel. — Naturwissenschaften 55: 450.
- Todt, D. (1968b). Zur Steuerung unregelmäßiger Verhaltensabläufe-Ergebnisse einer Analyse des Gesanges der Amsel. — In: Kybernetik (Mittelstaedt, H., ed.). München, p. 465-485.
- Todt, D. (1981). On functions of vocal matching: effect of counter-replies on song post choice and singing. — Z. Tierpsychol. 57: 73-93.
- Ward, S., Lampe, H.M. & Slater, P.J.B. (2004). Singing is not energetically demanding for pied flycatchers, *Ficedula hypoleuca*. — Behav. Ecol. 15: 477-484.
- Weatherhead, P.J. & Boag, P.T. (1995). Pair and extra-pair mating success relative to male quality in red-winged blackbirds. — Behav. Ecol. Sociobiol. 37: 81-91.
- Wysocki, D. & Kiriaka, B. (2007). Wing to tail length ratio in European blackbirds (*Turdus merula* L.) of different age. — Pol. J. Ecol. 55: 121-125.