Are bird song complexity and song sharing shaped by habitat structure? An information theory and statistical approach

Elodie Briefer [a,*], Tomasz S. Osiejuk [b], Fanny Rybak [a], Thierry Aubin [a]

[a] University Paris 11, NAMC, CNRS-UMR8620, Orsay, France
[b] Department of Behavioural Ecology, Institute of Environmental Biology, Faculty of Biology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznań, Poland

ARTICLE INFO

ABSTRACT

In songbirds, song complexity and song sharing are features of prime importance for territorial defence and mate attraction. These aspects of song may be strongly influenced by changes in social environment caused by habitat fragmentation. We tested the hypothesis that habitat fragmentation induced by human activities influences song complexity and song sharing in the skylark, a songbird with a very large repertoire and whose population recently underwent a large decline. We applied powerful mathematical and statistical tools to assess and compare song complexity and song sharing patterns of syllables and sequences of syllables in two populations: a declining population in a fragmented habitat, in which breeding areas are separated from each other by unsuitable surroundings, and a stable population in a continuous habitat. Our results show that the structure of the habitat influences song sharing, but not song complexity. Neighbouring birds shared more syllables and sequences of syllables in the fragmented habitat than in the continuous one. Habitat fragmentation seems thus to have an effect on the composition of elements in songs, but not on the number and complexity of these elements, which may be a fixed feature of song peculiar to skylarks.

1. Introduction

Bird songs are among the most complex sounds produced by animals. These elaborate signals of oscine birds (songbirds) are thought to have evolved largely through sexual selection. The functions of songs are territorial defence and mate attraction. Between species, songs show great variation: (1) at the level of song duration, from short and discrete song types to long and continuous ones; (2) at the level of repertoire size, from one element produced repeatedly to hundred of elements produced with immense variety; (3) or even at the level of song sharing, from one song type shared by all individuals of a given area to different sets of song types produced by each individual (Catchpole and Slater, 1995).

The repertoire size of a bird is assessed as the number of different elements it can produce. What constitutes an element is not always the same between species. As many songbird species produce short and discrete songs (discontinuous singers), repertoire size is usually measured as the number of different song types produced. However, it can also be measured as the number of different syllables produced, especially for species producing continuous songs (continuous singers, Catchpole and Slater, 1995). Despite the fact that approximately 80% of all songbird species have small or moderate song repertoire sizes (1–25 different song types per individual), most of them produce more than one song type (reviewed by Lambrechts and Dhondt, 1995; Stoddard, 1996). Various modes of repertoire acquisition exist: repertoire size appears to be partially under genetic control, at least in some species (Marler and Sherman, 1985), and may depend on condition during development of young males (Nowicki et al., 2000; Buchanan et al., 2003) or on adult body condition (Lampe and Espmark, 1994; Kipper et al., 2006). Repertoire appears to vary with age (Nottebohm and Nottebohm, 1978; Loffredo and Borgia, 1986), geography (Mulligan, 1966; Kroodsma, 1981), and population size or density (Bitterbaum and Baptista, 1979; Laiolo et al., 2008). Furthermore, the composition of a repertoire is based on songs of multiple tutors and is influenced by social interactions (Nordby et al., 1999).

The phenomenon of song sharing occurs when songs of neighbouring males are more similar to one another than to those of non-neighbouring males. Variation in song composition resulting from song sharing and occurring between close groups of birds is called local dialect or microdialect and is often expressed in the syntactic organization of songs (Mundinger, 1982). Whole songs (e.g. Griessmann and Naguib, 2002), individual syllables within songs (e.g. Isaac and Marler, 1963; Kreutzer, 1974) or groups of linked syllables (e.g. Becker, 1974) are shared by

individuals from the same population. Song sharing is thought to be a consequence of the cultural transmission of song through learning (Trainer, 1983; Kroodsma and Miller, 1996). Spatial patterns of song sharing can be influenced by the social structure (nest distribution), affiliative interactions and population density (Hausberger, 1997), the specific learning strategy, the timing of learning, the learning environment, migration and dispersal patterns (Nottebohm, 1969; Goodfellow and Slater, 1986; Slater, 1989; Koetz et al., 2007).

Repertoire size, or more broadly song complexity and song sharing are both known to play important roles in territory acquisition, territory defence and mate attraction (Beecher and Brenowitz, 2005), and may result from complex interactions between genes and environment (Catchpole and Slater, 1995).

One of the effects of human activities on environment is the fragmentation of animal species' habitat in patches separated by relatively inhospitable terrain. There is considerable evidence that habitat fragmentation can cause declines in avian populations (Bender et al., 1998; Schmiegelow and Monkkonen, 2002). The main reasons of these species' declines are a reduction of habitat areas (area effect) and an increase in distance between remaining habitat patches that influences therefore dispersal ability (isolation effect, Andrén, 1994; Opdam et al., 1994; reviewed in Saunders et al., 1991; Debinski and Holt, 2000). Habitat fragmentation has been suggested to condition song variation in some bird species (white crowned sparrow Zonotrichia leucophrys nuttalli, Trainer, 1983; corn bunting Miliaria calandra, Holland et al., 1996), but direct evidence has only been shown in the Dupont's lark Chersophilius duponti (Laiolo and Tella, 2005, 2007). In this species, song diversity has been shown to be an indicator of the viability of fragmented bird populations. Song diversity could thus be used as a warning signal of population decline (Laiolo et al., 2008).

To study the effect of habitat fragmentation on both song complexity and patterns of song element sharing, we analysed the song of the skylark an ideal model species because males produce highly complex songs. The skylark is a territorial songbird of open landscape. During the breeding season, from February to the end of July, pairs settle in adjacent territories (Delius, 1963). During the entire breeding season, males fiercely defend territories (Donald, 2004). To deter intruders, they produce a flight song (Delius, 1963; Hedenström, 1995) that reveals species identity via encoding of temporal parameters (Aubin and Brémond, 1983). Unlike songs of many songbirds which are relatively short and discontinuous (Catchpole and Slater, 1995) and which can be categorized in discrete song types, the skylark song is long and continuous. It consists of series of song units, named syllables, produced between 2 and 6 kHz. With an estimated repertoire of more than 300 different syllables per individual, this song is one of the most complex songs among songbirds, giving rise to a huge potential for variation at the syntactic level (Aubin, 1981, 1982). Neighbouring males in a given breeding area share numerous syllables and sequences of syllables that constitute a microdialect (Briefer et al., 2008). Since the mid-1970s, the European skylark population underwent strong declines. The major changes which have been implicated in the decline of skylarks are agricultural intensification, increasing monoculture inducing a loss of structural diversity of vegetation, elimination of lands not used for cultivation, use of pesticides and fertilisers, and increasing mechanisation of cultivation (Busche, 1989; Odderskaer et al. 1997; Chamberlain and Gregory, 1999; Tryjanowski, 2000).

We compared two skylark populations: one living in a fragmented habitat with a low population density in decline and one living in a continuous habitat with a high and stable population density. In the fragmented habitat, territories of males were concentrated in particular breeding areas and individuals were thus gathered in distinct small groups of 'neighbours' because of the heterogeneity of the landscape (presence of woods, villages and roads). In the continuous habitat, the whole area was suitable for breeding and territories were thus unbrokenly distributed in the studied area.

We developed a method for assessing song diversity. We measured song complexity and assessed patterns of syllables and sequences of syllables shared by individuals within each habitat. Song complexity was assessed by measuring repertoire size and song versatility, and by using information theory concepts (Shannon and Weaver, 1949). Patterns of syllable and sequence sharing were assessed using various statistical tools. We then compared results obtained in the two kinds of habitat. We expected the continuity of the habitat associated with a high population density to result in songs being more varied and complex compared to the fragmented habitat, as a consequence of more tutor songs available to young birds during song learning and of stronger male–male competition.

2. Theory

2.1. Information theory

Information theory was first developed by Shannon (1948) and Shannon and Weaver (1949) to encode signals that could efficiently transmit information across a variety of noisy communication channels. The information theory method examines the amount of information, defined as a measure of one's freedom of choice when selecting a message, at increasingly complex levels of signalling organisation of a communication system. The amount of information in a communication system is measured using entropy, also called uncertainty, and represents the degree of randomness in the system. According to the level of signalling organisation, several values of entropy can be measured. Maximum entropy or zero-order entropy assesses repertoire diversity. First-order entropy represents the simple repertoire internal organisation. Higher-order entropies assess the communication system complexity.

2.2. Markov chain process

Markov chain analyses are used to determine if events occurring in a sequence are independent of one another. However, a Markov process cannot depict the sequential organisation of these events. The alternative is to use transition matrices (e.g. Isaac and Marler, 1963; Lemon and Chatfield, 1971, 1973), or kinematic graphs (e.g. Ficken et al., 1994) to study sequential organisation of syllables in bird songs. Yet, the skylark repertoire is so large that these methods are not useful to describe songs of this species. Instead, we investigated Markov chain analyses using an information graph. According to this method, the amount of information contained in a sequence of elements is at maximum if each element's probability in the sequence is completely independent of the preceding element. Thus, by plotting the average amount of information (entropy) of element sequences against sequence size, we obtain an information graph. If the probability of any element is dependent on the element or elements immediately preceding it, the amount of information will decrease as sequence size increases. Lemon and Chatfield (1973) suggested that the portion of the information graph where information declines most sharply (where there is a higher negative slope, called 'Shannon's slope') indicates the degree of Markov chaining.

2.3. Zipf–Mandelbrot

'Zipf's law' was first proposed by Zipf (1949) to describe human language. One of his proposed linguistic generalizations was

called 'principle of least effort'. According to this law, if elements are ranked according to their probability of occurrence, from the commonest to the rarest, there must be an inverse monotonic relationship between rank of use and probability of occurrence. When plotted logarithmically, the function obtained should be linear. This function gives an indication of the non-randomness and the potential capacity for information transfer in a system. A regression coefficient (slope) of -1 indicates that there is a balance between diversity and redundancy in the system and thus, that the system has a high potential capacity for transferring information. However, if the slope ranges between -1 and 0, there is more diversity than redundancy. If the slope is smaller than -1, there is more redundancy than diversity in the system and thus, less potential capacity for transferring information. Mandelbrot (1953) found a rationale for Zipf's empirical generalization and developed a formula that could fit data from a greater variety of languages. This formula is: $P_i = i (r+k)^{-s}$, with P_i the probability of occurrence of elements, r their rank of use, i the intercept, k the curvature factor and s the asymptotic slope. The resulting curve is more hyperbolic than Zipf's law at the upper data (lowest rank elements) of Zipf's plot and has been shown to describe bird communication system well (Hailman et al., 1985). Furthermore, Mandelbrot's curve gives an indication about whether the system possesses generative properties, that is, whether the number of different elements in the system is unlimited. To show openness of a communication system, the plotted curve should reach a non-zero asymptotic slope (Hailman et al., 1985; Ficken et al., 1994). Indeed, a non-zero asymptotic slope indicates that, as larger samples of songs are recorded, new element types will be encountered in the population.

3. Methods

3.1. Study areas, subjects and song recordings

We studied two skylark populations, one in France (fragmented habitat) and one in Poland (continuous habitat). Skylark breeding population is still declining in France (about 12% of population reduction between 1994 and 2000, with a more important reduction in the northern part where our study was conducted, Eraud and Boutin, 2001), but it remains stable in Poland (BirdLife International, 2008; Donald, 2004).

In France, songs were recorded during the breeding season in 2005 in the fields surrounding the University of Paris 11, Orsay, Essone. The study plot covered 81 km² and consisted of arable fields, most of them being cereals crops, surrounded by habitats unsuitable for skylarks, like anthropogenic barriers (roads and houses) and few woodlands that covered about 50% of the area (assessed using aerial pictures, Fig. 1). Due to the heterogeneity of the landscape, this population was distributed in distinct small groups separated by few kilometres. We studied seven groups established in locations of 0.44 ± 0.07 km² (mean $\pm$ SE), containing 9.6 ± 0.7 individuals each, and separated by unsuitable areas (Figs. 1 and 2). Within these locations, the bird density was of 24.2 pairs/km². One to three individuals per group were recorded, making a total of 14 individuals.

In Poland, songs were recorded during the breeding season in 2006 in the fields around Odolanów, Wielkopolska. Large scale monitoring data for the whole country indicate that skylark has a continuous distribution across Poland (Stawarczyk, 2004; Tryjanowski, 2007). The study plot covered 24 km² and consisted of arable and pastoral fields. Groups of tree and shrubs covered about 2% of the area and anthropogenic barriers about 8% (assessed using aerial pictures, Fig. 1). Birds of this population were dispersed in the habitat without unsuitable areas between them. The bird density within this study area exceeded two times the bird density in the fragmented habitat (48.8 pairs/km²). To allow comparison at similar distances with the French population, we recorded 20 individuals established in two sets that covered 3.1 and 2.7 km² respectively and that were situated 2 km apart (Figs. 1 and 2). The populations of France and Poland were situated 1100 km apart.

During the breeding season, site fidelity is very strong (Jenny, 1990; Delius, 1963). Once territories are established, boundaries between adjoining territories are stable and males remain inside these boundaries (Aubin, 1981). We recorded songs at the middle of the breeding season, in May, when territories are stable. To avoid individual misidentification, birds established in neighbouring territories were recorded during a single visit. For each recorded bird, we took GPS coordinates (Garmin GPSMAP 76S) at the place where we had seen the bird singing. We then used these

Fig. 1. Aerial photographs of the two studied areas: (A) fragmented habitat (Orsay, Essone, France); (B) continuous habitat (Odolanów, Wielkopolska, Poland). The locations of the seven studied groups in the fragmented habitat and of the two studied sets of individuals in the continuous habitat are indicated (Google Earth version 4.3, Google, 2008).
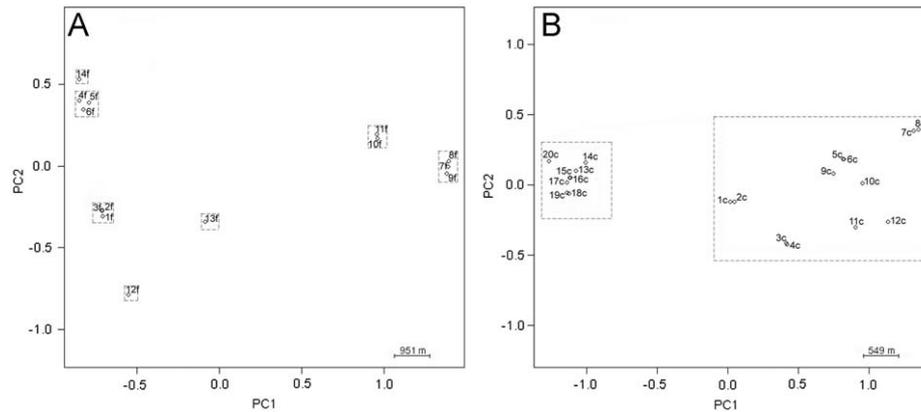
Fig. 2. Geographical distances between individuals. Plot, for the fragmented habitat (A) and continuous habitat (B), of the first principal component (PC1) against the second principal component (PC2) extracted from the analysis in principal coordinates. Individuals are identified with a number followed by 'f' in the fragmented habitat and followed by 'c' in the continuous one. For the fragmented habitat, dotted rectangles delimit the seven studied groups of individuals (1–3 individuals recorded per group). For the continuous habitat, dotted rectangles delimit the two studied sets of individuals. Scales were estimated by comparing distances separating pairs of individuals on the plots to real geographical distances between them. These scales are $\pm$ 10 m accurate.

GPS coordinates to calculate, in each kind of habitat, geographical distances between territories of all pairs of recorded individuals. We calculated distances between two GPS coordinates in metres with a calculator using a spherical earth assumption.

We recorded several songs per individual between 0700 and 1200 h Eastern Daylight Time using a Marantz PMD 690 digital recorder (sampling rate: 48 000 Hz) connected to a Sennheiser ME 64 K6 omni-directional microphone (frequency response: 30 Hz to 20 kHz $\pm$ 1 dB) mounted on a Telinga Universal parabola (diameter: 50 cm). We then transferred songs files to a computer and high-pass filtered them (cut-off frequency: 1600 Hz) to remove the background noise. We used the Avisoft SASLab pro version 4.31 (Specht, 2004) for subsequent analyses.

3.2. Analysis of songs

We selected for the analysis 100.33 $\pm$ 0.21 s (mean $\pm$ SE, N=34 individuals) of continuous song per individual of one good signal to noise ratio song (i.e. a song with a low level of background noise). Such selection allowed us to standardize the song duration analysed per individual. We visualized songs on a spectrogram (FFT-Length: 1024; Frame: 100%; Bandwidth: 61 Hz; Resolution: 46 Hz, Hamming window).

A syllable was defined as a continuous trace on the sound spectrogram or a group of continuous traces spaced out by less than 25 ms. As described elsewhere (Briefer et al., 2008), we labelled syllables on the sound spectrogram with a number according to their overall frequency modulation shapes. The same syllable found at several places in the song of one individual or in songs of different individuals was labelled with the same number. Then, we examined the sequential organization of syllables using a custom Matlab program (The MathWorks, Natick, MA, USA; see Lehongre et al., 2008). The numbers corresponding to the syllable labelling were inputted into the Matlab program in the order of production of the corresponding syllables in the songs. For each individual, the program classified sequences of numbers according to their length and the number of times they were repeated. We thus obtained, for each individual, the repertoire of all syllables and sequences of 2–10 syllables with the number of times they were repeated.

3.3. Song complexity

To assess the effect of the kind of habitat (continuous versus fragmented) on the complexity of songs, we used the repertoires

of syllables and of sequences of 2–10 syllables produced by each individual to calculate complexity and versatility of songs. We then applied information theory concepts (Shannon and Weaver, 1949) to calculate entropy. As skylarks do not sing discontinuous songs that can be classified like most other passerine birds, syllable types and not song types were the basic units for our measures.

3.3.1. Versatility: variety of syllables and complexity of sequences within songs

To assess versatility of songs, we used the following four measures (e.g. Kroodsma, 1978; Ince and Slater, 1985):

(a) *Repertoire size: syllable composition of songs.* We measured the *observed repertoire size* (*ObsR*) of each individual as the total number of different syllables produced. As skylark repertoire size is huge (Briefer et al., 2008), the *total repertoire size* (*R*) of each individual was estimated following the method developed by Wildenthal (1965) and Howard (1974). This method assumes that the curve of the number of new encountered syllable types plotted against the number of analysed syllables rises to meet an asymptote (*R*) following an exponential equation. This method is applied under the following assumptions: the probability of a new syllable type appearing is proportional to the number of unused syllable types remaining in the repertoire; the repertoire size is constant and fixed; and syllables are produced in random sequences. When 300 s of skylark song were analysed, the obtained repertoire curves indeed reached an asymptote (Briefer et al., 2008). We derived the more approximate exponential equation fitting skylark repertoire curves from Wildenthal (1965) equation as follows: $n=R\,(1-e^{-(T/\alpha)/R})$, with n=the number of new encountered syllable types, T=the total number of analysed syllables, R=the total repertoire size, and α=1.86. The latter value was estimated from the curves of nine individuals used in our previous study for which the total repertoire size (*R*) was studied by analysing more than 300 s of song.

(b) *Syllable type versatility: variety of syllables within sequences.* We calculated *syllable type versatility* (*SylVer*) for each individual as the mean number of different syllables in sequences of 10 syllables. *SylVer* ranges from 0 to 10, with 10 being the maximum versatility.

(c) *Transition versatility: repetition of syllables within sequences.* We calculated *transition versatility* (*TrVer*) for each individual, as the mean number of transitions between two different syllables in sequences of 10 syllables. *TrVer* ranges from 0 to 9, with 9 being the maximum versatility.

(d) *Total versatility: variety and repetitions of syllables within sequences.* We calculated the *total versatility* (*TotVer*) for each individual as the mean value of *SylVer ∗ TrVer* in sequences of 10 syllables. Total versatility ranges from 0 to 90, with 90 being the maximum versatility.

3.3.2. Information analysis: amount of information in sequences within songs

We applied the information theory concepts to skylark song by calculating the following entropy measures for each individual. Entropy values are measured in bits ('binary digit') per syllable type.

(a) *Zero-order entropy: syllable song diversity.* We calculated *zero-order entropy* (E_0) as follows: $E_0 = \log_2(k)$, with k being the observed repertoire size (*ObsR*). Additionally, we calculated the *estimated zero-order entropy* (*EstE_0*) by using the following equation: $E_0 = \log_2(k')$, with k' being the estimated repertoire size (R). These entropy values estimate song diversity.

(b) *First-order entropy: simple internal organisation of songs.* First-order entropy (E_1) takes into account the probability of occurrence of each syllable type in a given song (Pi), which is equal to the number of times that syllable type is observed divided by the total number of syllables in the song. We calculated E_1 as follows: $E_1 = -\sum P_i \log_2(P_i)$. Such entropy value assesses the simple internal organisation of a song.

(c) *Higher-order entropies: complex internal organisation of songs.* We then measured entropy values associated with each size of sequences composed of 2–10 syllables, by calculating *second-* (E_2) to *tenth-order* (E_{10}) *entropy* respectively. These *higher-order entropies* measure song complexity by assigning to each sequence of a given size a joint probability of occurrence and a conditional probability of occurrence in the song. The joint probability (P_{ijkl}) of a sequence type of four syllables (i–j–k–l), for example, is equal to the number of times this sequence type is observed divided by the total number of sequences of four syllables in the song. The conditional probability ($P_{l/ijk}$) of this sequence type is the probability that (i–j–k–l) occurs in the song given that the sequence type of three syllables (i–j–k) just occurred. $P_{l/ijk}$ is calculated as the joint probability of the sequence type of 4 syllables (i–j–k–l) divided by the joint probability of the sequence type of 3 syllables (i–j–k). E_2 to E_{10} were calculated as follows: $En = -\sum P_{ij} \log_2(P_{i/j})$, where P_{ij} is the joint probability, $P_{j/i}$ is the conditional probability, and n is the order of entropy. These entropy values measure the amount of information (freedom of choice) contained at increasing levels of song organisation.

(d) *Relative entropy and redundancy: proportion of free choice in building sequences.* Additionally, we calculated the relative entropy (*RE*) of songs, which is the ratio of the actual to the zero-order entropy: $REn = En/E_0$, with n being the order of entropy (1–10). *RE* is a measure of the proportion of free choice of syllables the sender has to build a sequence.

We then calculated the redundancy (*RD*), which is calculated as $RDn = 1 - REn$, with n being the order of entropy (1–10). *RD* is a measure of the fraction of the syllable sequence structure, which is determined not by free choice of the sender but rather by statistical rules governing the use of syllables.

3.3.3. Markov chain analysis: order of transition between syllables within sequences

We carried out a Markov chain analysis of songs by using an *information graph*. We drew such graph by plotting the mean values of entropy (E_0–E_{10}) obtained in each habitat against sequence size (Entropy orders: 0–10 respectively).

3.3.4. Zipf–Mandelbrot: openness of the communication system

To assess openness of the skylark communication system, we drew a *Zipf–Mandelbrot plot* in each kind of habitat by plotting the logarithm of the probability of occurrence of each syllable type against the logarithm if its rank of use in the population.

3.3.5. Statistical analyses

We compared all values of song complexity measured in the fragmented and continuous habitat with independent samples *t*-tests using Statistica version 6 (StatSoft, 2001).

3.4. Song sharing

To assess the effect of habitat type (continuous versus fragmented) on geographical patterns of song sharing, we calculated syllable and sequence sharing between pairs of individuals.

3.4.1. Syllables and sequences shared by individuals: coefficients of repertoire similarity

We measured syllable and sequence sharing between each pair of individuals within each habitat using coefficients of repertoire similarity (*RS*). Coefficient of syllable repertoire similarity and coefficients of similarity between repertoires of each sequence size (2–10 syllables) were calculated as follows:

$RS = Z/((X + Y) - Z)$, with X and Y being the total number of syllables or sequences produced by males x and y, and Z being the number of syllables or sequences shared by males x and y (Hultsch and Todt, 1981). *RS* values range from 0 to 1, with 1 being maximum sharing.

3.4.2. Statistical analyses

(a) *Analysis in principal coordinates.* We carried out an analysis in principal coordinates (PCO) for each habitat on the matrix of geographical distances calculated between pairs of individuals.

(b) *Hierarchical cluster analysis: spacing patterns of syllable and sequence sharing.* We carried out a hierarchical cluster analysis (HCA) using the average agglomeration method. Before performing the HCA, *RS* values were converted to distances by calculating $1 - RS$. Then, we made a HCA on each of the matrices of $1 - RS$ values calculated between repertoires of syllables and sequences of pairs of individuals.

(c) *Comparisons between syllables and sequences shared by close and distant individuals in the two habitats.* Conventional parametric and non-parametric tests are not suitable for analyses in which each individual is included several times in the different pair-wise comparisons (Sokal and Rohlf, 1995; Mundry, 1999). Thus, we compared *RS* values calculated between pairs of individuals situated at comparable geographic distances in the two kinds of habitat with two-tailed exact permutation tests using Monte Carlo method. We separated pairs of individuals in two classes according to the distances between them: pairs of birds situated at close geographical distance, corresponding to the within group distance in the fragmented habitat, were referred to as 'close individuals'; pairs of birds situated at a greater distance were referred to as 'distant individuals'. First, we made comparisons within each habitat between *RS* values calculated for pairs of close and distant individuals. Then, we made a comparison between *RS* values calculated for pairs of close individuals in the fragmented habitat and in the continuous one. The same comparison was made between *RS* values calculated for pairs of distant individuals.

(d) *Correlations between syllable and sequence sharing and geographical distances.* We tested correlations between *RS* values and geographical distances within each habitat for close and

distant individuals separately. As values of *RS* or geographical distances are not independent of each other (each individual is included several times), we used a permuted correlation test (1000 permutations) for vectors of numeric values containing distances or similarities. This test corresponds to a Mantel test, which is a statistical test of the correlation between two distance matrices dealing with the problem of dependance between values by using a randomization test (Mantel, 1967; see also Sokal and Rohlf, 1995; Rossi, 1996).

(e) *ANCOVA: effect of the habitat and of the proximity between birds on syllable and sequence sharing.* Additionally, we carried out analyses of covariance (ANCOVA) to assess the effect of the habitat (continuous vs. fragmented) and of the proximity between birds (close vs. distant) on syllable and sequence sharing by controlling for geographic distance effects. First, we performed two-way ANCOVAs on each habitat, with *RS* values as a dependant variable, the proximity between birds (close vs. distant) as a categorical predictor (fixed effect), and the distance between birds as a continuous predictor (covariate). Then, we performed a two-way ANCOVA including both habitats, with *RS* values as a dependant variable, the proximity between birds (close vs. distant) and the habitat (continuous vs. fragmented) as categorical predictors (fixed effects), and the distance between birds as a continuous predictor (covariate).

We carried out statistical analysis using R (R development core team, 2007, PCO and HCA, package ade4 version 1.4–5, Chessel et al., 2004; Permuted correlation test, package simba version 0.2–5, Jurasinski, 2007), StatXact version 3.1 (Permutation tests, Cytel, 1997), and Statistica version 6 (ANCOVAs, StatSoft, 2001).

4. Results

Abbreviations for measured parameters are listed in Table 1.

4.1. Song complexity

4.1.1. Versatility: variety of syllables and complexity of sequences within songs

In both habitats, the mean estimated repertoire size per individual was larger than 340 different syllables (Table 2). Mean values of *SylVer* and *TrVer* for each habitat were close to the maximum versatility (10 and 9 respectively), inducing high mean values of *TotVer* (Table 2). Mean *TrVer* values (*TrVer*=7.3 for each habitat) indicate that more than 81% (7.3/9) of transitions were between different syllables. There was a high correlation between observed and estimated syllable repertoire curves: mean $\pm$ SE Pearson correlation coefficient=0.992 $\pm$ 0.002 (range: 0.949–0.999, *N*=34 individuals, Fig. 3).

Comparisons between versatility values measured in the fragmented habitat and those values measured in the continuous one did not show any difference (independent samples *t*-test,

Table 1
List of abbreviations and their corresponding parameters.

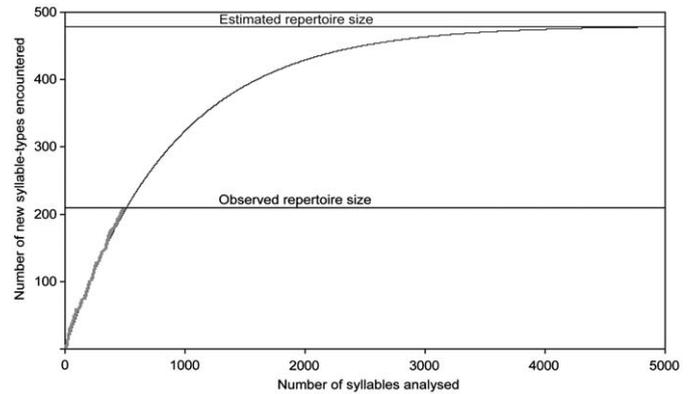| Abbreviation	Parameter
ObsR	Observed repertoire size
R	Total repertoire size
SylVer	Syllable type versatility
TrVer	Transition versatility
TotVer	Total versatility
En	*n*-Order entropy
REn	*n*-Order relative entropy
RDn	*n*-Order redundancy
RS	Coefficient of repertoire similarity

Fig. 3. Repertoire size: example of observed (grey) and estimated (black) repertoire curves obtained for one individual in the continuous habitat. This individual has an observed repertoire size of 207 syllables and an estimated repertoire size of 480 syllables.

df=32: *ObsR*, $t=-0.96$, $P=0.35$; *R*, $t=0.15$, $P=0.88$; *SylVer*, $t=-0.82$, $P=0.42$; *TrVer* $t=-0.12$, $P=0.91$; *TotVer* $t=-0.67$, $P=0.51$).

4.1.2. Information analysis: amount of information in sequences within songs

The mean $EstE_0$ value per individual for the two habitats was of 8.4 (Table 3). The mean RE_1 value for the two habitats was of 0.94 (Table 3), indicating that males had 94% of free choice of syllables when they build a sequence. The remaining fraction of the structure of the syllable sequence, which is determined not by the free choice of the bird (RD_1), was thus very low (mean RD_1 values < 0.07, Table 3). The fraction of free choice dropped at the level of sequences of two syllables (mean RE_2 value for the two habitats=0.13, Table 3). Entropy values did not differ between the fragmented and continuous habitat (independent samples *t*-test, df=32: $P \geq 0.099$ for all values).

4.1.3. Markov chain analysis: order of transition between syllables within sequences

Entropy values decreased as a function of orders (Fig. 4), indicating that the amount of information contained in sequences of syllables decreased as their size increased. Information declined most sharply, i.e. there was a higher negative slope, between first- (E_1) and second-order (E_2) entropy values. This indicates that the skylark song follows a first order Markov chain, with each syllable especially dependent on the syllable immediately preceding it. There was no subsequent drop between higher-order entropy values, indicating that each syllable was highly dependent on the syllable immediately preceding it, but not on previous ones.

4.1.4. Zipf–Mandelbrot: openness of the communication system

The *Zipf–Mandelbrot plot* showed that skylark song does not fit the 'principle of least effort' (Zipf's law), but is closer to Mandelbrot's modification with a hyperbolic curve at the lowest rank syllables (Fig. 5). Both curves obtained for the fragmented and continuous habitats seemed to reach a zero asymptotic slope.

4.2. Song sharing

As the information graph revealed a first order Markov chain, with most song information contained in the syllables and sequences of 2 syllables and little information contained in higher order sequences (3–10 syllables), we chose to group repertoires of 3–10 syllables and to carry out all subsequent analyses on the coefficient of syllable repertoire similarity (RS_1), the coefficient of similarity between repertoires of sequences of 2 syllables (RS_2), and the coefficient of similarity between repertoires of sequences

Table 2
Versatility.

Habitat	Nb Syl analysed	ObsR	R	Nb Seq analysed	SylVer	TrVer	TotVer
Fragmented ($N=14$)	498.5 ± 8.2	172.4 ± 8.6	348.6 ± 23.0	49.5 ± 0.8	7.0 ± 0.2	7.3 ± 0.2	54.6 ± 2.6
Continuous ($N=20$)	492.5 ± 7.5	181.7 ± 5.4	344.0 ± 20.2	48.8 ± 0.8	7.2 ± 0.1	7.3 ± 0.1	56.3 ± 1.1

Mean $\pm$ SE values of versatility obtained in the two kinds of habitat for each individual: total number of syllables analysed (*Nb Syl analysed*), observed repertoire size (*ObsR*), estimated repertoire size (*R*), total number of sequences of 10 syllables analysed (*Nb Seq analysed*), syllable type versatility (*SylVer*), transition versatility (*TrVer*) and total versatility (*TotVer*).

of 3–10 syllables (RS_{3-10}). We calculated RS_{3-10} by first grouping for each individual the repertoires of 3–10 syllables and then calculating coefficients of similarities between pairs of individuals.

4.2.1. Geographical distances between territories of individuals

Geographical distances between territories of pairs of individuals ranged from 16.3 to 10 195.8 m in the fragmented habitat, and from 15.0 to 5766.8 m in the continuous habitat (Fig. 2). In the fragmented habitat, within group distances ranged from 16.3 to 355.4 m, and between group distances ranged from 579.2 to 10 195.8 m. 'Close individuals' were defined as birds separated by distances corresponding to the within group distances in the fragmented habitat. We thus considered as 'close individuals' birds that were separated by distances shorter than 355.4 m, and as 'distant individuals' birds that were separated by greater distances.

4.2.2. Syllables and sequences shared by individuals: coefficients of repertoire similarity

(a) *Comparisons between syllables and sequences shared by close and distant individuals in the two habitats.* In both habitats, pairs of close individuals shared many sequences of 2–10 syllables, whereas distant individuals shared almost no sequence (Table 4). To compare habitats, RS calculated for pairs of distant individuals situated between 585.0 and 5445.4 m in the continuous habitat and between 579.21 and 5401.86 m in the fragmented habitat were included in the permutation tests. All RS calculated between close individuals were included in the permutation tests. In both kinds of habitat, pairs of close individuals shared more syllables and sequences than pairs of distant individuals. However, two pairs of close individuals in the fragmented habitat shared more syllables and more sequences than two pairs of close individuals in the continuous habitat, although these individuals were situated at comparable distances from each other in the two kinds of habitat.

Indeed, for both habitats, RS between close individuals were greater than RS between distant individuals (Permutation tests using Monte Carlo method: Fragmented habitat, $N=10$ pairs of close individuals and 37 pairs of distant individuals, $P < 0.0001$ for RS_1, RS_2 and RS_{3-10}; Continuous habitat, $N=24$ pairs of close individuals and 145 pairs of distant individuals, $P < 0.0001$ for RS_1, RS_2 and RS_{3-10}). RS_1, RS_2 and RS_{3-10} between close individuals were greater in the fragmented habitat compared to the continuous habitat (Permutation tests using Monte Carlo method, $N=34$ pairs: RS_1, $P < 0.05$; RS_2, $P < 0.01$; RS_{3-10}, $P < 0.001$). On the other hand, there was no differences between RS values calculated for pairs of distant individuals in the two kinds of habitat (Permutation tests using Monte Carlo method, $N=182$ pairs: RS_1, $P=0.09$; RS_2, $P=0.12$; RS_{3-10}, $P=0.51$).

(b) *Correlations between syllable and sequence sharing and geographical distances.* In the continuous habitat, the amount of syllables and sequences shared by distant individuals decreased as the distance between their territories increased. In the

fragmented habitat, only the amount of syllables shared by distant individuals depended on the distance separating them. At the opposite, the amount of sequences shared by distant individuals in this habitat and the amount of syllables and sequences shared by close individuals in both habitats did not depend on the distance (Fig. 6).

At more than 2000 m of distance, in both habitats, individuals shared almost no sequences of 3–10 syllables (RS_{3-10}), whereas some syllables (RS_1 around 0.1) and few sequences of 2 syllables (RS_2 around 0.03) were still shared at more than 10 000 m. Results of permuted correlation tests showed that, in the continuous habitat, RS_1, RS_2 and RS_{3-10} calculated between pairs of distant individuals were negatively correlated with geographical distances. In the fragmented habitat, RS_1 calculated between pairs of distant individuals was negatively correlated with the distance, but this was not the case for RS_2 and RS_{3-10} (weak correlations, Pearson correlation coefficients between -0.5 and 0). In both habitats, RS_1, RS_2 and RS_{3-10} calculated for pairs of close individuals were not significantly correlated with geographical distances (Fig. 6).

(c) *Spacing patterns of syllable and sequence sharing.* In both habitats, dendrograms of hierarchial cluster analyses showed differences between syllable and sequence sharing patterns (Fig. 7). The pattern of shared sequences of 2 syllables (dendrogram $1 - RS_2$) seemed to be intermediate between sharing patterns of syllables (dendrogram $1 - RS_1$) and of sequences of 3–10 syllables (dendrogram $1 - RS_{3-10}$), with some clusters of individuals being identical to the syllable sharing pattern ($1 - RS_1$) and other clusters being identical to the sharing pattern of sequence of 3–10 syllables ($1 - RS_{3-10}$).

Fragmented habitat. In the fragmented habitat, syllable sharing pattern reflected geographical distances, whereas sequence sharing patterns reflected the grouping of birds (close vs. distant individuals). Thus, the studied groups corresponded to discrete patches of acoustic variability at the level of sequences.

In each of the dendrograms of $1 - RS_1$, $1 - RS_2$ and $1 - RS_{3-10}$, individuals were clustered in groups reflecting geographical distances, except for two groups (7f–8f–9f and 10f–11f) that were different in $1 - RS_2$ (cf. Fig. 2). RS distance between groups was greater in the dendrogram of $1 - RS_{3-10}$ than in the dendrogram of $1 - RS_2$ and than in the dendrogram of $1 - RS_1$. In the dendrogram of $1 - RS_1$, individuals for which no other birds of their group had been recorded (12f, 13f and 14f) were clustered with the group that was at closest geographical distance (12f and 13f with the group 1f–2f–3f; 14f with the group 4f–5f–6f). In the dendrograms of $1 - RS_2$ and $1 - RS_{3-10}$, they were excluded from these clusters (Fig. 7).

Continuous habitat. No discrete patches of acoustic variability at the level of syllables or sequences were found in the continuous habitat, with high song sharing between birds situated at short distances and random variation among birds situated at longer distances. It seems that acoustic dissimilarities increased with geographical distance.

In each of the dendrograms of $1 - RS_1$, $1 - RS_2$ and $1 - RS_{3-10}$, individuals were clustered in two sets of individuals, reflecting

Table 3
Information analysis.

Habitat	E_0	$EstE_0$	E_1	E_2	E_3	RE_1	RE_2	RE_3	RD_1	RD_2	RD_3
Fragmented ($N=14$)	7.402 ± 0.081	8.398 ± 0.108	6.941 ± 0.102	0.952 ± 0.031	0.326 ± 0.024	0.937 ± 0.004	0.129 ± 0.004	0.045 ± 0.003	0.063 ± 0.004	0.871 ± 0.004	0.956 ± 0.004
Continuous ($N=20$)	7.492 ± 0.044	8.373 ± 0.093	7.047 ± 0.051	1.004 ± 0.024	0.286 ± 0.011	0.941 ± 0.002	0.134 ± 0.003	0.038 ± 0.002	0.059 ± 0.002	0.866 ± 0.003	0.962 ± 0.002

Mean $\pm$ SE values of entropy obtained in the two kinds of habitat for each individual: entropy (E) (in bits/syllable-type), relative entropy (RE) and Redundancy (RD). Only first- (E_1), second- (E_2) and third- (E_3) order entropy values are presented here. Higher-order entropy values are presented in the information graph (Fig. 4).
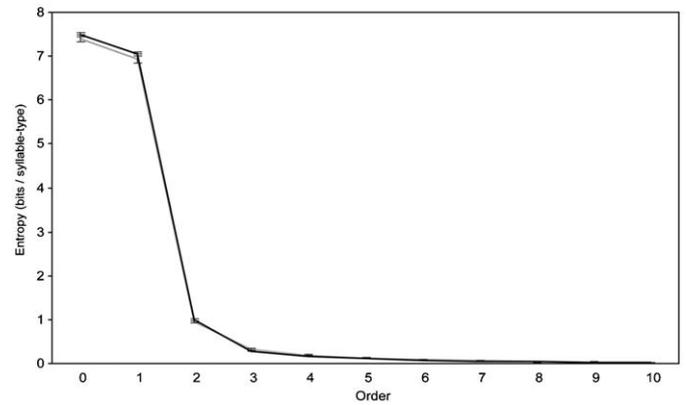

Fig. 4. Information graph: mean $\pm$ SE entropy values obtained in the fragmented habitat (grey, $N=14$ individuals) and in the continuous habitat (black, $N=20$ individuals) plotted as a function of entropy orders (size of syllable sequences).
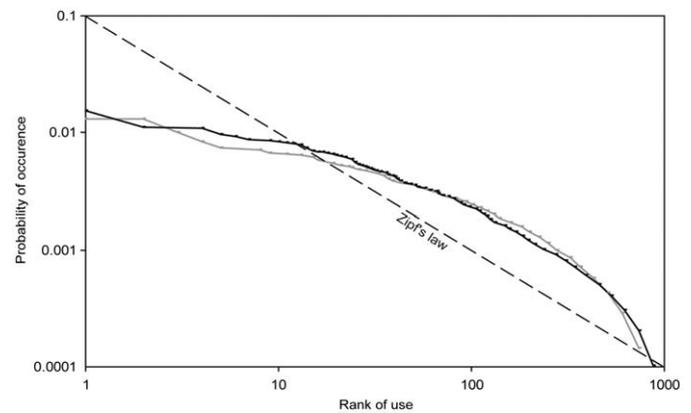

Fig. 5. Zipf plot: logarithm of the probability of occurrence of each syllable type encountered in the fragmented habitat (grey curve, $N=952$ syllable types) and in the continuous habitat (black curve, $N=1109$ syllable types) plotted against the logarithm of syllable type rank of use. The dotted line represents Zipf's law (regression coefficient $= -1.00$) that fits human language. The skylark song is more close to Mandelbrot's modification than Zipf's law, with a hyperbolic curve at the upper data (lowest rank elements).

geographical distances (cf. Fig. 2), except two individuals (1c and 2c) that were isolated in the dendrograms of $1-RS_2$ and $1-RS_{3-10}$. These two individuals were situated, within their set of individuals, at short geographical distances from the other set. In each of these dendrograms, individuals were often clustered two by two. Within these clusters, individuals were clustered with another individual situated at short geographical distance, but not always with the closest individual (e.g. 9c–10c and 19c–17c in $1-RS_1$; 5c–9c and 14c–15c in $1-RS_2$). Furthermore, these clusters of two individuals were not always identical in dendrograms $1-RS_1$; $1-RS_2$ and $1-RS_{3-10}$, except four of them (1c–2c, 3c–4c, 11c–12c and 7c–8c) that were identical in all dendrograms and three of them (5c–9c, 14c–15c and 18c–19c) that were identical in two dendrograms. However, none of the dendrograms of $1-RS_1$, $1-RS_2$ and $1-RS_{3-10}$ allowed us to distinguish a discrete cluster of more than two individuals (Fig. 7).

(d) *ANCOVA: effect of the habitat and of the proximity between birds on syllable and sequence sharing.* Results of two-way ANCOVAs corroborated with the cluster dendrograms of *RS* in showing that there was a continuous increase in acoustic dissimilarities with increasing geographical distances in the continuous habitat, and discrete patches of acoustic variability at the level of sequences (corresponding to the groups) in the fragmented habitat (Table 5). In this kind of habitat, sequence

Table 4
Repertoire similarity.

Habitat	C or D	*Nb Syl*	Nb Seq 2	*Nb Seq 3–10*	RS_1	RS_2	RS_{3-10}
Fragmented (*N*=91 pairs)	C (*N*=10)	80.4 ± 6.0	11.7 ± 1.2	21.6 ± 3.2	0.300 ± 0.016	0.231 ± 0.024	0.239 ± 0.029
	D (*N*=81)	38.8 ± 1.4	1.4 ± 0.1	0.4 ± 0.1	0.125 ± 0.004	0.025 ± 0.003	0.003 ± 0.001
Continuous (*N*=190 pairs)	C (*N*=24)	76.0 ± 3.2	11.0 ± 1.1	14.2 ± 1.8	0.257 ± 0.011	0.153 ± 0.015	0.126 ± 0.018
	D (*N*=166)	49.6 ± 1.0	2.8 ± 0.3	1.6 ± 0.3	0.159 ± 0.003	0.035 ± 0.003	0.013 ± 0.003

Mean ± SE, within each kind of habitat (fragmented and continuous), number of syllables (*Nb Syl*), number of sequences of 2 syllables (*Nb Seq 2*) and number of sequences of 3–10 syllables (*Nb Seq 3–10*) shared between pairs of close individuals (C) and between pairs of distant individuals (D), with the associated coefficients of repertoire similarity (RS_1, RS_2 and RS_{3-10} respectively).

sharing depended on the grouping of birds (same group: close individuals; different group: distant individuals), more than on the distance between them. Habitat fragmentation affected both syllable and sequence sharing.

In both habitats, the proximity between birds (close vs. distant) had a significant effect on RS_1, RS_2 and RS_{3-10}, which indicates that syllable and sequence sharing differed between pairs of close individuals and pairs of distant individuals, as shown by the permutation tests. The distance had a significant effect on RS_1 and RS_2 for both kinds of habitats, indicating that sharing of syllables and of sequences of 2 syllables by a pair of birds depended on the distance between them. The distance had also a significant effect on RS_{3-10} in the continuous, but not in the fragmented habitat. Sharing of sequences of 3–10 syllables thus depended on the distance between birds in the continuous, but not in the fragmented habitat. Furthermore, the ANCOVA carried out on RS_1, RS_2 and RS_{3-10} in both kinds of habitats showed that the kind of habitat had an effect on each of these values. The significant interaction effect between the kind of habitat and the proximity between birds on each RS value indicates that the amount of syllables and sequences shared by close or distant individuals differed between the two kinds of habitats (Table 5).

(e) *Syllable sharing between the two populations.* RS_1 calculated between repertoires of all syllables encountered in the fragmented habitat (French population) and in the continuous one (Polish population) was 0.32, RS_2 0.01 and RS_{3-10} 0. RS calculated between two individuals taken at random in each population were 0.15 for RS_1, 0.001 for RS_2 (one sequence of 2 syllables shared) and 0 for RS_{3-10}. Thus, the RS_1 value of 0.1 observed at more than 10 000 m in Fig. 6 seemed to persist over hundreds of kilometres.

5. Discussion

Mathematical tools like measures of versatility, information theory concepts, Markov chain analyses and Zipf'law allowed us to describe the complexity of skylark song according to the spatial distribution of individuals. Then, calculation of coefficients of repertoire similarity associated with various statistical tools (hierarchical cluster analyses, exact permutation tests, permuted correlation tests and ANCOVAs) enabled us to assess the influence of habitat structure on song sharing, a key feature in social interactions in birds.

5.1. Song complexity

Our results show that skylark songs are among the most complex acoustic signals compared to other songbird species. Indeed, the repertoire size of more than 300 different syllables is huge compared to species with a large repertoire that has been estimated in number of different syllables, like the thrush nightingales (49 basic song components, Griessmann and Naguib, 2002), the sedge warbler *Acrocephalus schoenobaenus* (37–54

syllables, Catchpole, 1980), the canary *Serinus canaria* (81 syllables, Leitner et al., 2001), the rock wren *Salpinctes obsoletus* (69–119 syllables, Kroodsma, 1975) or even the mockingbird *Mimus polyglottos* (66–244 syllables, Wildenthal, 1965).

We found a substantial immediate variety in the emission of syllables, as 81% of transitions (mean *TrVer* for the two habitats) are between different syllables. The potential amount of information in songs is thus particularly important at the level of syllables, with 94% (mean RE_1 for the two habitats) of free choice of syllable types. Skylark song can also be considered as highly versatile with a total average versatility of 55, exceeding the value of 50 suggested as a criterion for characterising songs as highly versatile by Ince and Slater (1985). Furthermore, information measures showed that syllable repertoire diversity (mean zero-order entropy for the two habitats: 7.4) and the amount of information contained at the level of syllables within songs (mean first-order entropy for the two habitats: 7.0) were comparable to or higher than values measured for the rufous-bellied thrush *Turdus rufiventris* (first order: 3.19, Da Silva et al., 2000), the veery *H. fuscescens* (first order: 2), the hermit thrush *H. guttata* (first order: 3.4), the wood thrush *Catharus mustelina* (first order: 3.5–4.5), the robin *T. migratorius* (first order, 3.5–4, Dobson and Lemon, 1978) or the European starling *Sturnus vulgaris* (zero-order: 7 and first-order: 6, Gentner and Hulse, 1998). First-order entropy values obtained for skylark songs are also higher than values found in non-bird species vocalizations, like male bullfrog *Rana castebeiana* advertisement calls (first order: 1.31, Suggs and Simmons, 2005), bottlenose dolphin *Tursiops truncatus* whistles (first order: 1.92), squirrel monkey *Saimiri sciureus* chucks (first order: 2.21, McCowan et al., 2002), humpback whale *Megaptera novaeangliae* songs (Suzuki et al., 2005), or even Russian, English or Arabic letters (first order: 4.03–4.35, McCowan et al., 2002).

The amount of information of skylark song drops at the level of sequences of two syllables (13% of free choice of sequence types, mean RE_2 for the two habitats), indicating a first order Markov chain, with each syllable being highly dependent on the syllable immediately preceding, but not on other ones. Thus, it seems that the use of sequences of two syllables is determined by more restrictive mechanical and/or biological constraints, whereas the use of syllables and chaining of more than two syllables is determined by the free choice of the bird.

Zipf–Mandelbrot plots showed that skylark syllable communication system resembles a 'closed' system more than an 'open' system with generative properties: Mandelbrot curves reached a zero asymptotic slope, indicating that the syllable repertoire of skylark populations seems to be fixed in time. As the likelihood of genetic determination of a whole syllable repertoire is unlikely in oscines, the existence of such a fixed repertoire could be explained by a learning strategy based on strict imitation by young birds of syllables produced by tutors, without improvisation. Indeed, in songbirds, individuals develop their songs by imitating tutors' songs, by improvising on tutors' songs or by inventing songs with minimal reference to external models (Beecher and Brenowitz,
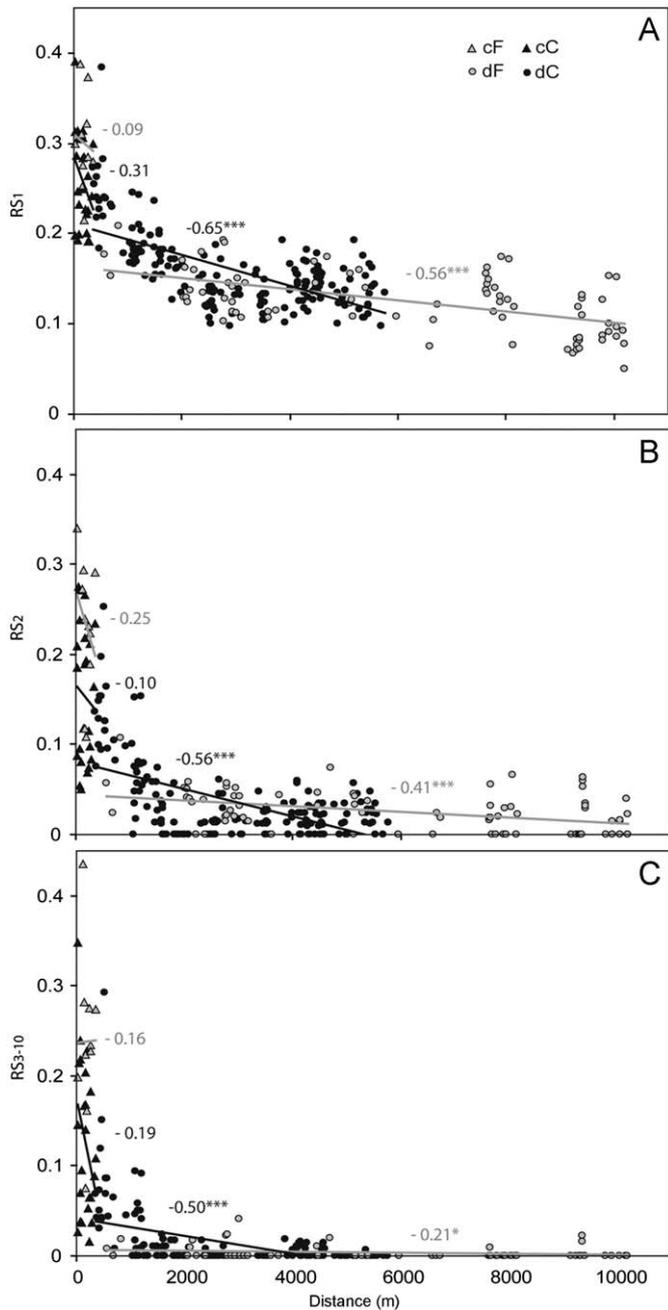
Fig. 6. Correlation between syllables and sequences shared by close and distant individuals and geographical distances separating these individuals. RS_1 (A graph), RS_2 (B graph) and RS_{3-10} (C graph) calculated between pairs of close individuals (cF) and distant individuals (dF) in the fragmented habitat (in grey) and between pairs of close individuals (cC) and distant individuals (dC) in the continuous habitat (in black) plotted as a function of geographical distances between individuals. Pearson correlation coefficients calculated for close and distant individuals with a permuted correlation test between RS values and distances are indicated above the regression lines for the fragmented habitat (grey) and continuous habitat (black). When correlations were significant, P values are indicated as follows: $^*P < 0.05$, $^{***}P < 0.001$.

2005). The song learning mechanism of the skylark has not been studied yet, and this would be a productive avenue for future research.

Our results do not seem support the prediction that song complexity may be influenced by habitat structure. Indeed, comparisons between songs produced by males living in either a continuous or a fragmented habitat revealed no differences in song complexity. The estimated breeding population density differed between the two habitats, with a lower density in the fragmented habitat (24.2 pairs/km^2) than in the continuous one (43–47 pairs/km^2). In some other bird species, repertoire size was shown to be correlated with population density (e.g. Bitterbaum and Baptista, 1979), and most of the time, a noticeable simplicity of songs was found in populations living on islands or isolated by unsuitable habitat (e.g. Naugler and Smith, 1991; Laiolo and Tella, 2007; review in Baker et al., 2006). Baker and Jenkins (1987) and Baker et al. (2006) proposed that such reduction of song complexity could be the consequence of a founder effect. This effect arises when a few colonists to an isolated area carry a small fraction of the mainland vocal diversity and establish an impoverished vocal tradition. In some other bird species, the opposite has been found, with the overall diversity of songs being larger in isolated populations (Schottler, 1995; Osiejuk et al., 2003). However, in the skylark, habitat fragmentation or density does not seem to have an effect on repertoire size or song complexity.

5.2. Song sharing

As predicted, our results show distinct patterns of song sharing depending on the kind of element considered and on the kind of habitat. First, at the level of syllables, we observed a gradual decline in syllable sharing as a function of geographical distances in both habitats. Birds situated at close geographical range ('close individuals') shared a lot more syllables than individuals situated at more than 500 m apart ('distant individuals'). At comparable distances, close individuals in the fragmented habitat shared more syllables than close individuals in the continuous one. We could notice that identical syllables were produced by individuals separated by large distances—until more than 10 km of distance—and even by individuals of the French and Polish populations situated 1100 km apart. Such syllable types could be determined by genetic limitations or song learning 'predispositions' of the species. Second, the geographical variations were more obvious at the sequence level than at the syllable level. Like for syllables, close individuals in the fragmented habitat shared more sequences than close individuals in the continuous habitat, and in addition, the profile of the plot of sequence sharing as a function of geographical distances differed between the two kinds of habitats. In the continuous habitat, there was a rapid decline in sequence sharing over short distances. In the fragmented habitat, we could observe two distinct categories of sequence sharing values without intermediate values. Indeed, there was a high sequence sharing at close distance and a very low sequence sharing at long distance.

The dendrograms that illustrated spatial patterns of sequence sharing showed distinct clusters corresponding to groups of close individuals in the fragmented habitat but not in the continuous one. These discrete patches of sequence variability found in the fragmented habitat indicate that sequence sharing is not related to the distance separating individuals but to the fact that individuals belong to the same group or not. In a previous study carried out in a fragmented habitat, we found that these sequences shared by close individuals established in the same group, referred to as 'neighbours', were used by birds to discriminate a neighbour from another individual coming from a distant group, referred to as a 'stranger' (Briefer et al., 2008). These sequences thus serve as a basis for group recognition that underlies the dear-enemy effect (i.e. reduced aggression toward familiar individuals compared to unfamiliar ones).

Groups of neighbours are particular communication networks within which males hold adjacent territories and sing within
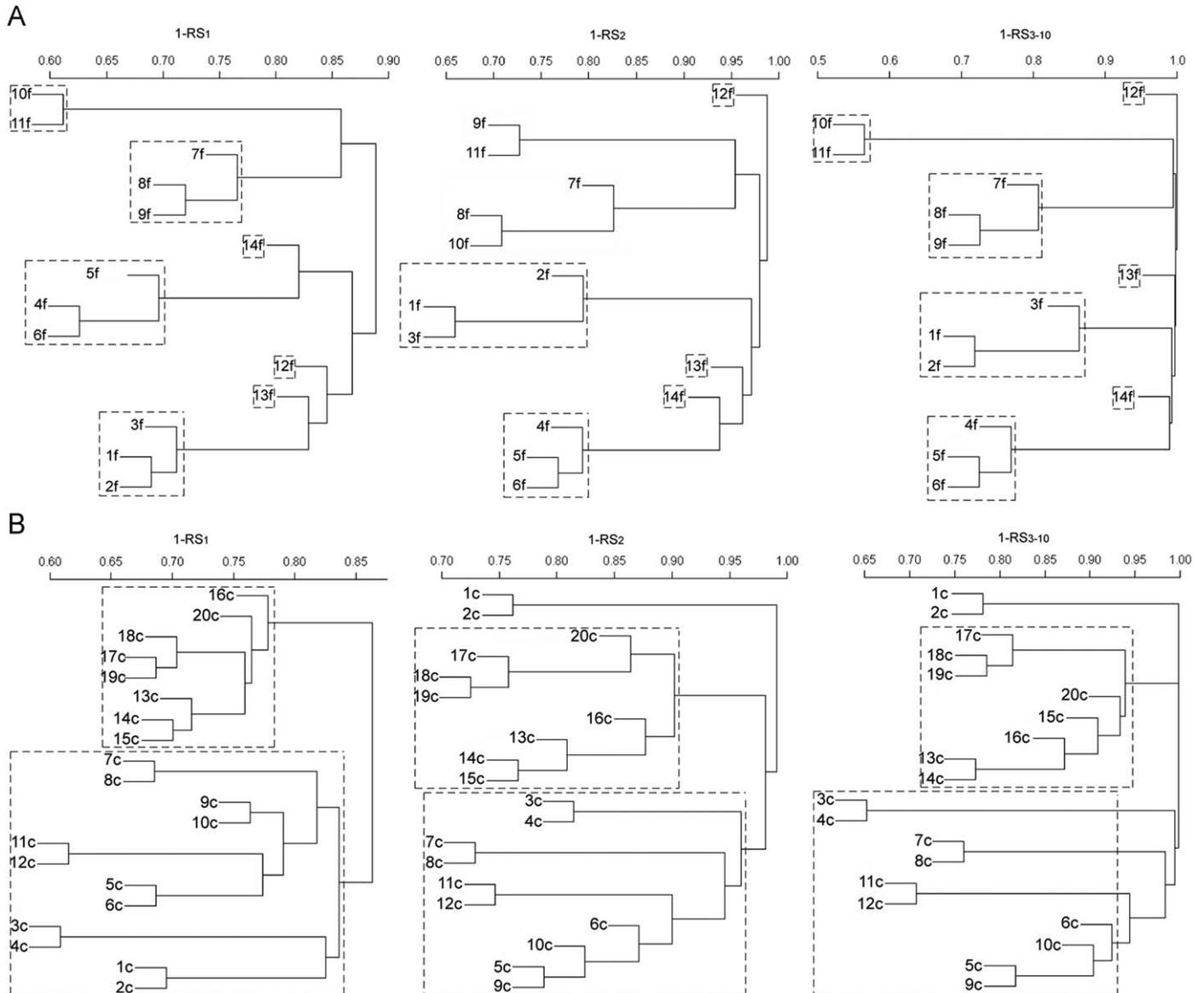
Fig. 7. Patterns of syllable and sequence sharing. Cluster dendrograms, for the fragmented habitat (A) and continuous habitat (B), obtained from hierarchical cluster analysis using the average agglomeration method on each matrix of $1 - RS_1$, $1 - RS_2$ and $1 - RS_{3-10}$ calculated between pairs of individuals. For the fragmented habitat, when clusters reflect geographical distances, dotted rectangles delimit the seven studied groups (1–3 individuals recorded per group) separated by unsuitable area. For the continuous habitat, dotted rectangles delimit the two sets of individuals (cf. Fig. 2).

signalling range of each other. Within such neighbourhoods, as within a continuous habitat in which males hold adjacent territories too, a stable spacing pattern between territories is essential to reduce the energetic costs of territorial defence (Naguib, 2005). In both habitats, song sharing could help to maintain a stable spacing pattern between territories by reducing aggressive interactions between neighbours (Baker et al., 1981). Indeed, in some species, territory tenure has been shown to be related to song sharing with neighbours (e.g. Beecher et al., 2000; Wilson et al., 2000). Selection pressure acting upon song structure may thus lead to an increase in song sharing between close birds.

As we do not have a replication of populations for each kind of habitat, and as our study compares two samples of individuals of limited size from different skylark populations (France and Poland), we cannot assert with certainty that the observed song differences between the two kinds of habitat are not the result of population or location differences (pseudo-replication, Hurlbert, 1984). However, although further studies using several popula-

tions could be useful to lend support to our results, the habitat fragmentation seemed here to induce on the one hand more sharing of syllables and sequences by close birds and on the other hand clear boundaries between sequence microdialects. Such enhanced song sharing between neighbouring males has been described in populations of birds living in isolated island sites (e.g. Morton, 1987; Lachlan and Slater, 2003) and in fragmented habitats (Laiolo and Tella, 2005). It has been hypothesized to be the result of a harsher competition for limited resources. However, in our study, population density was lower in the fragmented habitat than in the continuous one, and male–male competition can thus not explain our results.

The increases in song sharing between neighbouring skylarks may more likely be a side-effect of isolation associated with low population density and low levels of dispersal that reduce the number of song models available at hearing range from juveniles. Habitat fragmentation has been shown to block dispersal, and thus to induce a limited turnover during breeding and post-

Table 5
Effect of the kind of habitat on syllable and sequence sharing.

Habitat	RS	Effect	Df	F	P
Fragmented	RS_1	Distance	88	28.56	***
		Proximity		121.93	***
	RS_2	Distance		7.44	**
		Proximity		224.77	***
	RS_{3-10}	Distance		0.23	NS
		Proximity		360.66	***
Continuous	RS_1	Distance	187	66.92	***
		Proximity		41.78	***
	RS_2	Distance		29.29	***
		Proximity		70.73	***
	RS_{3-10}	Distance		14.78	***
		Proximity		58.04	***
Both habitats	RS_1	Distance	276	75.07	***
		Proximity		142.90	***
		Habitat		10.59	**
		Interaction		13.04	***
	RS_2	Distance		25.61	***
		Proximity		224.41	***
		Habitat		39.36	***
		Interaction		21.23	***
	RS_{3-10}	Distance		6.27	*
		Proximity		222.16	***
		Habitat		44.72	***
		Interaction		40.00	***

Abbreviations for P values: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, NS=non-significant. Results, within each habitat (fragmented and continuous) separately and in both kinds of habitats, of two-way analyses of covariance (ANCOVAs) on coefficients of repertoire similarity (RS_1, RS_2 and RS_{3-10}), with RS as a dependant variable, the distance between birds as a covariate, the proximity between birds (close vs. distant) as a fixed effect, and, for the ANCOVA carried out on both habitat, the kind of habitat as another fixed effect. For this ANCOVA, the interaction between the two fixed effects (the proximity between birds and the kind of habitat) is stated.

breeding seasons (Laiolo and Tella, 2005, 2008). Thus, as site fidelity is very high within and between breeding seasons (Jenny, 1990; Delius, 1963) and as males are strongly confined inside their territories (Aubin, 1981), the number of models to imitate when juveniles learn their songs may be reduced by habitat fragmentation. A reduction in population size and density, as suggested by Baker and Jenkins (1987) and Lynch (1996), may shape song sharing patterns and induce a kind of 'cultural consanguinity'.

Analyses of spatial patterns of syllable and of sequence sharing using dendrograms showed that these patterns differed. Indeed, dendrograms of syllables and of sequences clustered individuals in a different way. This could indicate that syllables and sequences do not result from the same learning process. The pattern of geographical variation is strongly influenced by dispersal and by the timing of learning. High level of sharing among immediate neighbours combined with a rapid decline in song sharing over short distances, as found for sequences in our study, may be the result of juvenile learning their song by imitation from their neighbours after dispersing. On the other hand, a gradual decline in sharing with distance, as found for syllables, occurs when juveniles learn their song before dispersal and move short distances to find a territory and settle preferentially near their tutors (Sokal and Neal, 1978; Goodfellow and Slater, 1986; Wilson et al., 2000; Diniz-Filho and Tellez, 2002; Koetz et al., 2007). The timing of song learning could thus differ between syllables and sequences, with syllables learned by juvenile in their birthplace before dispersal (which should occur at short distances), and phrases learned later by imitation of their neighbours. One year old male skylarks have indeed been shown to display regional philopatry (Delius, 1963).

To conclude, we found that the structure of the habitat influenced song sharing but not song complexity. Therefore, the environment could have an effect on songs at the level of sharing of elements (syllables and sequences of syllables), but not on the number and the complexity of these elements that seems to be a fixed characteristic of skylark songs. Thus, habitat fragmentation caused by anthropogenic barriers seems to induce an increase in the amount of song sharing to the detriment of variety among individuals.

References

Andrén, H., 1994. Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat: a review. Oikos 71, 355–366.

Aubin, T., 1981. Etude Expérimentale du Chant Territorial de l'Alouette des Champs (Alauda arvensis L.). Caractéristiques physiques, valeur sémantique et spécificité. Ph.D. thesis, University of Besançon, Nancy 1, Strasbourg.

Aubin, T., 1982. Habituation au chant territorial chez l'alouette des champs (Alauda arvensis L.); rôle de la diversité et de la monotonie. Biol. Behav. 7, 353–362.

Aubin, T., Brémond, J.-C., 1983. The process of species-specific song recognition in the skylark Alauda arvensis. An experimental study by means of synthesis. Z. Tierpsychol. 61, 141–152.

Baker, A.J., Jenkins, P.F., 1987. Founder effect and cultural evolution of songs in an isolated population of chaffinches, Fringilla coelebs, in the Chatham Islands. Anim. Behav. 35, 1793–1803.

Baker, M.C., Thompson, D.B., Sherman, G.L., Cunningham, M.A., 1981. The role of male interactions in maintaining population dialect structure. Behav. Ecol. Sociobiol. 8, 65–69.

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. Lin. Soc. 89, 331–342.

Becker, P.H., 1974. Der Gesang von Winter und Sommergoldhähnchen (Regulus regulus, Regulus ignicapillus) am westlichen. Bodensee. Vogelwarte 27, 233–243.

Beecher, M.D., Campbell, S.E., Nordby, J.C., 2000. Territory tenure in song sparrows is related to song sharing with neighbors, but not to repertoire size. Anim. Behav. 59, 29–37.

Beecher, M.D., Brenowitz, E.A., 2005. Functional aspects of song learning in songbirds. Trends Ecol. Evol. 20, 143–149.

Bender, D.J., Contreras, T.A., Fahrig, L., 1998. Habitat loss and population decline: a meta-analysis of the patch size effect. Ecology 79, 517–533.

BirdLife International, 2008. Species factsheet: Alauda arvensis. Downloaded from ⟨http://www.birdlife.org⟩ on 18/6/2008.

Bitterbaum, E., Baptista, L.F., 1979. Geographical variation in songs of Carolina House Finches (Carpodacus mexicanus). Auk 96, 462–474.

Briefer, E., Aubin, T., Lehongre, K., Rybak, F., 2008. How to identify dear-enemies: the group signature in the complex song of the skylark Alauda arvensis. J. Exp. Biol. 211, 317–326.

Buchanan, K.L., Spencer, K.A., Goldsmith, A.R., Catchpole, C.K., 2003. Song is an honest signal of past developmental stress in the European starling (Sturnus vulgaris). Proc. R. Soc. B 270, 1149–1156.

Busche, G., 1989. Drastische Bestandseinbussen des Feldlerche Alauda arvensis auf Grünlandflachen in Schleswig-Holstein. Vogelwelt 110, 51–59.

Catchpole, C.K., 1980. Sexual selection and the evolution of complex songs among European warblers of the genus Acrocephalus. Behaviour 74, 149–166.

Catchpole, C.K., Slater, P.J. (Eds.), 1995. Bird Song: Biological Themes and Variations. Cambridge University Press, Cambridge, pp. 45–69 131–137; 196–218.

Chamberlain, D.E., Gregory, R.D., 1999. Coarse and fine scale habitat associations of breeding skylarks Alauda arvensis in the UK. Bird Study 46, 34–47.

Chessel, D., Dufour, A.-B., Thioulouse, J., 2004. The ade4 package-I-One-table methods. R. News 4, 5–10.

Cytel Software Corp., 1997. StatXact v3.1. Cytel Software Corporation, Cambridge.

Da Silva, M.L., Piqueira, J.R.C., Vielliard, J.M.E., 2000. Using Shannon entropy on measuring the individual variability in the rufous-bellied thrush Turdus rufiventris vocal communication. J. Theor. Biol. 207, 57–64.

Debinski, D.M., Holt, R.D., 2000. A survey and overview of habitat fragmentation experiments. Conser. Biol. 14, 342–355.

Delius, J.D., 1963. Das Verhalten der Feldlerche. Z. Tierpsychol. 20, 297–348.

Diniz-Filho, J.A.F., Tellez, M.P., 2002. Spatial autocorrelation analysis and the identification of operational units for conservation in continuous populations. Conserv. Biol. 16, 924–935.

Dobson, C.W., Lemon, R.E., 1978. Markov sequences in songs of American thrushes. Behaviour 68, 86–104.

Donald, P.F., 2004. Song and song flight. In: The Skylark. T and AD Poyser, London, pp. 72–88.

Eraud, C., Boutin, J.-M., 2001. Ecologie de l'alouette des champs (Alauda arvensis) et perspectives de conservation des populations. In: Office National de la Chasse et de la Faune Sauvage (Ed.), Rapport Scientifique 2000. Office National de la Chasse et de la Faune Sauvage, Paris, pp. 51–53.

Ficken, M.S., Hailman, E.D., Hailman, J.P., 1994. The chick-a-dee call system of the Mexican chickadee. Condor 96, 70–82.

Gentner, T.Q., Hulse, S.H., 1998. Perceptual mechanisms for individual vocal recognition in European starlings, Sturnus vulgaris. Anim. Behav. 56, 579–594.

Goodfellow, D.J., Slater, P.J.B., 1986. A model of bird song dialects. Anim. Behav. 34, 1579–1580.

Griessmann, B., Naguib, M., 2002. Song sharing in neighboring and non-neighboring thrush nightingales (Luscina luscina) and its implications for communication. Ethology 108, 377–387.

Hailman, J.P., Ficken, J.P., Ficken, R.W., 1985. The 'chick-a-dee' calls of Parus articapillus: a recombinant system of animal communication compared with written English. Semiotica 56, 191–224.

Hausberger, M., 1997. Social influences on song acquisition and sharing in the European starling (Sturnus vulgaris). In: Snowdon, C.T., Hausberger, M. (Eds.), Social Influences on Vocal Development. Cambridge University Press, Cambridge, pp. 128–156.

Hedenström, A., 1995. Song flight performance in the skylark Alauda arvensis. J. Avian Biol. 26, 337–342.

Holland, J., McGregor, P.K., Rowe, C.L., 1996. Changes in microgeographic song variation in the corn bunting Miliaria calandra. J. Avian Biol. 27, 1059–1061.

Howard, R.D., 1974. The influence of a sexual selection and interspecific competition on mokingbirds song (Mimus polyglottos). Evolution 28, 428–438.

Hultsch, H., Todt, D., 1981. Repertoire sharing and song post distance in nightingales. Behav. Ecol. Sociobiol. 8, 182–188.

Hurlbert, S.H., 1984. Pseudoreplication and the design of ecological field experiments. Ecol. Mono. 54, 187–211.

Ince, S.A., Slater, P.J.B., 1985. Versatility and continuity in the songs of thrushes Turdus spp. Ibis 127, 355–364.

Isaac, D., Marler, P., 1963. Ordering of sequences of singing behaviour of mistle thrushes in relationship to timing. Anim. Behav. 11, 179–188.

Jenny, M., 1990. Populationsdynamik der Feldlerche Alauda arvensis in einer intensive genutzten Agrarlandschaft des schweizerischen Mittellandes. Orn. Beob. 87, 153–163.

Jurasinski, G., 2007. Simba: a collection of functions for similarity calculation of binary data. R Package Version 0.2-5.

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.

Koetz, A.H., Westcott, D.A., Congdon, B.C., 2007. Spatial pattern of song element sharing and its implications for song learning in the chowchilla, Orthonyx spaldingii. Anim. Behav. 74, 1019–1028.

Kreutzer, M., 1974. Stéréotypies et variations dans les chants de proclamation territoriale chez le Troglodyte (Troglodytes troglodytes). Rev. Comp. Anim. 8, 270–286.

Kroodsma, D.E., 1975. Song patterning in the rock wren. Condor 77, 294–303.

Kroodsma, D.E., 1978. Continuity and versatility in bird song: support for the monotony-threshold hypothesis. Nature 274, 681–683.

Kroodsma, D.E., 1981. Winter Wren singing behavior: a pinnacle of song complexity. Condor 82, 357–365.

Kroodsma, D.E., 1996. In: Kroodsma, D.E., Miller, E.H. (Eds.), Ecology and Evolution of Acoustic Communication in Birds. Cornell University Press, Ithaca, New York, pp. 3–19 125–147.

Lachlan, R.F., Slater, P.J.B., 2003. Song learning by chaffinches: how accurately and from where? A simulation analysis of patterns of geographical variation. Anim. Behav. 65, 957–969.

Laiolo, P., Tella, J.L., 2005. Habitat fragmentation affects culture transmission: patterns of song matching in Dupont's lark. J. Appl. Ecol. 42, 1183–1193.

Laiolo, P., Tella, J.L., 2007. Erosion of animal cultures in fragmented landscapes. Front. Ecol. Env. 5, 68–72.

Laiolo, P., Vögeli, M., Serrano, D., Tella, J.L., 2008. Song diversity predicts the viability of fragmented bird populations. Plos-One 3, e1822.

Lambrechts, M.M., Dhondt, A.A., 1995. Individual voice discrimination in birds. Curr. Ornithol. 12, 115–139.

Lampe, H.M., Espmark, Y.E., 1994. Song structure reflects male quality in pied flycatchers (Ficedula hypoleuca). Anim. Behav. 47, 869–876.

Lehongre, K., Aubin, T., Robin, S., Del Negro, C., 2008. Individual signature in canary songs: contribution of multiple levels of song structure. Ethology 114, 425–435.

Leitner, S., Voigt, C., Gahr, M., 2001. Seasonal changes in the song pattern of the non-domesticated island canary (Serinus canaria), a field study. Behaviour 138, 885–904.

Lemon, R.E., Chatfield, C., 1971. Organization of song in Cardinals. Anim. Behav. 19, 1–17.

Lemon, R.E., Chatfield, C., 1973. Organization of song of Rose-breasted Grosbeaks. Anim. Behav. 21, 28–44.

Lynch, A., 1996. The population memetics of birdsong. In: Kroodsma, D.E., Miller, E.H. (Eds.), Ecology and Evolution of Acoustic Communication in Birds. Cornell University Press, Ithaca, New York, pp. 181–197.

Loffredo, C.A., Borgia, G., 1986. Male courtship vocalization as cues for mate choice in the satin bowerbird, Ptilonorhynchus violaceaus. Auk 103, 189–195.

Mandelbrot, B., 1953. Contribution à la théorie mathématique des jeux de communication. Publ. L'inst. Stat. L'Univ. Paris 2, 5–50.

Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27, 209–220.

Marler, P., Sherman, V., 1985. Innate differences in singing behavior of sparrows reared in isolation from adult conspecific song. Anim. Behav. 33, 57–71.

McCowan, B., Doyle, L.R., Hanser, S.F., 2002. Using information theory to assess the diversity, complexity, and development of communication repertoires. J. Comp. Psychol. 116, 166–172.

Morton, E.S., 1987. The effect of distance and isolation on song-type sharing in the Carolina wren. Wilson Bull. 99, 601–610.

Mulligan, J.A., 1966. Singing behavior and its development in the song sparrow, Melospiza melodia. Univ. Calif. Publ. Zool. 81, 1–76.

Mundinger, P.C., 1982. Microgeographic and macrogeographic variation in the acquired vocalizations of birds. In: Kroodsma, D.E., Miller, E.H. (Eds.), Acoustic Communication in Birds II. Academic Press, New York, pp. 147–208.

Mundry, R., 1999. Testing related samples with missing values: a permutation approach. Anim. Behav. 58, 1143–1153.

Naguib, M., 2005. Singing interactions in songbirds: implications for social relations and territorial settlement. In: McGregor, P. (Ed.), Animal Communication Networks. Cambridge University Press, Cambridge, pp. 300–319.

Naugler, C.T., Smith, P.C., 1991. Song similarity in an isolated population of fox sparrow (Passerella iliaca). Condor 93, 1001–1003.

Nordby, J.C., Campbell, S.E., Beecher, M.D., 1999. Ecological correlates of song learning in song sparrows. Behav. Ecol. 10, 287–297.

Nottebohm, F., 1969. The song of the chingolo (Zonotrichia capensi) in Argentina: description and evaluation of a system of dialects. Condor 71, 299–315.

Nottebohm, F., Nottebohm, M.E., 1978. Relationship between song repertoire and age in the canary, Serinus canarius. Z. Tierpsychol. 46, 298–305.

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.

Odderskaer, P., Prang, A., Poulsen, J.G., Andersen, P.N., Elmegaard, N., 1997. Skylarks (Alauda arvensis) utilisation of micro-habitats in spring barley fields. Agric. Ecos. Env. 62, 21–29.

Opdam, P., Foppen, R., Reijnen, R., Schotman, A., 1994. The landscape ecological approach in bird conservation: integrating the metapopulation concept into spatial planning. Ibis 137, 139–146.

Osiejuk, T.S., Ratyńska, K., Cygan, J.P., Dale, S., 2003. Song structure and repertoire variation in Ortolan Bunting (Emberiza hortulana L.) from isolated Norwegian population. Ann. Zool. Fen. 40, 3–16.

R Development Core Team, 2007. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, URL ⟨http://www.R-project.org⟩.

Rossi, J.P., 1996. Statistical tool for soil biology. 11. Autocorrelogram and Mantel test. Eur. J. Soil Biol. 32, 195–203.

Saunders, D.A., Hobbs, R.J., Margules, C.R., 1991. Biological consequences of ecosystem fragmentation: a review. Conserv. Biol. 5, 18–32.

Schmiegelow, F.K.A., Monkkonen, M., 2002. Habitat loss and fragmentation in dynamic landscapes: avian perspectives from the boreal forest. Ecol. Appl. 12, 375–389.

Schottler, B., 1995. Songs of blue tits Parus caeruleus palmensis from las Palmas (Canary Islands)—a test of hypotheses. Bioacoustics 6, 135–152.

Shannon, C., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27, 379–423 623–656.

Shannon, C., Weaver, W., 1949. The Mathematical Theory of Communication. University of Illinois Press, Chicago.

Slater, P.J.B., 1989. Bird song learning: causes and consequences. Ethol. Ecol. Evol. 1, 19–43.

Sokal, R.R., Neal, O., 1978. Spatial autocorrelation in biology. II. Some implications and four applications of evolutionary interest. Biol. J. Lin. Soc. 10, 229–249.

Sokal, R.R., Rohlf, F.J., 1995. In: Freeman, W.H. (Ed.), Biometry III, New York.

Specht, R., 2004. Avisoft-SASLab pro v4.31. Avisoft, Berlin.

StatSoft, Inc., 2001. Statistica v6.0. for Windows (Computer program manual). StatSoft, Inc., Tulsa, Oklahoma.

Stawarczyk, T., 2004. Barycz Valley. In: Sidlo, P.O., Blaszkowska, B., Chylarecki, P. (Eds.), Important Bird Areas of European Union Importance in Poland. Polish Society for the Protection of Birds, BirdLife partner in Poland, Warszawa, pp. 432–434.

Stoddard, P.K., 1996. Vocal recognition of neighbors by territorial Passerines. In: Kroodsma, D.E., Miller, E.H. (Eds.), Ecology and Evolution of Acoustic Communication in Birds. Cornell University Press, Ithaca, New York, pp. 356–374.

Suggs, D.N., Simmons, A.M., 2005. Information theory analysis of patterns of modulation in the advertisement call of the male bullfrog, Rana castebeiana. J. Acoust. Soc. Am. 117, 2330–2337.

Suzuki, R., Buck, J.R., Tyack, P.L., 2005. Information entropy of humpback whale songs. J. Acoust. Soc. Am. 119, 1849–1866.

Trainer, J.M., 1983. Changes in song dialect distributions and microgeographic variation in song of white-crowned sparrows (Zonotrichia leucophrys nuttalli). Auk 100, 568–582.

Tryjanowski, P., 2000. Changes in breeding populations of some farmland birds in W Poland in relation to changes in crop structure, weather conditions and number of predators. Folia Zool. 49, 305–315.

Tryjanowski, P., 2004. Skylark—*Alauda arvensis*. In: Sikora, A., Rohde, Z., Gromadzki, M., Neubauer, G., Chylarecki, P. (Eds.), The Atlas of Breeding Birds in Poland 1985–2004. Bogucki Wydawnictwo Naukowe, Poznań, pp. 318–319.

Wildenthal, J.L., 1965. Structure in primary song of the Mockingbird (*Mimus polyglottos*). Auk 82, 161–189.

Wilson, P.L., Towner, M.C., Vehrencamp, S.L., 2000. Survival and song-type sharing in a sedentary subspecies of the song sparrow. Condor 102, 355–363.

Zipf, G.K., 1949. Human Behavior and the Principle of Least Effort. Addison-Wesley Press, Cambridge, MA.