



HAL
open science

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

Elodie Briefer, Tomasz S. Osiejuk, Fanny Rybak, Thierry Aubin

► To cite this version:

Elodie Briefer, Tomasz S. Osiejuk, Fanny Rybak, Thierry Aubin. Are bird song complexity and song sharing shaped by habitat structure? An information theory and statistical approach. *Journal of Theoretical Biology*, 2009, 262 (1), pp.151. 10.1016/j.jtbi.2009.09.020 . hal-00559159

HAL Id: hal-00559159

<https://hal.science/hal-00559159v1>

Submitted on 25 Jan 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Author's Accepted Manuscript

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

Elodie Briefer, Tomasz S. Osiejuk, Fanny Rybak, Thierry Aubin

PII: S0022-5193(09)00440-8
DOI: doi:10.1016/j.jtbi.2009.09.020
Reference: YJTBI5711



www.elsevier.com/locate/jtbi

To appear in: *Journal of Theoretical Biology*

Received date: 6 November 2008
Revised date: 5 September 2009
Accepted date: 11 September 2009

Cite this article as: Elodie Briefer, Tomasz S. Osiejuk, Fanny Rybak and Thierry Aubin, Are bird song complexity and song sharing shaped by habitat structure? An information theory and statistical approach, *Journal of Theoretical Biology*, doi:[10.1016/j.jtbi.2009.09.020](https://doi.org/10.1016/j.jtbi.2009.09.020)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

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

3
4
5
6
7 **Elodie Briefer^{a,1,*}, Tomasz S. Osiejuk^b, Fanny Rybak^a, Thierry Aubin^a**

8
9 ^a University Paris 11, NAMC, CNRS-UMR8620, Orsay, France

10 ^b Department of Behavioural Ecology, Institute of Environmental Biology, Faculty of
11 Biology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznań, Poland

12 ¹ Present address: Queen Mary University of London, School of Biological and Chemical
13 Sciences, Mile End Road, London E1 4NS, UK

14 * Corresponding author: E. Briefer (e.briefer@qmul.ac.uk)

15

15 **Abstract**

16

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

32

33

34 **Keywords:** *entropy, Zipf's law, habitat fragmentation, bird song, skylark*

35

35 **1. Introduction**

36

37 Bird songs are among the most complex sounds produced by animals. These
38 elaborate signals of oscine birds (songbirds) are thought to have evolved largely through
39 sexual selection. The functions of songs are territorial defence and mate attraction.

40 Between species, songs show great variation: 1) at the level of song duration, from short
41 and discrete song types to long and continuous ones; 2) at the level of repertoire size,
42 from one element produced repeatedly to hundred of elements produced with immense
43 variety; 3) or even at the level of song sharing, from one song type shared by all
44 individuals of a given area to different sets of song types produced by each individual
45 (Catchpole and Slater, 1995).

46 The repertoire size of a bird is assessed as the number of different elements it can
47 produce. What constitutes an element is not always the same between species. As many
48 songbird species produce short and discrete songs (discontinuous singers), repertoire size
49 is usually measured as the number of different song types produced. However, it can also
50 be measured as the number of different syllables produced, especially for species
51 producing continuous songs (continuous singers, Catchpole and Slater, 1995). Despite the
52 fact that approximately 80% of all songbird species have small or moderate song
53 repertoire sizes (1-25 different song types per individual), most of them produce more
54 than one song type (reviewed by Lambrechts and Dhondt, 1995; Stoddard, 1996).
55 Various modes of repertoire acquisition exist: repertoire size appears to be partially under
56 genetic control, at least in some species (Marler and Sherman, 1985), and may depend on
57 condition during development of young males (Nowicki et al., 2000; Buchanan et al.,

58 2003) or on adult body condition (Lampe and Espmark, 1994; Kipper et al., 2006).
59 Repertoire appears to vary with age (Nottebohm and Nottebohm, 1978; Loffredo and
60 Borgia, 1986), geography (Mulligan, 1966; Kroodsma, 1981), and population size or
61 density (Bitterbaum and Baptista, 1979; Laiolo et al., 2008). Furthermore, the
62 composition of a repertoire is based on songs of multiple tutors and is influenced by
63 social interactions (Nordby et al., 1999).

64 The phenomenon of song sharing occurs when songs of neighbouring males are
65 more similar to one another than to those of non-neighbouring males. Variation in song
66 composition resulting from song sharing and occurring between close groups of birds is
67 called local dialect or microdialect and is often expressed in the syntactic organization of
68 songs (Mundinger, 1982). Whole songs (e.g. Griessmann and Naguib, 2002), individual
69 syllables within songs (e.g. Isaac and Marler, 1963; Kreutzer, 1974) or groups of linked
70 syllables (e.g. Becker, 1974) are shared by individuals from the same population. Song
71 sharing is thought to be a consequence of the cultural transmission of song through
72 learning (Trainer, 1983; Kroodsma, 1996). Spatial patterns of song sharing can be
73 influenced by the social structure (nest distribution), affiliative interactions and
74 population density (Hausberger, 1997), the specific learning strategy, the timing of
75 learning, the learning environment, migration and dispersal patterns (Nottebohm, 1969;
76 Goodfellow and Slater, 1986; Slater, 1989; Koetz et al., 2007).

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

81 One of the effects of human activities on environment is the fragmentation of
82 animal species' habitat in patches separated by relatively inhospitable terrain. There is
83 considerable evidence that habitat fragmentation can cause declines in avian populations
84 (Bender et al., 1998; Schmiegelow and Monkkonen, 2002). The main reasons of these
85 species' declines are a reduction of habitat areas (area effect) and an increase in distance
86 between remaining habitat patches that influences therefore dispersal ability (isolation
87 effect, Andrén, 1994; Opdam et al., 1994; reviewed in Saunders et al., 1991; Debinski
88 and Holt, 2000). Habitat fragmentation has been suggested to condition song variation in
89 some bird species (white crowned sparrow *Zonotrichia leucophrys nuttalli*, Trainer,
90 1983; corn bunting *Miliaria calandra*, Holland et al., 1996), but direct evidence has only
91 been shown in the Dupont's lark *Chersophilus duponti* (Laiolo and Tella, 2005, 2007).
92 In this species, song diversity has been shown to be an indicator of the viability of
93 fragmented bird populations. Song diversity could thus be used as a warning signal of
94 population decline (Laiolo et al., 2008).

95 To study the effect of habitat fragmentation on both song complexity and patterns
96 of song element sharing, we analysed the songs of skylarks, an ideal model species
97 because males produce highly complex songs. The skylark is a territorial songbird of
98 open landscape. During the breeding season, from February to the end of July, pairs settle
99 in adjacent territories (Delius, 1963). During the entire breeding season, males fiercely
100 defend territories (Donald, 2004). To deter intruders, they produce a flight song (Delius,
101 1963; Hedenström, 1995) that reveals species identity via encoding of temporal
102 parameters (Aubin and Brémond, 1983). Unlike songs of many songbirds which are
103 relatively short and discontinuous (Catchpole and Slater, 1995) and which can be

104 categorized in discrete song types, the skylark song is long and continuous. It consists of
105 series of song units, named syllables, produced between 2 and 6 kHz. With an estimated
106 repertoire of more than 300 different syllables per individual, this song is one of the most
107 complex songs among songbirds, giving rise to a huge potential for variation at the
108 syntactic level (Aubin, 1981, 1982). Neighbouring males in a given breeding area share
109 numerous syllables and sequences of syllables that constitute a microdialect (Briefer et
110 al., 2008). Since the mid-1970s, the European skylark population underwent strong
111 declines. The major changes which have been implicated in the decline of skylarks are
112 agricultural intensification, increasing monoculture inducing a loss of structural diversity
113 of vegetation, elimination of lands not used for cultivation, use of pesticides and
114 fertilisers, and increasing mechanisation of cultivation (Busche, 1989; Odderskaer et al.
115 1997; Chamberlain and Gregory, 1999; Tryjanowski, 2000).

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

123 We developed a method for assessing song diversity. We measured song
124 complexity and assessed patterns of syllables and sequences of syllables shared by
125 individuals within each habitat. Song complexity was assessed by measuring repertoire
126 size and song versatility, and by using information theory concepts (Shannon and

127 Weaver, 1949). Patterns of syllable and sequence sharing were assessed using various
128 statistical tools. We then compared results obtained in the two kinds of habitat. We
129 expected the continuity of the habitat associated with a high population density to result
130 in songs being more varied and complex compared to the fragmented habitat, as a
131 consequence of more tutor songs available to young birds during song learning and of
132 stronger male-male competition.

133

134 **Theory**

135

136 *Information Theory*

137 Information theory was first developed by Shannon (1948; Shannon and Weaver,
138 1949) to encode signals that could efficiently transmit information across a variety of
139 noisy communication channels. The information theory method examines the amount of
140 information, defined as a measure of one's freedom of choice when selecting a message,
141 at increasingly complex levels of signalling organisation of a communication system. The
142 amount of information in a communication system is measured using entropy, also called
143 uncertainty, and represents the degree of randomness in the system. According to the
144 level of signalling organisation, several values of entropy can be measured. *Maximum*
145 *entropy* or *zero-order entropy* assesses repertoire diversity. *First-order entropy* represents
146 the simple repertoire internal organisation. *Higher-order entropies* assess the
147 communication system complexity.

148

149 *Markov chain process*

150 Markov chain analyses are used to determine if events occurring in a sequence are
151 independent of one another. However, a Markov process cannot depict the sequential
152 organisation of these events. The alternative is to use transition matrices (e.g. Isaac and
153 Marler, 1963; Lemon and Chatfield, 1971, 1973), or kinematic graphs (e.g. Ficken et al.,
154 1994) to study sequential organisation of syllables in bird songs. Yet, the skylark
155 repertoire is so large that these methods are not useful to describe songs of this species.
156 Instead, we investigated Markov chain analyses using an *information graph*. According
157 to this method, the amount of information contained in a sequence of elements is at
158 maximum if each element's probability in the sequence is completely independent of the
159 preceding element. Thus, by plotting the average amount of information (entropy) of
160 element sequences against sequence size, we obtain an *information graph*. If the
161 probability of any element is dependent on the element or elements immediately
162 preceding it, the amount of information will decrease as sequence size increases. Lemon
163 and Chatfield (1973) suggested that the portion of the *information graph* where
164 information declined most sharply (where there is a higher negative slope, called
165 'Shannon's slope') indicates the degree of Markov chaining.

166

167 *Zipf-Mandelbrot*

168 'Zipf's law' was first proposed by Zipf (1949) to describe human language. One
169 of his proposed linguistic generalizations was called 'principle of least effort'. According
170 to this law, if elements are ranked according to their probability of occurrence, from the
171 commonest to the rarest, there must be an inverse monotonic relationship between rank of
172 use and probability of occurrence. When plotted logarithmically, the function obtained

173 should be linear. This function gives an indication of the non-randomness and the
174 potential capacity for information transfer in a system. A regression coefficient (slope) of
175 -1 indicates that there is a balance between diversity and redundancy in the system and
176 thus, that the system has a high potential capacity for transferring information. However,
177 if the slope ranges between -1 and 0, there is more diversity than redundancy. If the slope
178 is smaller than -1, there is more redundancy than diversity in the system and thus, less
179 potential capacity for transferring information. Mandelbrot (1953) found a rationale for
180 Zipf's empirical generalization and developed a formula that could fit data from a greater
181 variety of languages. This formula is: $P_i = i (r + k)^{-s}$, with P_i the probability of
182 occurrence of elements, r their rank of use, i the intercept, k the curvature factor and s the
183 asymptotic slope. The resulting curve is more hyperbolic than Zipf's law at the upper data
184 (lowest rank elements) of Zipf's plot and has been shown to describe bird communication
185 system well (Hailman et al., 1985). Furthermore, Mandelbrot's curve gives an indication
186 about whether the system possesses generative properties, that is, whether the number of
187 different elements in the system is unlimited. To show openness of a communication
188 system, the plotted curve should reach a non-zero asymptotic slope (Hailman et al., 1985;
189 Ficken et al., 1994). Indeed, a non-zero asymptotic slope indicates that, as larger samples
190 of songs are recorded, new element types will be encountered in the population.

191

192 **2. Methods**

193

194 2.1. Study Areas, Subjects and Song Recordings

195

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

201 In France, songs were recorded during the breeding season in 2005 in the fields
202 surrounding the University of Paris 11, Orsay, Essone. The study plot covered 81 km²
203 and consisted of arable fields, most of them being cereals crops, surrounded by habitats
204 unsuitable for skylarks, like anthropogenic barriers (roads and houses) and few
205 woodlands that covered about 50% of the area (assessed using aerial pictures, Fig. 1).
206 Due to the heterogeneity of the landscape, this population was distributed in distinct
207 small groups separated by few kilometres. We studied 7 groups established in locations
208 of 0.44 ± 0.07 km² (mean \pm SE), containing 9.6 ± 0.7 individuals each, and separated by
209 unsuitable areas (Fig. 1; 2). Within these locations, the bird density was of 24.2
210 pairs/km². One to 3 individuals per groups were recorded, making a total of 14
211 individuals.

212 In Poland, songs were recorded during the breeding season in 2006 in the fields
213 around Odolanów, Wielkopolska. Large scale monitoring data for the whole country
214 indicate that skylark has a continuous distribution across Poland (Stawarczyk, 2004;
215 Tryjanowski, 2007). The study plot covered 24 km² and consisted of arable and pastoral
216 fields. Groups of tree and shrubs covered about 2% of the area and anthropogenic barriers
217 about 8% (assessed using aerial pictures, Fig. 1). Birds of this population were dispersed
218 in the habitat without unsuitable areas between them. The bird density within this study

219 area exceeded two times the bird density in the fragmented habitat (48.8 pairs/km²). To
220 allow comparison at similar distances with the French population, we recorded 20
221 individuals established in two sets that covered 3.1 km² and 2.7 km² respectively and that
222 were situated 2 km apart (Fig. 1; 2). The populations of France and Poland were situated
223 1100 km apart.

224 During the breeding season, site fidelity is very strong (Jenny, 1990; Delius,
225 1963). Once territories are established, boundaries between adjoining territories are stable
226 and males remain inside these boundaries (Aubin, 1981). We recorded songs at the
227 middle of the breeding season, in May, when territories are stable. To avoid individual
228 misidentification, birds established in neighbouring territories were recorded during a
229 single visit. For each recorded bird, we took GPS coordinates (Garmin GPSMAP 76S) at
230 the place where we had seen the bird singing. We then used these GPS coordinates to
231 calculate, in each kind of habitat, geographical distances between territories of all pairs of
232 recorded individuals. We calculated distances between 2 GPS coordinates in meters with
233 a calculator using a spherical earth assumption.

234 We recorded several songs per individual between 0700 and 1200 hours Eastern
235 Daylight Time using a Marantz PMD 690 digital recorder (sampling rate: 48000 Hz)
236 connected to a Sennheiser ME 64 K6 omni-directional microphone (frequency response:
237 30 Hz to 20 kHz \pm 1dB) mounted on a Telinga Universal parabola (diameter: 50 cm). We
238 then transferred songs files to a computer and high-pass filtered them (cut-off frequency:
239 1600 Hz) to remove the background noise. We used the Avisoft SASLab pro version 4.31
240 (Specht, 2004) for subsequent analyses.

241

242 2.2. Analysis of songs

243

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

249 A syllable was defined as a continuous trace on the sound spectrogram or a group
250 of continuous traces spaced out by less than 25 ms. As described elsewhere (Briefer et al.,
251 2008), we labelled syllables on the sound spectrogram with a number according to their
252 overall frequency modulation shapes. The same syllable found at several places in the
253 song of one individual or in songs of different individuals was labelled with the same
254 number. Then, we examined the sequential organization of syllables using a custom
255 Matlab program (The MathWorks, Natick, MA, USA; see Lehongre et al., 2008). The
256 numbers corresponding to the syllable labelling were inputted into the Matlab program in
257 the order of production of the corresponding syllables in the songs. For each individual,
258 the program classified sequences of numbers according to their length and the number of
259 times they were repeated. We thus obtained, for each individual, the repertoire of all
260 syllables and sequences of 2 to 10 syllables with the number of times they were repeated.

261

262 2.3. Song complexity

263

264 To assess the effect of the kind of habitat (continuous versus fragmented) on the
265 complexity of songs, we used the repertoires of syllables and of sequences of 2 to 10
266 syllables produced by each individual to calculate complexity and versatility of songs.
267 We then applied information theory concepts (Shannon and Weaver, 1949) to calculate
268 entropy. As skylarks do not sing discontinuous songs that can be classified like most
269 other passerine birds, syllable types and not song types were the basic units for our
270 measures.

271

272 *Versatility: variety of syllables and complexity of sequences within songs*

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

275

276 (a) *Repertoire size: syllable composition of songs*

277 We measured the *observed repertoire size* (**ObsR**) of each individual as the total
278 number of different syllables produced. As skylark repertoire size is huge (Briefer et al.,
279 2008), the *total repertoire size* (**R**) of each individual was estimated following the method
280 developed by Wildenthal (1965) and Howard (1974). This method assumes that the curve
281 of the number of new encountered syllable types plotted against the number of analysed
282 syllables rises to meet an asymptote (**R**) following an exponential equation. This method
283 is applied under the following assumptions: the probability of a new syllable type
284 appearing is proportional to the number of unused syllable types remaining in the
285 repertoire; the repertoire size is constant and fixed; and syllables are produced in random
286 sequences. When 300 s of skylark song were analysed, the obtained repertoire curves

287 indeed reached an asymptote (Briefer et al., 2008). We derived the more approximate
 288 exponential equation fitting skylark repertoire curves from Wildenthal (1965) equation as
 289 follows: $n = R(1 - e^{-(T/\alpha)/R})$, with n = the number of new encountered syllable types,
 290 T = the total number of analysed syllables, R = the total repertoire size, and $\alpha = 1.86$.
 291 The latter value was estimated from the curves of 9 individuals used in our previous study
 292 for which the total repertoire size (R) was studied by analysing more than 300 s of song.

293
 294 (b) *Syllable type versatility: variety of syllables within sequences*

295 We calculated *syllable type versatility* (**SylVer**) for each individual as the mean
 296 number of different syllables in sequences of 10 syllables. **SylVer** ranges from 0 to 10,
 297 with 10 being the maximum versatility.

298
 299 (c) *Transition versatility: repetition of syllables within sequences*

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

303
 304 (d) *Total versatility: variety and repetitions of syllables within sequences*

305 We calculated the *total versatility* (**TotVer**) for each individual as the mean value
 306 of **SylVer** * **TrVer** in sequences of 10 syllables. Total versatility ranges from 0 to 90, with
 307 90 being the maximum versatility.

308
 309 *Information analysis: amount of information in sequences within songs*

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

313

314 (a) *Zero-order entropy: syllable song diversity*

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

319

320 (b) *First-order entropy: simple internal organisation of songs*

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

326

327 (c) *Higher-order entropies: complex internal organisation of songs*

328 We then measured entropy values associated with each size of sequences
 329 composed of 2 to 10 syllables, by calculating *second-* (E_2) to *tenth-order* (E_{10}) *entropy*
 330 respectively. These *higher-order entropies* measure song complexity by assigning to each
 331 sequence of a given size a joint probability of occurrence and a conditional probability of
 332 occurrence in the song. The joint probability (P_{ijkl}) of a sequence type of 4 syllables (i-j-

333 k-l), for example, is equal to the number of times this sequence type is observed divided
 334 by the total number of sequences of 4 syllables in the song. The conditional probability
 335 ($P_{l/ijk}$) of this sequence type is the probability that (i-j-k-l) occurs in the song given that
 336 the sequence type of 3 syllables (i-j-k) just occurred. $P_{l/ijk}$ is calculated as the joint
 337 probability of the sequence type of 4 syllables (i-j-k-l) divided by the joint probability of
 338 the sequence type of 3 syllables (i-j-k). E_2 to E_{10} were calculated as follows: $E_n = -\sum P_{ij}$
 339 $\log_2(P_{ij})$, where P_{ij} is the joint probability, $P_{j/i}$ is the conditional probability, and n is the
 340 order of entropy. These entropy values measure the amount of information (freedom of
 341 choice) contained at increasing levels of song organisation.

342

343 (d) *Relative entropy and redundancy: proportion of free choice in building*
 344 *sequences*

345 Additionally, we calculated the relative entropy (**RE**) of songs, which is the ratio
 346 of the actual to the zero-order entropy: $RE_n = E_n / E_0$, with n being the order of entropy
 347 (1 to 10). **RE** is a measure of the proportion of free choice of syllables the sender has to
 348 build a sequence.

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

353

354 *Markov chain analysis: order of transition between syllables within sequences*

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

358

359 *Zipf-Mandelbrot: openness of the communication system*

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

363

364 *Statistical analyses*

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

368

369 2.4. Song sharing

370

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

374

375 *Syllables and sequences shared by individuals: coefficients of repertoire similarity*

376 We measured syllable and sequence sharing between each pair of individuals
377 within each habitat using coefficients of repertoire similarity (RS). Coefficient of syllable

378 repertoire similarity and coefficients of similarity between repertoires of each sequence
379 size (2 to 10 syllables) were calculated as follows:

380 $RS = Z / ((X + Y) - Z)$, with X and Y being the total number of syllables or
381 sequences produced by males x and y , and Z being the number of syllables or sequences
382 shared by males x and y (Hultsch and Todt, 1981). RS values range from 0 to 1, with 1
383 being maximum sharing.

384

385 *Statistical analyses*

386

387 (a) *Analysis in principal coordinates*

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

390

391 (b) *Hierarchical cluster analysis*

392 We carried out a hierarchical cluster analysis (HCA) using the average
393 agglomeration method. Before performing the HCA, RS values were converted to
394 distances by calculating $1 - RS$. Then, we made a HCA on each of the matrices of $1 - RS$
395 values calculated between repertoires of syllables and sequences of pairs of individuals.

396

397 (c) *Comparisons between syllables and sequences shared by close and distant*
398 *individuals in the two habitats*

399 Conventional parametric and non-parametric tests are not suitable for analyses in
400 which each individual is included several times in the different pair-wise comparisons

401 (Sokal and Rohlf, 1995; Mundry, 1999). Thus, we compared **RS** values calculated
402 between pairs of individuals situated at comparable geographic distances in the two kinds
403 of habitat with two-tailed exact permutation tests using Monte Carlo method. We
404 separated pairs of individuals in two classes according to the distances separating them:
405 pairs of birds situated at close geographical distance, corresponding to the within group
406 distance in the fragmented habitat, were referred to as ‘close individuals’; pairs of birds
407 situated at a greater distance were referred to as ‘distant individuals’. First, we made
408 comparisons within each habitat between **RS** values calculated for pairs of close and
409 distant individuals. Then, we made a comparison between **RS** values calculated for pairs
410 of close individuals in the fragmented habitat and in the continuous one. The same
411 comparison was made between **RS** values calculated for pairs of distant individuals.

412

413 (d) *Correlations between syllable and sequence sharing and geographical*
414 *distances*

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

423

424 (e) ANCOVA: effect of the habitat and of the proximity between birds on syllable
425 and sequence sharing

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

436

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

441

442 **3. Results**

443

444 Abbreviations for measured parameters are listed in Table 1.

445

446 3.1. Song complexity

447

448 *Versatility: variety of syllables and complexity of sequences within songs*

449 In both habitats, the mean estimated repertoire size per individual was larger than
 450 340 different syllables (Table 2). Mean values of **SylVer** and **TrVer** for each habitat were
 451 close to the maximum versatility (10 and 9 respectively), inducing high mean values of
 452 **TotVer** (Table 2). Mean **TrVer** values (**TrVer** = 7.3 for each habitat) indicate that more
 453 than 81% (7.3/9) of transitions were between different syllables. There was a high
 454 correlation between observed and estimated syllable repertoire curves: mean \pm SE =
 455 0.992 ± 0.002 (range: 0.949 - 0.999, $N = 34$ individuals, Fig. 3).

456 Comparisons between versatility values measured in the fragmented habitat and
 457 those values measured in the continuous one did not show any difference (independent
 458 samples t-test, $df = 32$: **ObsR**, $t = -0.96$, $P = 0.35$; **R**, $t = 0.15$, $P = 0.88$; **SylVer**, $t = -0.82$,
 459 $P = 0.42$; **TrVer** $t = -0.12$, $P = 0.91$; **TotVer** $t = -0.67$, $P = 0.51$).

460

461 *Information analysis: amount of information of sequences within songs*

462 The mean **EstE₀** value per individual for the 2 habitats was of 8.4 (Table 3). The
 463 mean **RE₁** value for the 2 habitats was of 0.94 (Table 3), indicating that males had 94% of
 464 free choice of syllables to build a sequence. The remaining fraction of the structure of the
 465 syllable sequence, which is determined not by the free choice of the bird (**RD₁**), was thus
 466 very low (mean **RD₁** values < 0.07 , Table 3). The fraction of free choice dropped at the
 467 level of sequences of two syllables (mean **RE₂** value for the 2 habitats = 0.13, Table 3).
 468 Entropy values did not differ between the fragmented and continuous habitat
 469 (independent samples t-test, $df = 32$: $P \geq 0.099$ for all values).

470

471 *Markov chain analysis: order of transition between syllables within sequences*

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

480

481 *Zipf-Mandelbrot: openness of the communication system*

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

486

487 3.2. Song sharing

488

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

493 repertoire similarity (RS_1), the coefficient of similarity between repertoires of sequences
494 of 2 syllables (RS_2), and the coefficient of similarity between repertoires of sequences of
495 3 to 10 syllables (RS_{3-10}). We calculated RS_{3-10} by first grouping for each individual the
496 repertoires of 3 to 10 syllables and then calculating coefficients of similarities between
497 pairs of individuals.

498

499 *Geographical distances between territories of individuals*

500 Geographical distances between territories of pairs of individuals ranged from
501 16.3 to 10195.8 m in the fragmented, and from 15.0 to 5766.8 m in the continuous habitat
502 (Fig. 2). In the fragmented habitat, within group distances ranged from 16.3 to 355.4 m,
503 and between group distances ranged from 579.2 to 10195.8 m. ‘Close individuals’ were
504 defined as birds separated by distances corresponding to the within group distances in the
505 fragmented habitat. We thus considered as ‘close individuals’ birds that were separated
506 by distances shorter than 355.4 m, and as ‘distant individuals’ birds that were separated
507 by greater distances.

508

509 *Syllables and sequences shared by individuals: coefficients of repertoire similarity*

510 (a) *Comparisons between syllables and sequences shared by close and distant*
511 *individuals in the two habitats*

512 In both habitats, pairs of close individuals shared many sequences of 2 to 10
513 syllables, whereas distant individuals shared almost no sequence (Table 4). To compare
514 habitats, RS calculated for pairs of distant individuals situated between 585.0 and
515 5445.4m in the continuous habitat and between 579.21 and 5401.86m in the fragmented

516 habitat were included in the permutation tests. All **RS** calculated between close
 517 individuals were included in the permutation tests. In both kinds of habitat, pairs of close
 518 individuals shared more syllables and sequences than pairs of distant individuals.
 519 However, two pairs of close individuals in the fragmented habitat shared more syllables
 520 and more sequences than two pairs of close individuals in the continuous habitat,
 521 although these individuals were situated at comparable distances from each other in the
 522 two kinds of habitat.

523 Indeed, for both habitats, **RS** between close individuals were greater than **RS**
 524 between distant individuals (Permutation tests using Monte Carlo method: Fragmented
 525 habitat, $N = 10$ pairs of close individuals and 37 pairs of distant individuals, $P < 0.0001$
 526 for **RS₁**, **RS₂** and **RS₃₋₁₀**; Continuous habitat, $N = 24$ pairs of close individuals and 145
 527 pairs of distant individuals, $P < 0.0001$ for **RS₁**, **RS₂** and **RS₃₋₁₀**). **RS₁**, **RS₂** and **RS₃₋₁₀**
 528 between close individuals were greater in the fragmented habitat compared to the
 529 continuous habitat (Permutation tests using Monte Carlo method, $N = 34$ pairs: **RS₁**, $P <$
 530 0.05 ; **RS₂**, $P < 0.01$; **RS₃₋₁₀**, $P < 0.001$). On the other hand, there was no differences
 531 between **RS** values calculated for pairs of distant individuals in the two kinds of habitat
 532 (Permutation tests using Monte Carlo method, $N = 182$ pairs: **RS₁**, $P = 0.09$; **RS₂**, $P =$
 533 0.12 ; **RS₃₋₁₀**, $P = 0.51$).

534

535 *(b) Correlations between syllable and sequence sharing and geographical*
 536 *distances*

537 In the continuous habitat, the amount of syllables and sequences shared by distant
 538 individuals decreased as the distance between their territories increased. In the

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

543 At more than 2000 m of distance, in both habitats, individuals shared almost no
 544 sequences of 3 to 10 syllables (RS_{3-10}), whereas some syllables (RS_1 around 0.1) and few
 545 sequences of 2 syllables (RS_2 around 0.03) were still shared at more than 10000 m.
 546 Results of permuted correlation tests showed that, in the continuous habitat, RS_1 , RS_2 and
 547 RS_{3-10} calculated between pairs of distant individuals were negatively correlated with
 548 geographical distances. In the fragmented habitat, RS_1 calculated between pairs of distant
 549 individuals was negatively correlated with the distance, but this was not the case for RS_2
 550 and RS_{3-10} (weak correlations, Pearson correlation coefficients between -0.5 and 0). In
 551 both habitats, RS_1 , RS_2 and RS_{3-10} calculated for pairs of close individuals were not
 552 significantly correlated with geographical distances (Fig. 6).

553

554 *(c) Spacing patterns of syllable and sequence sharing*

555 In both habitats, dendrograms of hierarchical cluster analyses showed differences
 556 between syllable and sequence sharing patterns (Fig. 7). The pattern of shared sequences
 557 of 2 syllables (dendrogram 1- RS_2) seemed to be intermediate between sharing patterns of
 558 syllables (dendrogram 1- RS_1) and of sequences of 3 to 10 syllables (dendrogram 1- RS_{3-10}).
 559 10), with some clusters of individuals being identical to the syllable sharing pattern (1-
 560 RS_1) and other clusters being identical to the sharing pattern of sequence of 3 to 10
 561 syllables (1- RS_{3-10}).

562 *Fragmented habitat*

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

567 In each of the dendrograms of 1-*RS*₁, 1-*RS*₂ and 1-*RS*₃₋₁₀, individuals were
568 clustered in groups reflecting geographical distances, except for two groups (7f - 8f - 9f
569 and 10f - 11f) that were different in 1-*RS*₂ (cf. Fig. 2). *RS* distance between groups was
570 greater in the dendrogram of 1-*RS*₃₋₁₀ than in the dendrogram of 1-*RS*₂ and than in the
571 dendrogram of 1-*RS*₁. In the dendrogram of 1-*RS*₁, individuals for which no other birds
572 of their group had been recorded (12f, 13f and 14f) were clustered with the group that
573 was at closest geographical distance (12f and 13f with the group 1f - 2f - 3f; 14f with the
574 group 4f - 5f - 6f). In the dendrograms of 1-*RS*₂ and 1-*RS*₃₋₁₀, they were excluded from
575 these clusters (Fig. 7).

576 *Continuous habitat*

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

581 In each of the dendrograms of 1-*RS*₁, 1-*RS*₂ and 1-*RS*₃₋₁₀, individuals were
582 clustered in two sets of individuals, reflecting geographical distances (cf. Fig. 2), except
583 two individuals (1c and 2c) that were isolated in the dendrograms of 1-*RS*₂ and 1-*RS*₃₋₁₀.
584 These two individuals were situated, within their set of individuals, at short geographical

585 distances from the other set. In each of these dendrograms, individuals were often
586 clustered two by two. Within these clusters, individuals were clustered with another
587 individual situated at short geographical distance, but not always with the closest
588 individual (e.g. 5c - 9c and 14c - 15c in 1-**RS**₂; 9c - 10c and 19c - 17c in 1-**RS**₃₋₁₀).
589 Furthermore, these clusters of two individuals were not always identical in dendrograms
590 1-**RS**₁, 1-**RS**₂ and 1-**RS**₃₋₁₀, except four of them (1c - 2c, 3c - 4c, 11c - 12c and 7c - 8c)
591 that were identical in all dendrograms and three of them (5c - 9c, 14c - 15c and 18c - 19c)
592 that were identical in two dendrograms. However, neither of the dendrograms of 1-**RS**₁,
593 1-**RS**₂ and 1-**RS**₃₋₁₀ allowed us to distinguish a discrete cluster of more than two
594 individuals (Fig. 7).

595

596 *(d) ANCOVA: effect of the habitat and of the proximity between birds on syllable*
597 *and sequence sharing*

598 Results of two-way ANCOVAs corroborated with the cluster dendrograms of **RS**
599 in showing that there was a continuous increase in acoustic dissimilarities with increasing
600 geographical distances in the continuous habitat, and discrete patches of acoustic
601 variability at the level of sequences (corresponding to the groups) in the fragmented
602 habitat (Table 5). In this kind of habitat, sequence sharing depended on the proximity
603 between birds (same group: close individuals; different group: distant individuals), more
604 than on the distance between them. Habitat fragmentation affected both syllable and
605 sequence sharing.

606 In both habitats, the proximity between birds (close vs. distant) had a significant
607 effect on **RS**₁, **RS**₂ and **RS**₃₋₁₀, which indicates that syllable and sequence sharing differed

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

619

620 *(e) Syllable sharing between the two populations*

621 RS_1 calculated between repertoires of all syllables encountered in the fragmented
622 habitat (French population) and in the continuous one (Polish population) was 0.32, RS_2
623 0.01 and RS_{3-10} 0. RS calculated between two individuals taken at random in each
624 population were 0.15 for RS_1 , 0.001 for RS_2 (one sequence of 2 syllables shared) and 0
625 for RS_{3-10} . Thus, the RS_1 value of 0.1 observed at more than 10000 m in Fig. 6 seemed to
626 persist over hundreds of kilometres.

627

628 **4. Discussion**

629

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

637

638 4.1. Song complexity

639

640 Our results show that skylark songs are among the most complex acoustic signals
641 compared to other songbird species. Indeed, the repertoire size of more than 300 different
642 syllables is huge compared to species with a large repertoire that has been estimated in
643 number of different syllables, like the thrush nightingales (49 basic song components,
644 Griessmann and Naguib, 2002), the sedge warbler *Acrocephalus schoenobaenus* (37-54
645 syllables, Catchpole, 1980), the canary *Serinus canaria* (81 syllables, Leitner et al.,
646 2001), the rock wren *Salpinctes obsoletus* (69-119 syllables, Kroodsma, 1975) or even
647 the mockingbird *Mimus polyglottos* (66-244 syllables, Wildenthal, 1965).

648 We found a substantial immediate variety in the emission of syllables, as 81% of
649 transitions (mean $TrVer$ for the 2 habitats) are between different syllables. The potential
650 amount of information of songs is thus particularly important at the level of syllables,
651 with 94% (mean RE_I for the 2 habitats) of free choice of syllable types. Skylark song can
652 also be considered as highly versatile with a total average versatility of 55, exceeding the

653 value of 50 suggested as a criterion for characterising songs as highly versatile by Ince
654 and Slater (1985). Furthermore, information measures showed that syllable repertoire
655 diversity (zero-order entropy: 7.4) and the amount of information contained at the level of
656 syllables within songs (first-order entropy: 7.0) were comparable to or higher than values
657 measured for the rufous-bellied thrush *Turdus rufiventris* (first order: 3.19, Da Silva et
658 al., 2000), the veery *H. fuscescens* (first order: 2), the hermit thrush *H. guttata* (first
659 order: 3.4), the wood thrush *Catharus mustelina* (first order: 3.5 - 4.5), the robin *T.*
660 *migratorius* (first order: 3.5 - 4) (Dobson and Lemon, 1978) or the European starling
661 *Sturnus vulgaris* (zero-order: 7 and first-order: 6, Gentner and Hulse, 1998). First-order
662 entropy values obtained for skylark songs are also higher than values found in non-bird
663 species vocalizations, like male bullfrog *Rana catesbeiana* advertisement calls (first
664 order: 1.31, Suggs and Simmons, 2005), bottlenose dolphin *Tursiops truncatus* whistles
665 (first order: 1.92), squirrel monkey *Saimiri sciureus* chucks (first order: 2.21) (McCowan
666 et al., 2002), humpback whale *Megaptera novaeangliae* songs (Suzuki et al., 2005), or
667 even Russian, English or Arabic letters (first order: 4.03 – 4.35, McCowan et al., 2002).

668 The amount of information of skylark song drops at the level of sequences of two
669 syllables (13% of free choice of sequence types, mean RE_2 for the 2 habitats), indicating a
670 first order Markov chain, with each syllable being highly dependent on the syllable
671 immediately preceding, but not on other ones. Thus, it seems that the use of sequences of
672 two syllables is determined by more restrictive mechanical and/or biological constraints,
673 whereas the use of syllables and chaining of more than two syllables is determined by the
674 free choice of the bird.

675 Zipf-Mandelbrot plots showed that skylark communication system resembles a
676 'closed' system more than an 'open' system with generative properties: Mandelbrot
677 curves reached a zero asymptotic slope, indicating that the syllable repertoire of skylark
678 populations seems to be fixed in time. As the likelihood of genetic determination of a
679 whole syllable repertoire is unlikely in oscines, the existence of such a fixed repertoire
680 could be explained by a learning strategy based on strict imitation of tutors by young
681 birds, without improvisation. Indeed, in songbirds, individuals develop their songs by
682 imitating tutors' songs, by improvising on tutors' songs or by inventing songs with
683 minimal reference to external models (Beecher and Brenowitz, 2005). The song learning
684 mechanism of the skylark has not been studied yet, and this would be a productive
685 avenue for future research.

686 Our results do not seem support the prediction that song complexity may be
687 influenced by habitat structure. Indeed, comparisons between songs produced by males
688 living in either a continuous or a fragmented habitat revealed no differences in song
689 complexity. The estimated breeding population density differed between the two habitats,
690 with a lower density in the fragmented habitat (24.2 pairs/km²) than in the continuous one
691 (43-47 pairs/km²). In some other bird species, repertoire size was shown to be correlated
692 with population density (e.g. Bitterbaum and Baptista, 1979), and most of the time, a
693 noticeable simplicity of songs was found in populations living on islands or isolated by
694 unsuitable habitat (e.g. Naugler and Smith, 1991; Laiolo and Tella, 2007; review in Baker
695 et al., 2006). Baker and Jenkins (1987) and Baker et al. (2006) proposed that such
696 reduction of song complexity could be the consequence of a founder effect. This effect
697 arises when a few colonists to an isolated area carry a small fraction of the mainland

698 vocal diversity and establish an impoverished vocal tradition. In some other bird species,
699 the opposite has been found, with the overall diversity of songs being larger in isolated
700 populations (Schottler, 1995; Osiejuk et al., 2003). However, in the skylark, habitat
701 fragmentation or density does not seem to have an effect on repertoire size or song
702 complexity.

703

704 4.2. Song sharing

705

706 As predicted, our results show distinct patterns of song sharing depending on the
707 kind of element considered and on the kind of habitat. First, at the level of syllables, we
708 observed a gradual decline in syllable sharing as a function of geographical distances in
709 both habitats. Birds situated at close geographical range ('close individuals') shared a lot
710 more syllables than individuals situated at more than 500m apart ('distant individuals').
711 At comparable distances, close individuals in the fragmented habitat shared more
712 syllables than close individuals in the continuous one. We could notice that identical
713 syllables were produced by individuals separated by large distances - until more than
714 10km of distance - and even by individuals of the French and Polish populations situated
715 1100km apart. Such syllable types could be determined by genetic limitations or song
716 learning 'predispositions' of the species. Second, the geographical variations were more
717 obvious at the sequence level than at the syllable level. Like for syllables, close
718 individuals in the fragmented habitat shared more sequences than close individuals in the
719 continuous habitat, and in addition, the profile of the plot of sequence sharing as a
720 function of geographical distances differed between the two kinds of habitats. In the

721 continuous habitat, there was a rapid decline in sequence sharing over short distances. In
722 the fragmented habitat, we could observe two distinct categories of sequence sharing
723 values without intermediate values. Indeed, there was a high sequence sharing at close
724 distance and a very low sequence sharing at long distance.

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

736 Groups of neighbours are particular communication networks within which males
737 hold adjacent territories and sing within signalling range of each other. Within such
738 neighbourhoods, as within a continuous habitat in which males hold adjacent territories
739 too, a stable spacing pattern between territories is essential to reduce the energetic costs
740 of territorial defence (Naguib, 2005). In both habitats, song sharing could help to
741 maintain a stable spacing pattern between territories by reducing aggressive interactions
742 between neighbours (Baker et al., 1981). Indeed, in some species, territory tenure has
743 been shown to be related to song sharing with neighbours (e.g. Beecher et al., 2000;

744 Wilson et al., 2000). Selection pressure acting upon song structure may thus lead to an
745 increase in song sharing between close birds.

746 As we do not have a replication of populations for each kind of habitat, and as our
747 study compares two samples of individuals of limited size from different skylark
748 populations (France and Poland), we cannot assert with certainty that the observed song
749 differences between the two kinds of habitat are not the result of population or location
750 differences (pseudo-replication, Hurlbert, 1984). However, although further studies using
751 several populations could be useful to lend support to our results, the habitat
752 fragmentation seemed here to induce on the one hand more sharing of syllables and
753 sequences by close birds and on the other hand clear boundaries between sequence
754 microdialects. Such enhanced song sharing between neighbouring males has been
755 described in populations of birds living in isolated island sites (e.g. Morton, 1987;
756 Lachlan and Slater, 2003) and in fragmented habitats (Laiolo and Tella, 2005). It has
757 been hypothesized to be the result of a harsher competition for limited resources.
758 However, in our study, population density was lower in the fragmented habitat than in the
759 continuous one, and male-male competition can thus not explain our results.

760 The increases in song sharing between neighbouring skylarks may more likely be
761 a side-effect of isolation associated with low population density and low levels of
762 dispersal that reduce the number of song models available at hearing range from
763 juveniles. Habitat fragmentation has been shown to block dispersal, and thus to induce a
764 limited turnover during breeding and post-breeding seasons (Laiolo and Tella, 2005,
765 2008). Thus, as site fidelity is very high within and between breeding seasons (Jenny,
766 1990; Delius, 1963) and as males are strongly confined inside their territories (Aubin,

767 1981), the number of models to imitate when juveniles learn their songs may be reduced
768 by habitat fragmentation. A reduction in population size and density, as suggested by
769 Baker and Jenkins (1987) and Lynch (1996), may shape song sharing patterns and induce
770 a kind of “cultural consanguinity”.

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

787

788 To conclude, we found that the structure of the habitat influenced song sharing
789 but not song complexity. Therefore, the environment could have an effect on songs at the

790 level of sharing of elements (syllables and sequences of syllables), but not on the number
791 and the complexity of these elements that seems to be a fixed characteristic of skylark
792 songs. Thus, habitat fragmentation caused by anthropogenic barriers seems to induce an
793 increase in the amount of song sharing to the detriment of variety among individuals.

794

795 **Acknowledgments**

796

797 This study was supported by the CNRS and the University of Paris 11. EB was funded by
798 a grant from the French Minister of Research and Technology during data collection. We
799 thank Katia Lehongre for running the custom Matlab program and Selvino de Kort for
800 valuable comments on earlier version of the manuscript and linguistic correction. We are
801 very grateful to F. Stephen Dobson and Paola Laiolo for revision comments and helpful
802 suggestions.

803

804 **References**

805

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

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

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

- 813 Aubin, T., Brémond, J.-C., 1983. The process of species-specific song recognition in the
814 skylark *Alauda arvensis*. An experimental study by means of synthesis. *Z. Tierpsychol.*
815 61, 141-152.
- 816 Baker, A. J., Jenkins, P. F., 1987. Founder effect and cultural evolution of songs in an
817 isolated population of chaffinches, *Fringilla coelebs*, in the Chatham Islands. *Anim.*
818 *Behav.* 35, 1793-1803.
- 819 Baker, M. C., Thompson, D. B., Sherman, G. L., Cunningham, M. A., 1981. The role of
820 male interactions in maintaining population dialect structure. *Behav. Ecol. Sociobiol.* 8,
821 65-69.
- 822 Baker, M. C., Baker, M. S. A., Tilghman, L. M., 2006. Differing effects of isolation on
823 evolution of bird songs: examples from an island-mainland comparison of three species.
824 *Biol. J. Lin. Soc.* 89, 331-342.
- 825 Becker, P. H., 1974. Der Gesang von Winter und Sommergoldhähnchen (*Regulus*
826 *regulus*, *Regulus ignicapillus*) am westlichen Bodensee. *Vogelwarte* 27, 233-243.
- 827 Beecher, M. D., Campbell, S. E., Nordby, J. C., 2000. Territory tenure in song sparrows is
828 related to song sharing with neighbors, but not to repertoire size. *Anim. Behav.* 59, 29-37.
- 829 Beecher, M. D., Brenowitz, E. A., 2005. Functional aspects of song learning in songbirds.
830 *Trends Ecol. Evol.* 20, 143-149.
- 831 Bender, D. J., Contreras, T. A., Fahrig, L., 1998. Habitat loss and population decline: a
832 meta-analysis of the patch size effect. *Ecology* 79, 517-533.
- 833 BirdLife International 2008. Species factsheet: *Alauda arvensis*. Downloaded from
834 <http://www.birdlife.org> on 18/6/2008.

- 835 Bitterbaum, E., Baptista, L. F., 1979. Geographical variation in songs of Carolina House
836 Finches (*Carpodacus mexicanus*). *Auk* 96, 462-474.
- 837 Briefer, E., Aubin, T., Lehongre, K., Rybak, F., 2008. How to identify dear-enemies: the
838 group signature in the complex song of the skylark *Alauda arvensis*. *J. Exp.Biol.* 211,
839 317-326.
- 840 Buchanan, K. L., Spencer, K. A., Goldsmith, A. R., Catchpole, C. K., 2003. Song is an
841 honest signal of past developmental stress in the European starling (*Sturnus vulgaris*).
842 *Proc. R. Soc. B* 270, 1149-1156.
- 843 Busche, G., 1989. Drastische Bestandseinbussen des Feldlerche *Alauda arvensis* auf
844 Grünlandflächen in Schleswig-Holstein. *Vogelwelt* 110, 51-59.
- 845 Catchpole, C. K., 1980. Sexual selection and the evolution of complex songs among
846 European warblers of the genus *Acrocephalus*. *Behaviour* 74, 149-166.
- 847 Catchpole, C. K., Slater, P. J., (Eds.) 1995. Bird Song: Biological themes and variations.
848 Cambridge University Press, Cambridge, pp. 45-69; 131-137; 196-218.
- 849 Chamberlain, D. E., Gregory, R. D., 1999. Coarse and fine scale habitat associations of
850 breeding skylarks *Alauda arvensis* in the UK. *Bird Study* 46, 34-47.
- 851 Chessel, D., Dufour, A.-B., Thioulouse, J., 2004. The ade4 package-I- One-table
852 methods. *R News*. 4, 5-10.
- 853 Cytel Software Corp., 1997. StatXact v3.1. Cytel Software Corporation, Cambridge.
- 854 Da Silva, M. L., Piqueira, J. R. C., Vielliard, J. M. E., 2000. Using Shannon entropy on
855 measuring the individual variability in the rufous-bellied thrush *Turdus rufiventris* vocal
856 communication. *J. Theor. Biol.* 207, 57-64.

- 857 Debinski, D. M., Holt, R. D., 2000. A survey and overview of habitat fragmentation
858 experiments. *Conser. Biol.* 14, 342-355.
- 859 Delius, J. D., 1963. Das Verhalten der Feldlerche. *Z. Tierpsychol.* 20, 297-348.
- 860 Diniz-Filho, J. A. F., Tellez, M. P., 2002. Spatial autocorrelation analysis and the
861 identification of operational units for conservation in continuous populations. *Cons. Biol.*
862 16, 924-935.
- 863 Dobson, C. W., Lemon, R. E., 1978. Markov sequences in songs of American thrushes.
864 *Behaviour* 68, 86-104.
- 865 Donald, P. F., 2004. Song and song flight. In *The skylark*, pp. 72-88. London: T and AD
866 Poyser.
- 867 Eraud, C., Boutin, J.-M., 2001. Ecologie de l'alouette des champs (*Alauda arvensis*) et
868 perspectives de conservation des populations. In : Office National de la Chasse et de la
869 Faune Sauvage (Ed.), Rapport scientifique 2000. Office National de la Chasse et de la
870 Faune Sauvage, Paris, pp. 51-53.
- 871 Ficken, M. S., Hailman, E. D., Hailman, J. P., 1994. The chick-a-dee call system of the
872 Mexican chickadee. *Condor* 96, 70-82.
- 873 Gentner, T. Q., Hulse, S. H., 1998. Perceptual mechanisms for individual vocal
874 recognition in European starlings, *Sturnus vulgaris*. *Anim. Behav.* 56, 579-594.
- 875 Goodfellow, D. J., Slater, P. J. B., 1986. A model of bird song dialects. *Anim. Behav.* 34,
876 1579-1580.
- 877 Griessmann, B., Naguib, M., 2002. Song sharing in neighboring and non-neighboring
878 thrush nightingales (*Luscinia luscinia*) and its implications for communication. *Ethology*
879 108, 377-387.

- 880 Hailman, J. P., Ficken, J. P., Ficken, R. W., 1985. The 'chick-a-dee' calls of Parus
881 articipillus: A recombinant system of animal communication compared with written
882 English. *Semiotica* 56, 191-224.
- 883 Hausberger, M., 1997. Social influences on song acquisition and sharing in the European
884 starling (*Sturnus vulgaris*). In: Snowdon, C. T., Hausberger, M. (Eds.) Social Influences
885 on Vocal Development. Cambridge University Press, Cambridge, pp. 128-156.
- 886 Hedenström, A., 1995. Song flight performance in the skylark *Alauda arvensis*. *J. Avian*
887 *Biol.* 26, 337-342.
- 888 Holland, J., McGregor, P. K., Rowe, C. L., 1996. Changes in microgeographic song
889 variation of the corn bunting *Miliaria calandra*. *J. Avian Biol.* 27, 1059-1061.
- 890 Howard, R. D., 1974. The influence of a sexual selection and interspecific competition on
891 mockingbirds song (*Mimus polyglottos*). *Evolution* 28, 428-438.
- 892 Hultsch, H., Todt, D., 1981. Repertoire sharing and song post distance in nightingales.
893 *Behav. Ecol. Sociobiol.* 8, 182-188.
- 894 Hurlbert, S. H., 1984. Pseudoreplication and the design of ecological field experiments.
895 *Ecol. Mono.* 54, 187-211.
- 896 Ince, S. A., Slater, P. J. B., 1985. Versatility and continuity in the songs of thrushes
897 *Turdus* spp. *Ibis* 127, 355-364.
- 898 Isaac, D., Marler, P., 1963. Ordering of sequences of singing behaviour of mistle thrushes
899 in relationship to timing. *Anim. Behav.* 11, 179-188.
- 900 Jenny, M., 1990. Populationsdynamik der Feldlerche *Alauda arvensis* in einer intensive
901 genutzten Agrarlandschaft des schweizerischen Mittellandes. *Orn. Beob.* 87, 153-163.

- 902 Jurasinski, G., 2007. simba: A Collection of functions for similarity calculation of binary
903 data. R package version 0.2-5.
- 904 Kipper, S., Mundry, R., Sommer, C., Hultsch, H., Todt, D., 2006. Song repertoire size is
905 correlated with body measures and arrival date in common nightingales, *Luscinia*
906 *megarhynchos*. *Anim. Behav.* 71, 211-217.
- 907 Koetz, A. H., Westcott, D. A., Congdon, B. C., 2007. Spatial pattern of song element
908 sharing and its implications for song learning in the chowchilla, *Orthonyx spaldingii*.
909 *Anim. Behav.* 74, 1019-1028.
- 910 Kreutzer, M., 1974. Stéréotypies et variations dans les chants de proclamation territoriale
911 chez le Troglodyte (*Troglodytes troglodytes*). *Rev. Comp. Animal* 8, 270-286.
- 912 Kroodsma, D. E., 1975. Song patterning in the rock wren. *Condor* 77, 294-303.
- 913 Kroodsma, D. E., 1978. Continuity and versatility in bird song: support for the
914 monotony-threshold hypothesis. *Nature* 274, 681-683.
- 915 Kroodsma, D. E., 1981. Winter Wren singing behavior: a pinnacle of song complexity.
916 *Condor* 82, 357-365.
- 917 Kroodsma, D. E., 1996. Kroodsma, D. E., Miller, E. H. (Eds) Ecology and Evolution of
918 Acoustic Communication in Birds. Cornell University Press, Ithaca, New York. pp. 3-19;
919 125-147.
- 920 Lachlan, R. F., Slater, P. J. B., 2003. Song learning by chaffinches: how accurately and
921 from where? A simulation analysis of patterns of geographical variation. *Anim. Behav.*
922 65, 957-969.
- 923 Laiolo, P., Tella, J. L., 2005. Habitat fragmentation affects culture transmission: patterns
924 of song matching in Dupont's lark. *J. Appl. Ecol.* 42, 1183-1193.

- 925 Laiolo, P., Tella, J. L., 2007. Erosion of animal cultures in fragmented landscapes. *Front.*
926 *Ecol. Env.* 5, 68-72.
- 927 Laiolo, P., Vögeli M., Serrano D., Tella J. L., 2008. Song diversity predicts the viability
928 of fragmented bird populations. *Plos-One* 3, e1822
- 929 Lambrechts, M. M., Dhondt, A. A., 1995. Individual voice discrimination in birds. *Curr.*
930 *Ornithol.* 12, 115-139.
- 931 Lampe, H. M., Espmark, Y. E., 1994. Song structure reflects male quality in pied
932 flycatchers (*Ficedula hypoleuca*). *Anim. Behav.* 47, 869-876.
- 933 Lehongre, K., Aubin, T., Robin, S., Del Negro, C., 2008. Individual signature in canary
934 songs: contribution of multiple levels of song structure. *Ethology* 114, 425-435.
- 935 Leitner, S., Voigt, C., Gahr, M., 2001. Seasonal changes in the song pattern of the non-
936 domesticated island canary (*Serinus canaria*), a field study. *Behaviour* 138, 885-904.
- 937 Lemon, R. E., Chatfield, C., 1971. Organization of song in Cardinals. *Anim. Behav.* 19, 1-
938 17.
- 939 Lemon, R. E., Chatfield, C., 1973. Organization of song of Rose-breasted Grosbeaks.
940 *Anim. Behav.* 21, 28-44.
- 941 Lynch, A., 1996. The population memetics of birdsong. In: Kroodsma, D. E, Miller, E. H.
942 (Eds.), *Ecology and Evolution of Acoustic Communication in Birds*. Cornell University
943 Press, Ithaca, New York, pp. 181-197.
- 944 Loffredo, C. A., Borgia, G., 1986. Male courtship vocalization as cues for mate choice in
945 the satin bowerbird, *Ptilonorhynchus violaceus*. *Auk* 103, 189-195.
- 946 Mandelbrot, B., 1953. Contribution à la théorie mathématique des jeux de
947 communication. *Publ. L'inst. Stat. L'Univ. Paris* 2, 5-50.

- 948 Mantel, N., 1967. The detection of disease clustering and a generalized regression
949 approach. *Cancer Res.* 27, 209-220.
- 950 Marler, P., Sherman, V., 1985. Innate differences in singing behavior of sparrows reared
951 in isolation from adult conspecific song. *Anim. Behav.* 33, 57-71.
- 952 McCowan, B., Doyle, L. R., Hanser, S. F., 2002. Using information theory to assess the
953 diversity, complexity, and development of communication repertoires. *J. Comp. Psychol.*
954 116, 166-172.
- 955 Morton, E. S., 1987. The effect of distance and isolation on song-type sharing in the
956 Carolina wren. *Wilson Bull.* 99, 601-610.
- 957 Mulligan, J. A., 1966. Singing behavior and its development in the song sparrow,
958 *Melospiza melodia*. *Univ. Calif. Publ. Zool.* 81, 1-76.
- 959 Munding, P. C., 1982. Microgeographic and macrogeographic variation in the acquired
960 vocalizations of birds. In: Kroodsma, D. E., Miller, E. H. (Eds.), *Acoustic communication*
961 *in birds II*. Academic Press, New York, pp. 147-208.
- 962 Mundry, R., 1999. Testing related samples with missing values: a permutation approach.
963 *Anim. Behav.* 58, 1143-1153.
- 964 Naguib, M., 2005. Singing interactions in songbirds: implications for social relations and
965 territorial settlement. In: McGregor, P. (Ed.), *Animal Communication Networks*.
966 Cambridge University Press, Cambridge, pp. 300-319.
- 967 Naugler, C. T., Smith, P. C., 1991. Song similarity in an isolated population of fox
968 sparrow (*Passerella iliaca*). *Condor* 93, 1001-1003.
- 969 Nordby, J. C., Campbell, S. E., Beecher, M. D., 1999. Ecological correlates of song
970 learning in song sparrows. *Behav. Ecol.* 10, 287-297.

- 971 Nottebohm, F., 1969. The song of the chingolo (*Zonotrichia capensi*) in Argentina:
972 description and evaluation of a system of dialects. *Condor* 71, 299-315.
- 973 Nottebohm, F., Nottebohm, M. E., 1978. Relationship between song repertoire and age in
974 the canary, *Serinus canarius*. *Z. Tierpsychol.*, 46, 298-305.
- 975 Nowicki, S., Hasselquist, D., Bensch, S., Peters, S., 2000. Nestling growth and song
976 repertoire size in great reed warblers: evidence for song learning as an indicator
977 mechanism in mate choice. *Proc. R. Soc. Lond. B.* 267, 2419-2424.
- 978 Odderskaer, P., Prang, A., Poulsen, J. G., Andersen, P. N., Elmegaard, N., 1997. Skylarks
979 (*Alauda arvensis*) utilisation of micro-habitats in spring barley fields. *Agric. Ecos. Env.*
980 62, 21-29.
- 981 Opdam, P., Foppen, R., Reijnen R., Schotman, A., 1994. The landscape ecological
982 approach in bird conservation: integrating the metapopulation concept into spatial
983 planning. *Ibis* 137, 139-146.
- 984 Osiejuk, T. S., Ratynska, K., Cygan, J. P., Dale, S., 2003. Song structure and repertoire
985 variation in Ortolan Bunting (*Emberiza hortulana L.*) from isolated Norwegian
986 population. *An. Zool. Fen.* 40, 3-16.
- 987 R Development Core Team, 2007. R: A language and environment for statistical
988 computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-
989 0, URL <http://www.R-project.org>.
- 990 Rossi, J. P., 1996. Statistical tool for soil biology .11. Autocorrelogram and Mantel test.
991 *Europ. J. Soil Biol.* 32, 195-203.
- 992 Saunders, D. A., Hobbs, R. J., Margules, C. R., 1991. Biological consequences of
993 ecosystem fragmentation: a review. *Conser. Biol.* 5, 18-32.

- 994 Schmiegelow, F. K. A., Monkkonen, M., 2003. Habitat loss and fragmentation in
995 dynamic landscapes: avian perspectives from the boreal forest. *Ecol. Appl.* 12, 375-389.
- 996 Schottler, B., 1995. Songs of blue tits *Parus caeruleus palmensis* from las Palmas
997 (Canary. islands) — a test of hypotheses. — *Bioacoustics* 6, 135-152.
- 998 Shannon, C., 1948. A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379-
999 423, 623-656.
- 1000 Shannon, C., Weaver, W., 1949. The Mathematical Theory of Communication.
1001 University of Illinois Press, Chicago.
- 1002 Slater, P. J. B., 1989. Bird song learning: causes and consequences. *Ethol. Ecol. Evol.* 1,
1003 19-43.
- 1004 Sokal, R. R., Neal, O., 1978. Spatial autocorrelation in biology. II. Some implications and
1005 four applications of evolutionary interest. *Biol. J. Lin. Soc.* 10, 229-249.
- 1006 Sokal, R. R., Rohlf, F. J., 1995. Freeman, W. H. (Ed.), Biometry III, New York.
- 1007 Specht, R., 2004. Avisoft-SASLab pro v4.31. Avisoft, Berlin.
- 1008 StatSoft, Inc., 2001. Statistica v6.0. for Windows (Computer program manual). StatSoft,
1009 Inc., Tulsa: Oklahoma.
- 1010 Stawarczyk, T., 2004. Barycz valley. In: Sidło, P. O., Błaszowska, B., Chylarecki, P.
1011 (Eds.), Important Bird Areas of European Union importance in Poland. Polish Society for
1012 the Protection of Birds, BirdLife partner in Poland. Warszawa, pp. 432-434.
- 1013 Stoddard, P. K., 1996. Vocal recognition of neighbors by territorial Passerines. In:
1014 Kroodsma, D. E, Miller, E. H. (Eds.), Ecology and Evolution of Acoustic
1015 Communication in Birds. Cornell University Press, Ithaca, New York, pp. 356-374.

- 1016 Suggs, D. N., Simmons, A. M., 2005. Information theory analysis of patterns of
1017 modulation in the advertisement call of the male bullfrog, *Rana catesbeiana*. *J. Acoust.*
1018 *Soc. Am.* 117, 2330-2337.
- 1019 Suzuki, R., Buck, J. R., Tyack, P. L., 2005. Information entropy of humpback whale
1020 songs. *J. Acoust. Soc. Am.* 119, 1849-1866.
- 1021 Trainer, J. M., 1983. Changes in song dialect distributions and microgeographic variation
1022 in song of white-crowned sparrows (*Zonotrichia leucophrys nuttalli*). *Auk* 100, 568-582.
- 1023 Tryjanowski, P., 2000. Changes in breeding populations of some farmland birds in W
1024 Poland in relation to changes in crop structure, weather conditions and number of
1025 predators. *Folia Zool.* 49, 305-315.
- 1026 Tryjanowski, P., 2007. Skylark – *Alauda arvensis*. In: Sikora, A., Rohde, Z., Gromadzki,
1027 M., Neubauer, G., Chylarecki, P. (Eds.), The atlas of breeding birds in Poland 1985-
1028 2004. Bogucki Wydawnictwo Naukowe, Poznań, pp. 318-319.
- 1029 Wildenthal, J. L., 1965. Structure in primary song of the Mockingbird (*Mimus*
1030 *polyglottos*). *Auk* 82 161-189.
- 1031 Wilson, P. L., Towner, M. C., Vehrencamp, S. L., 2000. Survival and song-type sharing
1032 in a sedentary subspecies of the song sparrow. *Condor* 102, 355-363.
- 1033 Zipf, G. K., 1949. Human Behavior and the Principle of Least Effort. Mass. : Addison-
1034 Wesley Press, Cambridge.
- 1035

1035 **Tables**

1036

1037 Table 1

1038 List of abbreviations and their corresponding parameters.

Abbreviation	Parameter
<i>ObsR</i>	observed repertoire size
<i>R</i>	total repertoire size
<i>SylVer</i>	syllable type versatility
<i>TrVer</i>	transition versatility
<i>TotVer</i>	total versatility
<i>En</i>	n-order entropy
<i>REn</i>	n-order relative entropy
<i>RDn</i>	n-order redundancy
<i>RS</i>	coefficient of repertoire similarity

1039

Accepted manuscript

1039 Table 2

1040 Versatility. Mean \pm SE values of versatility obtained in the two kinds of habitat for each
 1041 individual: Total number of syllables analysed (*Nb Syl analysed*), Observed repertoire
 1042 size (*ObsR*), Estimated repertoire size (*R*), Total number of sequences of 10 syllables
 1043 analysed (*Nb Seq analysed*), Syllables versatility (*SylVer*), Transition versatility (*TrVer*)
 1044 and Total versatility (*TotVer*).

1045

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

1046

1046 Table 3
 1047 Information analysis. Mean \pm SE values of entropy obtained in the two kinds of habitat
 1048 for each individual: Entropy (E) (in bits/syllable-type), Relative entropy (RE) and
 1049 Redundancy (RD). Only first- (E_1), second- (E_2) and third- (E_3) order entropy values are
 1050 presented here. Higher-order entropy values are presented in the information graph (Fig.
 1051 4).

1052

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	8.398	6.941	0.952	0.326	0.937	0.129	0.045	0.063	0.871	0.956
	\pm	\pm	\pm	\pm	\pm	\pm	\pm	\pm	\pm	\pm	\pm
	0.081	0.108	0.102	0.031	0.024	0.004	0.004	0.003	0.004	0.004	0.004
Continuous ($N = 20$)	7.492	8.373	7.047	1.004	0.286	0.941	0.134	0.038	0.059	0.866	0.962
	\pm	\pm	\pm	\pm	\pm	\pm	\pm	\pm	\pm	\pm	\pm
	0.044	0.093	0.051	0.024	0.011	0.002	0.003	0.002	0.002	0.003	0.002

1053

1053 Table 4
 1054 Repertoire similarity. Mean \pm SE, within each kind of habitat (fragmented and
 1055 continuous), number of syllables (*Nb Syl*), number of sequences of 2 syllables (*Nb Seq 2*)
 1056 and number of sequences of 3 to 10 syllables (*Nb Seq 3-10*) shared between pairs of close
 1057 individuals (**C**) and between pairs of distant individuals (**D**), with the associated
 1058 coefficients of repertoire similarity (*RS₁*, *RS₂* and *RS₃₋₁₀* respectively).
 1059

Habitat	C or D	<i>Nb Syl</i>	<i>Nb Seq 2</i>	<i>Nb Seq 3-10</i>	<i>RS₁</i>	<i>RS₂</i>	<i>RS₃₋₁₀</i>
Fragmented (<i>N</i> = 91 pairs)	C (<i>N</i> =10)	80.4 \pm 6.0	11.7 \pm 1.2	21.6 \pm 3.2	0.300 \pm 0.016	0.231 \pm 0.024	0.239 \pm 0.029
	D (<i>N</i> =81)	38.8 \pm 1.4	1.4 \pm 0.1	0.4 \pm 0.1	0.125 \pm 0.004	0.025 \pm 0.003	0.003 \pm 0.001
Continuous (<i>N</i> = 190 pairs)	C (<i>N</i> =24)	76.0 \pm 3.2	11.0 \pm 1.1	14.2 \pm 1.8	0.257 \pm 0.011	0.153 \pm 0.015	0.126 \pm 0.018
	D (<i>N</i> =166)	49.6 \pm 1.0	2.8 \pm 0.3	1.6 \pm 0.3	0.159 \pm 0.003	0.035 \pm 0.003	0.013 \pm 0.003

1060

1060 Table 5

1061 Effect of the kind of habitat on syllable and sequence sharing. Results, within each
 1062 habitat (fragmented and continuous) separately and in both kinds of habitats, of two-way
 1063 analyses of covariance (ANCOVAs) on coefficients of repertoire similarity (RS_1 , RS_2 and
 1064 RS_{3-10}), with RS as a dependant variable, the distance between birds as a covariate, the
 1065 proximity between birds (close vs. distant) as a fixed effect, and, for the ANCOVA
 1066 carried out on both habitat, the kind of habitat as another fixed effect. For this ANCOVA,
 1067 the interaction between the 2 fixed effects (the proximity between birds and the kind of
 1068 habitat) is stated.

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	***

1069

1070 Abbreviations for P values: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, NS = Non Significant.

1071 **Figure Legends**

1072

1073 Fig 1. Aerial photographs of the 2 studied areas: (A) Fragmented habitat (Orsay, Essone,
1074 France); (B) Continuous habitat (Odolanów, Wielkopolska, Poland). The locations of the
1075 7 studied groups in the fragmented habitat and of the 2 studied sets of individuals in the
1076 continuous habitat are indicated (Google Earth version 4.3., Google, 2008).

1077

1078 Fig. 2. Geographical distances between individuals. Plot, for the fragmented habitat (A)
1079 and continuous habitat (B), of the first principal component (PC1) against the second
1080 principal component (PC2) extracted from the analysis in principal coordinates.

1081 Individuals are identified with a number followed by 'f' in the fragmented habitat and
1082 followed by 'c' in the continuous one. For the fragmented habitat, dotted rectangles
1083 delimit the 7 studied groups of individuals (1 to 3 individuals recorded per group). For
1084 the continuous habitat, dotted rectangles delimit the 2 studied sets of individuals.

1085 Scales were estimated by comparing distances separating pairs of individuals on the plots to real
1086 geographical distances between them. These scales are ± 10 m accurate.

1087

1088 Fig. 3. Repertoire size. Example of observed (grey) and estimated (black) repertoire
1089 curves obtained for one individual in the continuous habitat. This individual has an
1090 observed repertoire size of 207 syllables and an estimated repertoire size of 480 syllables.

1091

1092 Fig. 4. Information graph. Mean \pm SE entropy values obtained in the fragmented habitat
1093 (grey, $N = 14$ individuals) and in the continuous habitat (black, $N = 20$ individuals)
1094 plotted as a function of Entropy orders (size of syllable sequences).

1095

1096 Fig. 5. Zipf plot. Logarithm of the probability of occurrence of each syllable type
 1097 encountered in the fragmented habitat (grey curve, $N = 952$ syllable types) and in the
 1098 continuous habitat (black curve, $N = 1109$ syllable types) plotted against the logarithm of
 1099 syllable type rank of use. The dotted line represents Zipf's law (regression coefficient = -
 1100 1.00) that fits human language. The skylark song is more close to Mandelbrot's
 1101 modification than Zipf's law, with a hyperbolic curve at the upper data (lowest rank
 1102 elements).

1103
 1104 Fig. 6. Correlation between syllables and sequences shared by close and distant
 1105 individuals and geographical distances separating these individuals. RS_1 (A graph), RS_2
 1106 (B graph) and RS_{3-10} (C graph) calculated between pairs of close individuals (cF) and
 1107 distant individuals (dF) in the fragmented habitat (in grey) and between pairs of close
 1108 individuals (cC) and distant individuals (dC) in the continuous habitat (in black) plotted
 1109 as a function of geographical distances between individuals. Pearson correlation
 1110 coefficients calculated for close and distant individuals with a permuted correlation test
 1111 between RS values and distances are indicated above the regression lines for the
 1112 fragmented habitat (grey) and continuous habitat (black).

1113 When correlations were significant, P values are indicated as follows: * $P < 0.05$, *** $P < 0.001$.

1114

1115 Fig. 7. Patterns of syllable and sequence sharing. Cluster dendrograms, for the
 1116 fragmented habitat (A) and continuous habitat (B), obtained from hierarchical cluster
 1117 analysis using the average agglomeration method on each matrix of $1-RS_1$, $1-RS_2$ and $1-$
 1118 RS_{3-10} calculated between pairs of individuals. For the fragmented habitat, when clusters
 1119 reflect geographical distances, dotted rectangles delimit the 7 studied groups (1 to 3

- 1120 individuals recorded per group) separated by unsuitable area. For the continuous habitat,
1121 dotted rectangles delimit the 2 sets of individuals (cf. Fig. 2).

Accepted manuscript















