

Territorial song does not isolate Yellowhammers (*Emberiza citrinella*) from Pine Buntings (*E. leucocephalos*)

DIETER THOMAS TIETZE, CHRISTINE WASSMANN & JOCHEN MARTENS

Institut für Zoologie, Johannes Gutenberg-Universität, 55099 Mainz, Germany.
mail(at)dieterthomastietze.de

Accepted on November 21, 2011.

Published online at www.vertebrate-zoology.de on April 05, 2012.

> Abstract

Western Palearctic Yellowhammer and eastern Palearctic Pine Bunting are considered a pair of young sister species. A huge area of sympatry is in western Siberia where they hybridize extensively despite striking plumage differences. The somehow dominant Yellowhammer pushes the hybrid zone further east. Territorial songs of the two taxa are almost indistinguishable. We tested 15 song types from both species from various localities inside and outside the hybrid zone in playback experiments on German Yellowhammer males. All songs elicited territorial behavior in almost every experiment. Reaction intensity depended on one out of 25 investigated sonographic parameters and was the weaker the further away the recording site of the playback was from the experimental site regardless of the taxon. Songs from the hybrid zone elicited disproportionately strong response so that a reinforcement of reproductive isolation through song can be excluded. Song seems to be more conserved than morphological features and has thus prevented a complete genetic divergence in this pair of subspecies groups.

> Key words

Emberiza citrinella; *Emberiza leucocephalos*; reproductive isolation; song differentiation; semispecies; subspecies group.

Introduction

In the breeding season, male songbirds use their songs to mark and defend their territories against conspecific competitors and to attract mates by advertising their qualities (CATCHPOLE & SLATER 1995: 116, 142). Both purposes require a species-specific song which restricts variability to a certain amount. Conversely, these territorial songs should have become so much differentiated during the speciation process that they work as isolating mechanism even between sympatric sister species. Thus in birds vocalizations provide behavioral traits which beside morphological and genetic characters may give hints to taxonomic relationships, specifically in cases of unclear taxonomic ranking of allopatric or parapatric populations. If taxa in question do not get in contact in nature, reciprocal playback experiments can simulate a secondary contact. But even in the case of secondary contact with hybridization such experiments help to find break points across the hybrid zones or to test for reinforcement of isolating mechanisms such as song between recently diverged species.

Yellowhammer *Emberiza citrinella* LINNAEUS, 1758 (YH) and Pine Bunting *E. leucocephalos* GMELIN, 1771 (PB) are such a pair of closely related sister species (ALSTRÖM *et al.*, 2008; IRWIN *et al.*, 2009). But they hybridize over a large zone of sympatry (PORTENKO & STÜBS, 1971). This is quite surprising given the striking morphological differences which apparently cannot maintain reproductive isolation. While YH males in breeding plumage show their characteristic yellow on head and breast, PBs lack lipochrome and thus have all yellow replaced by white (JOHANSEN, 1944). Birds in less conspicuous plumage might be hard to assign to either species. The YH is distributed in open landscape from the European Atlantic coast east to Central Siberia, while the PB breeds in open coniferous and mixed forests from the Asian Pacific coast west to West Siberia, there is even a partial overlap in ecological niche (CRAMP & PERRINS, 1994). Their areas overlap from the Ural Mountains roughly 4000 km east to Lake Baikal in an up to 1800 km wide band (PORTENKO & STÜBS, 1971). LÖHRL (1967) as-

sumed that this secondary contact became established after the last glaciation. YH and PB hybridize in the center of the southern part of that huge zone of overlap; hybrids are fertile so that various intermediate phenotypes can be observed (2.5% of the total sympatric population F1 hybrids, 15% PBs and 20% YHs with heterospecific morphological features) and the hybrid zone has continued to grow (KLEINSCHMIDT, 1903; GLUTZ VON BLOTZHEIM & BAUER, 1997; PANOV *et al.*, 2003). Isolating mechanisms fail only in some places (GLUTZ VON BLOTZHEIM & BAUER, 1997), it has remained unclear why. PANOVA *et al.* (2003) showed in experiments that isolation is maintained by plumage coloration rather than by song.

Songs of both species are very similar from auditory impression and sonagrams (Fig. 1) at first glance (LÖHRL, 1967): In a first phrase one syllable (1–3 elements which are inborn) is repeated several times, the second (and last) phrase consists of one extended element of constant frequency that might be preceded by a slightly shorter element of higher frequency. This last part needs to be learnt and is thus used to define local dialects. Volume also increases to a maximum during the final phrase (GLAUBRECHT, 1989; WALLSCHLÄGER, 1998; CARO *et al.*, 2009).

We used song recordings from various YH and PB populations from within and outside the hybrid zone to work out the degree of reproductive isolation that can be deduced from these vocalizations in behavioral experiments. Furthermore we are interested in whether differently intensive response to the playbacks can be explained by certain sonographic parameters or the distance between recording and playback sites.

Material and methods

Definitions

Territorial song: Vocalization consisting of various subunits, e.g. a sequence of verses delimited by pauses, uttered by males during the breeding season in order to defend a territory or to attract a mate.

Call: Vocalization consisting of one or few elements, not bound to the breeding season, but potentially context-specific, e.g. alarm call.

Verse (of territorial song): Sequence of elements, syllables or phrases, separated from other verses by distinct pauses.

Element: Smallest unit of a vocalization, recognizable as a continuous blackening in a sonagram.

Syllable: Sequence of few elements of different kind.

Phrase: Sequence of elements or syllables of same kind.

Motif: Distinct sequence of elements of different kind.

Song type: Variant of verse.

Dialect: Regional variation of the song in populations within a species, reciprocal understanding is not diminished, dialect areas may overlap (WICKLER, 1986).

Regiolect: Macrogeographic variation of the song, no overlap of regiolect areas (except for secondary contact), understanding between populations of different regiolects restricted or even impossible (MARTENS, 1996).

Material

We used 16 different recordings of YH and PB song (Table 1, Fig. 1) representing all three YH and both PB subspecies *sensu* DICKINSON (2003) to prepare playbacks in Avisoft SASlab Pro (Sound Analysis and Synthesis Laboratory) v4.37 (SPECHT, 2005). Selection of verses was based on song type and geographic distribution of the two taxa such that possibly some variation within a population repertoire is reflected (Table 1). For double control playback we used a typical YH verse from Germany (playback 0). We removed potentially disturbing background noise from the recordings. Each playback finally consisted of ten copies of a single verse interrupted by 6 s of silence each.

Field experiments

The general design of the playback experiments followed THIELCKE (1969; cf. TIETZE *et al.*, 2011). We performed the experiments between Mainz and Kirchheimbolanden, Rheinland-Pfalz, Germany (ca. 20 km around 49° 50' N, 8° 9' E) between May 17th and June 17th 2005. Experiments were only started where a YH was detected or habitat was suitable. Every individual was tested only once to avoid habituation. Playbacks were presented in randomized order. We hid a speaker under a bush and connected it via a 10-m cable with either a Sony Portable MiniDisc Recorder MZ-N707 or a Sony Discman D-143. We presented the playback, continued to observe the bird's behavior for another 2 min, presented control playback 0, and noted reactions again for 2 min. Because experiments were performed by multiple observers, we used only few simple measures of response: Horizontal approach to

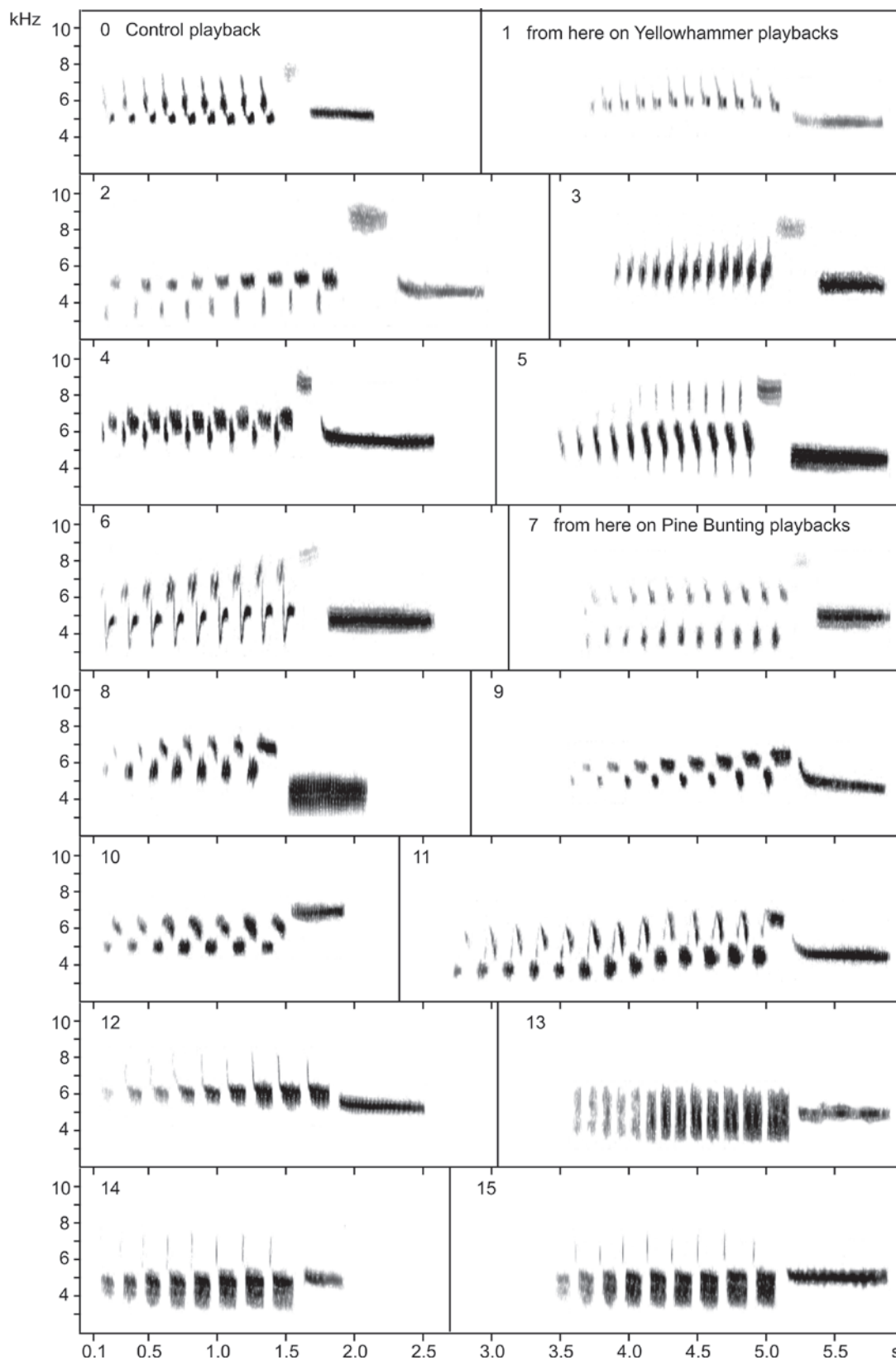


Fig. 1. Sonagram plate of the Yellowhammer (0–6) and Pine Bunting (7–15) verses used for playback experiments. For details see Table 1.

the speaker was given up to three points (3 for 5–0 m, 2 for 10–5 m, 1 for 20–10 m), all other reactions (vertical approach to the speaker, territorial song,

calls, search flight, wing twitch) only one point each, resulting in maximally eight points of reaction intensity.

Table 1. Recordings used as playbacks and mean intensity of reaction to those: SDB = ID in J. Martens's sound database, n = number of successful field experiments with this playback, V = Wilcoxon test statistic, comparing reaction intensity to playback and corresponding control playback. Playback 0 was used for a double-control experiment. *omitted since population introduced from the UK.

Playback	Taxon	SDB	Recording locality (distance to 0 in km)	Year	Recordist	n	Playback	Control	Plb./Ctr.	V	P
0	<i>E. c. cinerea</i>	759	Mainz, Rheinland-Pfalz, Germany (0)	1993	J. Martens	15	5.00 ± 1.31	5.20 ± 1.21	0.99 ± 0.30	22	0.565
1	<i>E. c. caliginosa</i>	678	Farewell Spit, South Island, New Zealand (*)	1990	J. Martens	12	3.83 ± 1.95	4.33 ± 1.50	0.95 ± 0.46	16	0.468
2	<i>E. c. caliginosa</i>	772	Farewell Spit, South Island, New Zealand (*)	1990	J. Martens	10	2.70 ± 2.41	4.60 ± 0.84	0.62 ± 0.55	3	0.073
3	<i>E. c. erythrogenys</i>	765	Akademgorodok, Novosibirsk Oblast, Russia (4849)	1986	J. Martens	12	4.25 ± 1.71	4.50 ± 0.80	0.97 ± 0.37	16	0.831
4	<i>E. c. erythrogenys</i>	1259	Soyamy, Tuva, Russia (5629)	1973	B. Vepintsev	12	4.00 ± 1.76	4.50 ± 1.45	0.94 ± 0.43	11	0.354
5	<i>E. c. erythrogenys</i>	1676	Outskirts of Tomsk, Tomsk Oblast, Russia (4895)	2002	S. Gashkov	14	4.00 ± 2.00	4.79 ± 1.12	0.92 ± 0.60	20	0.261
6	<i>E. c. erythrogenys</i>	1677	Outskirts of Tomsk, Tomsk Oblast, Russia (4895)	2003	S. Gashkov	15	4.73 ± 1.62	5.20 ± 0.86	0.93 ± 0.32	12	0.429
7	<i>E. l. leucocephalus</i>	1410	Outskirts of Tomsk, Tomsk Oblast, Russia (4895)	2003	S. Gashkov	15	4.20 ± 2.11	4.73 ± 1.53	0.93 ± 0.43	16	0.433
8	<i>E. l. leucocephalus</i>	1260	Erzin, Tuva, Russia (5829)	1975	B. Vepintsev	14	4.64 ± 1.50	4.29 ± 0.99	1.17 ± 0.62	47	0.546
9	<i>E. l. leucocephalus</i>	978	Issyk Kul basin, Kyrgyzstan (5198)	1993	J. Martens	15	3.80 ± 2.01	4.60 ± 1.35	0.82 ± 0.33	6	0.053
10	<i>E. l. leucocephalus</i>	978	Issyk Kul basin, Kyrgyzstan (5198)	1993	J. Martens	10	3.80 ± 1.75	4.30 ± 1.25	0.92 ± 0.42	5	0.586
11	<i>E. l. leucocephalus</i>	610	Issyk Kul basin, Kyrgyzstan (5198)	1993	J. Martens	15	4.40 ± 1.24	4.20 ± 1.52	1.32 ± 1.09	27	0.631
12	<i>E. l. fronto</i>	1019	Qinghai Lake, Qinghai, China (7093)	1996	J. Martens	14	3.64 ± 1.86	3.79 ± 1.81	1.01 ± 0.33	16	0.774
13	<i>E. l. fronto</i>	1019	Qinghai Lake, Qinghai, China (7093)	1996	J. Martens	15	2.87 ± 1.96	4.20 ± 1.52	0.76 ± 0.53	17	0.044
14	<i>E. l. fronto</i>	902	Xining, Qinghai, China (7219)	1996	J. Martens	14	3.36 ± 1.01	3.71 ± 1.38	1.15 ± 0.91	11	0.609
15	<i>E. l. fronto</i>	903	Xining, Qinghai, China (7219)	1996	J. Martens	15	3.20 ± 2.11	3.27 ± 1.83	1.31 ± 1.44	34	1.000

Statistical analysis

We correlated reaction intensities with day of the year and time of the day. We compared the intensity of behavioral reaction to playback with the one to the corresponding control playback for all different playbacks (Wilcoxon rank-sum test). Further pairwise comparisons between reasonable groups of playbacks were among reaction score to actual playback and ratios of playback/control experiments (Mann-Whitney *U* test). Mean reaction intensity was furthermore correlated with 25 standard (cf. TIETZE *et al.*, 2008, 2011) and YH/PB specific verse parameters (RUBTSOV *et al.*, 2007; CARO *et al.*, 2009) as well as geographic distance of the experimental site from the corresponding recording locality (Table 1). All statistical analyses were done in R v2.11.1 (R DEVELOPMENT CORE TEAM, 2010).

Results

Field experiments

In total we performed 217 successful experiments (six others excluded, because no reaction at all was observed), i.e. 10–15 per playback. YH males reacted to 206 of the presented test playbacks (95%). Rarely did they start own singing during the presentation of the playback, instead approached and replied with own song in the 2-min break. Reaction to the first (test) and second (control) playback in a given experiment were highly significantly correlated with each other (Pearson's product-moment correlation: $t_{215} = 5.14$, $r = 0.33$ $P < 0.0001$; Spearman's rank correlation: $S = 1134502$, $\rho = 0.33$, $P < 0.0001$). In general there was no significant difference between reaction intensities to control and test playbacks per given playback (Wilcoxon rank-sum test: $P > 0.05$; Table 1, cf. Fig. 2) with only one exception (playback 13). While there was no significant correlation between reaction intensity to test playbacks pooled and either time of the day or calendar date (Spearman's rank correlation: $P > 0.5$), reaction to (local) control playbacks got more intensive from beginning to end of the field season (Spearman's rank correlation: $S = 1410484$, $\rho = 0.17$, $P = 0.011$ roughly by 1.5 points) and less intensive from morning towards evening (Spearman's rank correlation: $S = 1967245$, $\rho = -0.16$, $P = 0.02$; roughly by 1.5 points).

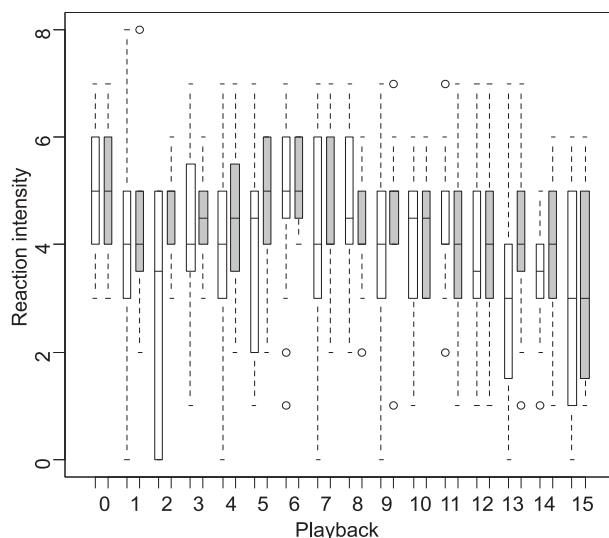


Fig. 2. Intensity of reaction to the playbacks (open boxes, cf. Table 1) and corresponding control playbacks (gray boxes); left-most pair (0) indicates double-control experiment.

Comparison of playbacks

We grouped the experiments by origin of the playback verses (species, recording in sympatry with or allopatry from the other species) and tested reaction intensity to the test playback as well as the ratio of reaction to test and control playback of the two groups against each other. Differences were normally not significant (Table 2; Mann-Whitney U test: $P > 0.05$), but raw intensities differed significantly between (1) PB sympatric with YH and PB allopatric from YH and (2) YH from Germany and YH from New Zealand.

Search for causation

Only one out of the 25 investigated verse parameters correlated significantly with mean reaction intensity to the corresponding playback (maximum frequency of phrase 1, Table 3). Recordings of both species are randomly distributed along the regression line in Fig. 3a. The number of elements in phrase 2 and the length of the pause between phrase 1 and the last element were the only parameters of the test playback that significantly correlated with the mean reaction intensity to the control playback following the actual test playback (Table 3). Both reaction intensity to test and control playbacks were significantly negatively correlated with the geographic distance between recording and playback locality of the respective experimental series (Table 3, Fig. 3b). This includes PBs of *ssp. fronto* from Qinghai, which is separated by a large geographical disjunction from nominate *leucocephalos*. Most playbacks from the zone of sympatry (six out of nine) – of both species – are situated above the

Table 3. Pearson's correlation coefficients for the meaningful pairwise correlations of mean reaction intensities, distance from recording site (see Table 1), and 25 sonographic parameters taken from the 16 playback verses; $P < 0.05$, unless values in italics.

Parameter	1	2	3
1. Response to test playback	1.00	0.57	-0.76
2. Response to control playback	0.57	1.00	-0.77
3. Distance from recording site	-0.76	-0.77	1.00
4. Maximum frequency of phrase 1	0.57	<i>-0.10</i>	<i>-0.08</i>
5. Number of elements in phrase 2	<i>0.30</i>	0.70	-0.58
6. Length of pause between phrase 1 and last element	<i>0.05</i>	0.54	-0.55

regression line, thus elicited stronger response than predicted by the overall linear model. The only verse parameters that were significantly (negatively) correlated with the distance from the recording site were again number of elements in phrase 2 and the length of the pause between phrase 1 and the last element (Table 3).

Discussion

Understanding the reaction to the playbacks

All playbacks elicited territorial response from Central European YH males. Apparently, all tested songs were recognized as conspecific. Furthermore, the intensity of the reaction to the control playback depended on the intensity of the reaction to the preceding test playback. Maybe YH males had been tuned to a certain level of behavioral response by the first playback, especially in cases of a more pronounced end part of the verse. This is also corroborated by the fact that the intensities of reaction to test playback and corresponding control playback did not differ significantly. A simpler explanation would be that more aggressive males react generally more intensively to conspecific songs than do less aggressive ones or males without a territory. Thus individual mood may matter more than (negligible) dialect differences.

The intensity of the reaction to the control playback slightly changed over time, i.e. with season during which the experiments were carried out. We accounted for that through a randomized order in which we presented the test playbacks. Nevertheless, we have not found such a temporal relationship with our actual test playbacks. Intensity of the reaction to the latter rather varied among the different playbacks: On the one hand, reactions were stronger the higher-pitched the

Table 2. Comparison of reaction intensity to test playback and of ratio of test and control playback reaction between various groups of playbacks (Mann-Whitney *U* tests): Pb = playback according to Table 1, n = number of experiments, W = test statistic.

Group1	Pb1	n1	Group2	Pb2	n2	Intensity		Ratio	
						W	P	W	P
Yellowhammer (YH)	0–6	90	Pine Bunting (PB)	7–15	127	6515	0.075	5528	0.676
YH allopatric from PB	0–2	37	YH sympatric with PB	3–6	53	912	0.568	896	0.483
YH sympatric with PB	3–6	53	PB sympatric with YH	7–8	29	731	0.715	697	0.484
PB sympatric with YH	7–8	29	PB allopatric from YH	9–15	98	1809	0.024	1581	0.349
YH allopatric from PB	0–2	37	PB allopatric from YH	9–15	98	2096	0.157	1699	0.567
YH allopatric from PB	0–2	37	PB sympatric with YH	7–8	29	476	0.427	450	0.257
YH from Germany	0	15	YH from New Zealand	1–2	22	248	0.010	186	0.520

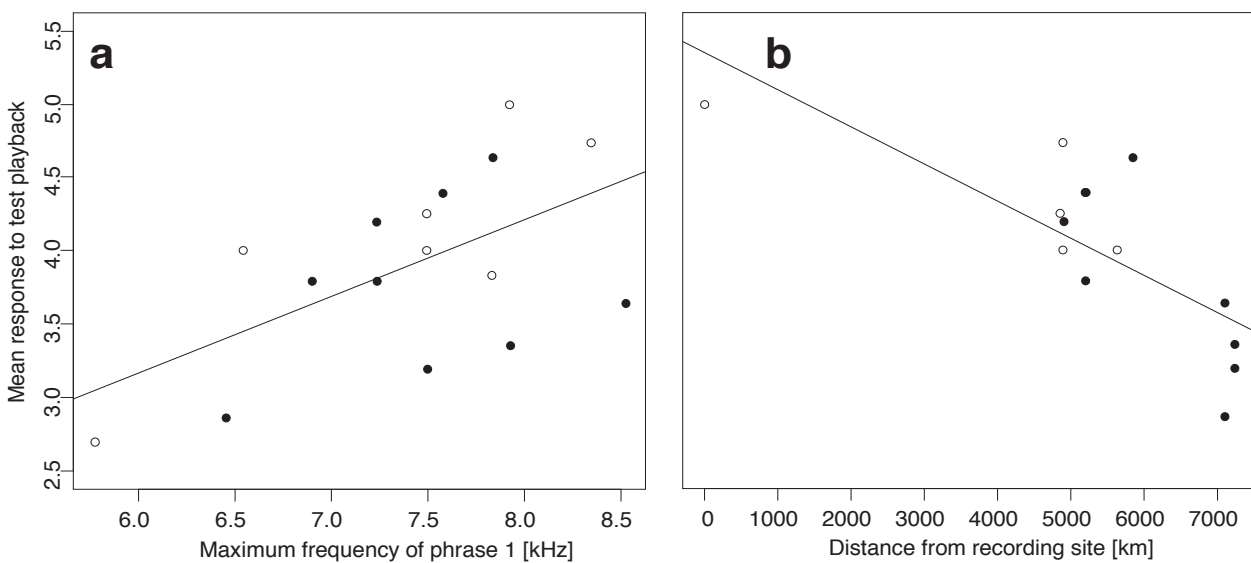


Fig. 3. Mean intensity of reaction to test playback correlated with a) the maximum frequency of phrase 1 ($t_{14} = 2.62$, $r = 0.57$, $P = 0.020$) and b) the distance between recording and playback sites ($t_{12} = -4.08$, $r = -0.76$, $P < 0.005$). \circ = Yellowhammer, \bullet = Pine Bunting. Recordings from populations introduced to New Zealand were excluded from partial figure b.

syllable repetitions of the first phrase of the playback were (control playback being third highest-pitched). On the other hand, response was the weaker the larger the distances to the recording sites were (also reflected in control reaction intensity). We consider this causal for the fact that songs from PB in sympatry with YH elicited stronger responses than songs from PB in allopatry, since the latter live even further east from the tested population than the former. Thus we have not actually found differential behavior towards songs from either species.

Standard verse parameters as analyzed here were generally not apt to explain the differences in reaction intensity. Only the number of elements in the test playback increased the response to the following control playback. The major significant explanation for the variation in reaction intensity was the geographic distance between the experimental site and the locality where the respective playback was recorded – regardless of which species sang the song. YH males learn their songs from older males during the breeding

season in which they hatch themselves (DIESELHORST, 1971; SCHÖN, 1989), disperse, and propagate that song in the following years with minor learning errors. Thus over time local dialects get established (WICKLER, 1986), which have been investigated in quite a number of YH studies (WALLSCHLÄGER, 1983, 1998; GLAUBRECHT, 1989, 1991; WONKE & WALLSCHLÄGER, 2009). Intensive gene flow between carriers of different dialects has been proven in YH (WONKE, 2008). Over larger distances song differences might reach a level where they might restrict reproduction of different regiolect carriers and thus interrupt gene flow (MARTENS, 1996): This might be the case in Chinese PBs with respect to Central European YHs. But the whole process is only driven by random changes over geographic distance and is apparently not interrupted by the species limits between YH and PB. The other way around dialects in the YH/PB species complex do not interrupt gene flow (MARTENS, 1993).

Since songs from the hybrid zone elicited disproportionately strong response in Central European YH

compared to all other playbacks, we exclude that a reinforcement of reproductive isolation (character displacement in song sensu THIELCKE, 1986) occurs in the hybrid zone. It could rather be a song convergence to compensate for the morphological differences as GRABOVSKY & PANOV (1992) found in the contact zone of Pied *Oenanthe pleschanka* and Black-eared Wheatears *Oe. hispanica*. Great Tits in the Amur region also show this phenomenon (Päckert *et al.*, 2005). In Iran, HAFFER (1977) described not only the wheatear contact zone, but also that of another *Emberiza* species pair with significant differences in male plumage but low genetic distance (ALSTRÖM *et al.*, 2007), *E. bruniceps*/*E. melanocephala*, where a third of the population consists of hybrids. Their territorial songs are also almost indistinguishable (GLUTZ VON BLOTZHEIM & BAUER, 1997). Songs of *Hippolais polyglotta* and *H. icterina* clearly differ in allopatric populations, but converge in sympatry; *H. icterina* adjusted temporal parameters and *H. polyglotta* syntax, the latter pushing the hybrid zone forward (SECONDI *et al.*, 2003). But song convergence needs not necessarily coincide with morphological divergence (e.g. the two *Regulus* species in northern Spain; Becker 1977). Simply the ratio of con- and heterospecific tutors in a population may influence learning of heterospecific song features.

Song parameters correlated with geographic distance are unexpectedly others than the one correlated with response to the test playbacks. Instead, the former ones, shorter pauses between the first phrase and the last element (e.g. because of the lack of the first of two elements in phrase 2), cause weaker response to the control playback. A proximate argument is that a phrase-1 parameter influences the response to the test playback and a phrase-2 parameter the response to the control playback. But we consider this little convincing because of the 2-min break between the two partial experiments, which should suffice to calm down the bird.

In addition to the interspecific focus of this paper we found that YHs descending from birds introduced from the British Isles to New Zealand in 1865 (GLUTZ VON BLOTZHEIM & BAUER 1997: 1434) were less well understood in Central Europe. This might either be due to bad recording quality or to island effects. The latter might either refer to the post-glacial expansion of YHs to the British Isles or the historical "isolation" through the founder effect of transferring European birds to New Zealand. On islands miscopied variants of songs have often been found in other passerine species, too, e.g. in treecreepers of genus *Certhia* in New and Old Worlds (BAPTISTA & JOHNSON, 1982; THIELCKE, 1986; TIETZE *et al.*, 2008) or kinglets of genus *Regulus* on the Balears and Azores (PÄCKERT *et al.*, 2001; PÄCKERT & MARTENS, 2004).

Song divergence in Yellowhammer and Pine Bunting

We cannot confirm the statement by CRAMP & PERRINS (1994) that PB song has fewer elements or motifs in phrase 1, but our – randomly chosen! – PB playbacks have only one (i.e. the final) element in phrase 2 except for playback 7 from the hybrid zone (Fig. 1). RUBTSOV (2007) studied YH and PB songs from Moscow east to eastern Siberia: (1) Main frequency of the final element decreased eastward, while its duration increased; we do not find the former in our dataset and only an insignificant tendency for the latter what may be due to our small sample size. (2) The further apart two study populations were the fewer similar song types they shared. This was true in conspecific and heterospecific comparisons, but on a 50% lower level in intraspecific analysis. (3) Neither direct sonagram parameters nor principal components were apt to separate the species on the level of the total trans-Eurasian range, but different features worked in specific populations. In a study across a hybrid zone from pure YHs to pure PBs (PANOV *et al.*, 2007), values of the two main factors in a multivariate analysis of sonographic parameters gradually decreased. All these findings support our own results that there is no steep break in song between YH and PB, but rather a gradual transition between neighboring dialects.

Warbler finches (*Certhidea olivacea* and *C. fusca*) from the Galápagos Islands have also not yet established acoustic barriers although a basal split in a phylogenetic tree separates them (GRANT & GRANT, 2002). As WALLSCHLÄGER (1983) already stated for the genus *Emberiza* in general, learnt song is more conservative than morphological traits, thus not a pacemaker of evolution here as is otherwise rather typical in passerine birds. Nominate Reed Buntings *Emberiza schoeniclus* also failed to discriminate own song from that of morphologically strongly differing subspecies *intermedia* and vice versa although response to foreign song was slightly weaker (MATESSI *et al.*, 2000, 2001). But we are well aware that field experiments on males are only a surrogate for the actually decisive sexual selection in songs by females. BAKER *et al.* (1987) showed that female YHs prefer local to foreign dialect and otherwise songs with larger repertoire. This finding does not challenge our conclusions, but rather helps to understand why YHs could have a selective advantage over PBs in the hybrid zone.

Genetic background

Karyotypes do not differ between the two species (RADZHABLI *et al.*, 1970). Mitochondrial differentiation is low (ALSTRÖM *et al.*, 2008; IRWIN *et al.*, 2009;

own unpublished cytochrome-*b* sequences from even more populations than sampled for the playbacks) and with under 1% markedly ranges below species level according to HELBIG *et al.*'s (1995) compilation. Such low distances hint to post-glacial speciation which has normally not occurred in the Holarctic (KLICKA & ZINK, 1997). But some nuclear DNA (AFLP and sex-linked CHD1Z) hints to a much older species split (3–6 Ma versus 14–97 ka) that is comparable to those in other *Emberiza* species pairs (IRWIN *et al.*, 2009), while other nuclear DNA (autosomal ODC in ALSTRÖM *et al.*, 2008) yet has much lower distance values for this species pair than for all others in the genus. IRWIN *et al.* (2009) explain this phenomenon with a rapid introgression of mitochondria from one species into the other, maybe caused by a selective sweep. It remains an open question which mitochondrial gene could have had what kind of advantageous mutation to not only rapidly spread within one species, but also within the other after crossing the species boundary. At least IRWIN *et al.* (2009) did not deny an ongoing gene flow in their nuclear marker gene. Sex-linked loci introgress less than autosomal loci due to the deleterious expression in the heterogametic sex, but almost as much as mitochondrial loci (CARLING & BRUMFIELD, 2008). The 20 informative AFLP markers used by IRWIN *et al.* (2009), chosen to best distinguish between allopatric YH and allopatric PB (only 10% of the genome indicated genetic differences between YH and PB), assigned intermediate or rather YH values to phenotypic hybrids on the primary axis of variation in a principal-component analysis. They concluded that there was more gene flow from the contact zone to pure YH rather than PB populations (although the hybrid zone moves eastward) and explain this finding with either higher fitness of YH backcrosses or larger dispersal distances in YH. The latter can be rejected due to more pronounced migratory behavior in extant PB (JOHANSEN, 1944; CRAMP & PERRINS, 1994): long-distance migration of Siberian birds to India and Pakistan and of Chinese birds to S China. Accordingly, the PB has longer wings at lower weight (CRAMP & PERRINS 1994), but YH has longest wings in easternmost ssp. *erythrogegens* (also hinting to a cline!). A higher fitness of YH backcrosses is rather backed by the fact that the YH has expanded its range eastward, the number of hybrids has risen and pure PB phenotypes have disappeared from certain regions (PANOV *et al.*, 2003).

Presence or absence of lipochrome not only make a distinction between YH and PB, but also between two subspecies groups in the Great Tit *Parus major* which diverged over roughly 1 Ma (PÄCKERT *et al.*, 2005). Lineage sorting in the Orchard Oriole (*Icterus spurius*) group is incomplete, nevertheless plumages diverged and migratory behavior is different (BAKER

et al., 2003). A rapid morphological change was also shown in two *Luscinia svecica* subspecies (QUESTIAU *et al.*, 1998). Thus distinctive plumage characters can change quickly and cannot be used as indicators for a long genetic isolation.

Taxonomic implications

We support the assumption of PORTENKO & STÜBS (1971) that YH and PB were not separated long enough during recent glaciations to establish effective isolating mechanisms, especially in territorial song, to prevent hybridization between the two taxa. Nevertheless single observations (e.g. MAUERSBERGER, 1971) might hint to effectiveness, but other factors such as visual reciprocal appraisal of countersinging males were not excluded. Lack of reproductive isolation due to territorial song might be part of the explanation for such a uniquely large hybrid or intergradation zone.

SHORT (1969) defined semispecies as basically allopatric, but allowed for some sympatry, and demanded them to form a zone of overlap and hybridization with competition and reinforcement of isolating mechanisms. This does not perfectly fit the case of YH and PB, thus we consider the two alternatives: Allopecies occur allopatrically with almost no hybridization – an even less appropriate description of this case. Subspecies (groups) below the taxonomic species limit have rather primary intergradation zones or at least a pure hybrid zone (cf. HEWITT, 1989). Although YH and PB do not hybridize throughout the zone of sympatry we consider the definition of subspecies groups as closest to the facts. We are quite sure that numerous similar cases are just not under debate, because intergrading subspecies normally do not differ so strikingly as do YH and PB in plumage coloration. Our suggestion is in accordance with the guidelines for assigning species rank of the British Ornithologists' Union (Helbig *et al.*, 2002) which state that “diagnosable populations joined by a cline may be treated as subspecies”.

YH and PB appear as very young species with no fully effective isolating mechanisms, explicitly territorial song. Hybridization is intensifying and YH expanding into PB range, because YH dominates in genetic (IRWIN *et al.*, 2009), morphological (lipochrome), and vocal respect (“sexier” songs). Under the assumption that the natural process will lead to a full merging of the two species, a treatment as two semispecies rather than as one biological species puts too much imagination of nearly perfect reproductive isolation into the mind and should thus be abandoned.

Acknowledgments

J.M. made the recordings during numerous field trips in the last half century, supported by various foundations and helpers acknowledged elsewhere (see MARTENS *et al.* [2011] for details). Especially to be mentioned are Feldbausch-Stiftung and Wagner-Stiftung, both at the department of biology at Mainz university, which repeatedly granted research funds for J.M. S. GASHKOV provided indispensable recordings from W Siberia. N. BREITBACH, F. FELGENHAUER, and I. LAUBE performed the field experiments with C.W. Student representatives of the department of biology at Mainz university collected money to make the field work for this study (part of C.W.'s Diplom thesis) possible. D.T.T. profited from a grant from the German Academic Exchange Service to visit Tomsk and from the Deutsche Forschungsgemeinschaft (Ti 679/2-1) while writing this paper. The authors express their gratitude to all persons mentioned. The experiments comply with German law. The authors declare that they have no conflict of interest.

References

- ALSTRÖM, P., OLSSON, U., LEI, F., WANG, H.-T., GAO, W. & SUNDBERG, P. (2008): Phylogeny and classification of the Old World Emberizini (Aves, Passeriformes). – *Molecular Phylogenetics and Evolution*, **47** (3): 960–973.
- BAKER, M.C., BJERKE, T.K., LAMPE, H.U. & ESPMARK, Y.O. (1987): Sexual response of female yellowhammers to differences in regional song dialects and repertoire sizes. – *Animal Behaviour*, **35** (2): 395–401.
- BAKER, J.M., LÓPEZ-MEDRANO, E., NAVARRO-SIGÜENZA, A.G., ROJAS-SOTO, O. & OMLAND, K.E. (2003): Recent speciation in the Orchard Oriole group: divergence of *Icterus spurius spurius* and *Icterus spurius fuertesi*. – *Auk*, **120** (3): 848–859.
- BAPTISTA, F. & KREBS, R. (2000): Vocalizations and relationships of brown creepers, *Certhia americana*: a taxonomic mystery. – *Ibis*, **142** (3): 457–465.
- BECKER, P.H. (1977): Geographische Variation des Gesanges von Winter- und Sommergoldhähnchen (*Regulus regulus*, *R. ignicapillus*). – *Vogelwarte*, **29** (1), 1–37.
- CARLING, M.D. & BRUMFIELD, R.T. (2008): Haldane's rule in an avian system: Using cline theory and divergence population genetics to test for differential introgression of mitochondrial, autosomal, and sex-linked loci across the *Passerina* bunting hybrid zone. – *Evolution*, **62** (10): 2600–2615.
- CARO, S.P., KEULEN, C. & PONCIN, P. (2009): Song repertoires in a western European population of Yellowhammers *Emberiza citrinella*. – *Acta Ornithologica* (Warsaw), **44** (1): 9–16.
- CATCHPOLE, C. K. & SLATER, P. J. B. (1995): Bird song. Biological themes and variations. – Cambridge University Press, Cambridge. VIII, 248 pp.
- CRAMP, S. & PERRINS, C.M. (eds.) (1994): Handbook of the birds of Europe, the Middle East and North Africa. The birds of the Western Palearctic. 9. Buntings and New World warblers. – Oxford University Press, Oxford. 488 pp.
- DICKINSON, E.C. (ed.) (2003): The Howard & Moore complete checklist of birds of the world. 3rd edn. – Christopher Helm, London. 1039 pp.
- DIESELHORST, G. (1971): Jugendgesang und Entwicklung des Gesanges in einer Population der Goldammer. – *Vogelwelt*, **92**: 201–226.
- GLAUBRECHT, M. (1989): Geographische Variabilität des Gesanges der Goldammer, *Emberiza citrinella*, im norddeutschen Dialekt-Grenzgebiet. – *Journal für Ornithologie*, **130** (3): 277–292.
- GLAUBRECHT, M. (1991): Gesangsvariation der Goldammer (*Emberiza citrinella*) in Norddeutschland und auf den dänischen Inseln. – *Journal für Ornithologie*, **132** (4): 441–445.
- GLUTZ VON BLOTZHEIM, U.N. & BAUER, K.M. (eds.) (1997): Handbuch der Vögel Mitteleuropas. Band 14/III. Passeriformes (5. Teil): Emberizidae – Icteridae. – AULA, Wiebelsheim. 719 pp.
- GRABOVSKY, V.I. & PANOV, E.N. (1992): Song convergence in Pied Wheatear (*Oenanthe pleschanka*) and Black-eared Wheatear (*O. hispanica*) in secondary contact zones. – *Zoologicheskii Zhurnal*, **71** (4): 75–84.
- GRANT, B.R. & GRANT, P.R. (2002): Lack of premating isolation at the base of a phylogenetic tree. – *American Naturalist*, **160** (1): 1–19.
- HAFFER, J. (1977): Secondary contact zones of birds in northern Iran. – *Bonner Zoologische Monographien*, **10**: 1–64.
- HELBIG, A.J., SEIBOLD, I., MARTENS, J. & WINK, M. (1995): Genetic differentiation and phylogenetic relationships of Bonelli's Warbler *Phylloscopus bonelli* and Green Warbler *P. nitidus*. – *Journal of Avian Biology*, **26** (2): 139–153.
- Helbig, A.J., Knox, A.G., Parkin, D.T., Sangster, G. & Collinson, M. (2002): Guidelines for assigning species rank. – *Ibis*, **144** (3): 518–525.
- HEWITT, G. M. (1989): The subdivision of species by hybrid zones. In: OTTE, D. (ed.) Speciation and its consequences. – Sinauer, Sunderland, Mass. pp 85–110.
- IRWIN, D.E., RUBTSOV, A.S. & PANOV, E.N. (2009): Mitochondrial introgression and replacement between Yellowhammers (*Emberiza citrinella*) and Pine Buntings (*Emberiza leucocephalos*) (Aves: Passeriformes). – *Biological Journal of the Linnean Society*, **98** (2): 422–438.
- JOHANSEN, H. (1944): Die Vogelfauna Westsibiriens. II. Teil. Systematik und Verbreitung, Oekologie und Biologie der Einzelarten. Fringillidae. Artengruppe *Emberiza citrinella* – *E. leucocephalos*. – *Journal für Ornithologie*, **92** (1/2): 1–105.
- KLEINSCHMIDT, O. (1903): [Plate VII for *Emberiza citrinella* and *Emberiza leucocephalos*]. – *Journal für Ornithologie*, **51**.
- KLICKA, J. & ZINK, R.M. (1997): The importance of recent ice ages in speciation: a failed paradigm. – *Science*, **277** (5332): 1666–1669.

- LÖHRL, H. (1967): Zur Verwandtschaft von Fichtenammer (*Emberiza leucocephala*) und Goldammer (*Emberiza citrinella*). – Vogelwelt, **88** (4/5): 148–153.
- MARTENS, J. (1993): Lautäußerungen von Singvögeln und die Entstehung neuer Arten. – Forschungsmagazin Johannes Gutenberg-Universität Mainz, **9** (2): 34–44.
- MARTENS, J. (1996): Vocalizations and speciation of Palearctic birds. In: KROODSMA, D. E. & MILLER, E. H. (eds.): Ecology and evolution of acoustic communication in birds. – Cornell University Press, Ithaca, NY, pp 221–240.
- MARTENS, J., TIETZE, D.T. & PÄCKERT, M. (2011): Phylogeny, biodiversity, and species limits of passerine birds in the Sino-Himalayan region – a critical review. – Ornithological Monographs, **70**: 64–94.
- MATESSI, G., DABELSTEEN, T. & PILASTRO, A. (2000): Responses to playback of different subspecies songs in the Reed Bunting *Emberiza schoeniclus*. – Journal of Avian Biology, **31** (1): 96–101.
- MATESSI, G., DABELSTEEN, T. & PILASTRO, A. (2001): Subspecies song discrimination in a Mediterranean population of the Reed Bunting *Emberiza schoeniclus intermedia*. – Italian Journal of Zoology, **68** (4): 311–314.
- MAUERSBERGER, G. (1971): Ist *Emberiza leucocephalos* eine Subspecies von *E. citrinella*? – Journal für Ornithologie, **112** (2): 232–233.
- PÄCKERT, M. & MARTENS, J. (2004): Song dialects on the Atlantic islands: goldcrests of the Azores (*Regulus regulus azoricus*, *R. r. sanctae-mariae*, *R. r. inermis*). – Journal of Ornithology, **145** (1): 23–30.
- PÄCKERT, M., MARTENS, J. & HOFMEISTER, T. (2001): Lautäußerungen der Sommergoldhähnchen von den Inseln Madeira und Mallorca (*Regulus ignicapillus madeirensis*, *R. i. balearicus*). – Journal für Ornithologie, **142** (1): 16–29.
- PÄCKERT, M., MARTENS, J., NAZARENKO, A.A., VALCHUK, O.P., PETRI, B. & VEITH, M. (2005): The Great Tit (*Parus major*) – a misclassified ring species. – Biological Journal of the Linnean Society, **86** (2): 153–174.
- PANOV, E.N., ROUBTSOV, A.S. & MONZIKOV, D.G. (2003): Hybridization between Yellowhammer and Pine Bunting in Russia. – Dutch Birding, **25** (1): 17–31.
- PANOV, E.N., ROUBTSOV, A.S. & MORDKOVICH, M.V. (2007): New data on interrelationships between two bunting species (*Emberiza citrinella*, *Emberiza leucocephalos*) interbreeding in the zone of their overlapping ranges. – Zoologicheskii Zhurnal, **86** (11): 1362–1378.
- PORTENKO, L.A. & STÜBS, J. (1971): *Emberiza citrinella* L. und *Emberiza leucocephala* Gmelin. In: STRESEMANN, E., PORTENKO, L. A. & MAUERSBERGER, G. (eds.): Atlas der Verbreitung paläarktischer Vögel, 3. Lieferung. – Berlin.
- QUESTIAU, S., EYBERT, M.-C., GAGINSKAYA, A. R., GIELLY, L. & TABERLET, P. (1998): Recent divergence between two morphologically differentiated subspecies of bluethroat (Aves: Muscicapidae: *Luscinia svecica*) inferred from mitochondrial DNA sequence variation. – Molecular Ecology, **7** (2): 239–245.
- R DEVELOPMENT CORE TEAM (2010): 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>.
- RADZHABLI, S.I., PANOV, E.N. & BULATOVA, N.S. (1970): Comparative studies of the karyotypes of two closely related species of buntings (*Emberiza citrinella* L. and *E. leucocephalos* Gm.) hybridised in areas of overlap. – Zoologicheskii Zhurnal, **49** (12): 1857–1863.
- RUBTSOV, A.S. (2007): The variation of songs in the Yellowhammer (*Emberiza citrinella*) and Pine Bunting (*Emberiza leucocephala*) as an evidence of the population structure and evolutionary history of the species. – Zoologicheskii Zhurnal, **86** (7): 863–876.
- SCHÖN, R. (1989): Dialekte, Individualität und Gesangslernen bei der Goldammer (*Emberiza citrinella* L.). – Dissertation, Wien.
- SECONDI, J., BRETAGNOLLE, V., COMPAGNON, C. & FAIVRE, B. (2003): Species-specific song convergence in a moving hybrid zone between two passerines. – Biological Journal of the Linnean Society, **80** (3): 507–517.
- SHORT, L.L. (1969): Taxonomic aspects of avian hybridization. – Auk, **86** (1): 84–105.
- SPECHT, R. (2005): Avisoft-SASLab Pro (Sound Analysis and Synthesis Laboratory). Vers. 4.37d. – Specht, Berlin.
- THIELCKE, G. (1969): Die Reaktion von Tannen- und Kohlmeise (*Parus ater*, *P. major*) auf den Gesang nahverwandter Formen. – Journal für Ornithologie, **110** (2): 148–157.
- THIELCKE, G. (1986): Waldbaumläufer (*Certhia familiaris*) singen bei Sympatrie mit dem Gartenbaumläufer (*C. brachydactyla*) nicht kontrastreicher. – Journal für Ornithologie, **127** (1): 43–49.
- TIETZE, D.T., MARTENS, J., SUN, Y.-H. & PÄCKERT, M. (2008): Evolutionary history of treecreeper vocalisations (Aves: *Certhia*). – Organisms, Diversity & Evolution, **8** (4): 305–324.
- TIETZE, D.T., MARTENS, J., SUN, Y.-H., SEVERINGHAUS, L.L. & PÄCKERT, M. (2011): Song evolution in the Coal Tit (*Parus ater*). – Journal of Avian Biology, **42** (3): 214–230.
- WALLSCHLÄGER, D. (1983): Vergleich von Gesangsstrukturen zentralasiatischer Ammern (*Emberiza*). – Mitteilungen aus dem Zoologischen Museum in Berlin, **59** (Supplement 7): 85–116.
- WALLSCHLÄGER, D. (1998): Was ist ein Goldammerdialekt? – Brandenburgische Umwelt-Berichte, **3**: 59–66.
- WICKLER, W. (1986): Dialekte im Tierreich. Ihre Ursachen und Konsequenzen. – Aschendorff, Münster: IV, 84 pp.
- WONKE, G. (2008): Gesangsdialekte bei der Goldammer „*Emberiza citrinella*“. Kultur als Mechanismus der Artaufspaltung? – WiKu-Verlag, Duisburg: XIX, 118, XLI pp.
- WONKE, G. & WALLSCHLÄGER, D. (2009): Song dialects in the Yellowhammer *Emberiza citrinella*: bioacoustic variation between and within dialects. – Journal of Ornithology, **150** (1): 117–126.