

The Humane Society Institute for Science and Policy
Animal Studies Repository

2013

Does True Syntax or Simple Auditory Object Support the Role of Skylark Song Dialect?

Elodie F. Briefer

University of Paris-Sud, elodie.briefer@usys.ethz.ch

Fanny Rybak

University of Paris-Sud

Thierry Aubin

University of Paris-Sud

Follow this and additional works at: http://animalstudiesrepository.org/acwp_asie

 Part of the [Animal Studies Commons](#), [Comparative Psychology Commons](#), and the [Other Animal Sciences Commons](#)

Recommended Citation

Briefer, E. F., Rybak, F., & Aubin, T. (2013). Does true syntax or simple auditory object support the role of skylark song dialect?. *Animal Behaviour*, 86(6), 1131-1137.

This Article is brought to you for free and open access by the Humane Society Institute for Science and Policy. It has been accepted for inclusion by an authorized administrator of the Animal Studies Repository. For more information, please contact eyahner@humanesociety.org.

Does True Syntax or Simple Auditory Object Support the Role of Skylark Song Dialect?

Elodie F. Briefer¹, Fanny Rybak^{1,2}, Thierry Aubin^{1,2}

¹ University of Paris-Sud

² French National Center for Scientific Research

KEYWORDS

Alauda arvensis, dialect, playback experiment, skylark, songbird, temporal ordering

ABSTRACT

*Parallels between birdsong and human language are numerous and include particular temporal arrangements of acoustic units and the existence of dialects. In animal communication, modifications of the temporal ordering of existing acoustic units have rarely been clearly linked with changes in information content, particularly in a natural environment. Here, we show that the organization of birdsong units ('syllables') in sequences supports interindividual relationships within skylark communities. We manipulated the temporal arrangement of song dialect variants ('shared phrases') in the skylark, *Alauda arvensis*, a songbird with a very large repertoire of syllables and complex song. When tested with playback experiments performed in the field, skylarks were able to perceive subtle differences in the ordering of syllables. Modifications of the syllable ordering within shared phrases changed the information content from 'group member' to 'unfamiliar individual' and induced more aggressive reactions than shared phrases with a preserved syllable arrangement. Shared phrases often varied between individuals in the number of successive repetitions of similar syllable types, but were very consistent in terms of syllable type ordering. Our results indicate that skylarks rely not simply on the composition in syllable types of shared phrases to recognize group members, but on syllable type ordering. Shared phrases could be perceived by birds as 'auditory objects' embedded within songs. Alternatively, birds might identify incorrect phrases using grammatical rules governing the succession of syllables composing the phrases shared by their group. The presence of between-individual variation in phrase length, associated with consistent syllable type ordering revealed by our analysis, suggests that the latter hypothesis is more likely. Our results show that birds perceive disruptions in the natural temporal pattern of song units, and that this temporal pattern is behaviourally salient and carries information.*

Striking genetic, neural and behavioural similarities exist between vocal learning in birds and human infants. This includes innate predispositions to learn species-specific signals, a sensitive period in early life and a memorization phase followed by a motor phase of production. It also includes the necessity for auditory feedback during the memorization phase, complex neural substrates and a role of social interactions (Doupe & Kuhl 1999; Beecher & Burt 2004; Beecher & Brenowitz 2005; Nottebohm 2005; Bolhuis et al. 2010). As a consequence of vocal learning, birdsong is characterized, like human language,

by geographical variation in element composition called 'dialects' (Catchpole & Slater 1995). Furthermore, the songs are computational, that is, they consist of discrete song units, named syllables, organized in a particular temporal arrangement, in a similar way to how phonemes are organized in a particular order to form words ('phonology'; Yip 2006). This temporal arrangement of elements into larger units in birds has been termed 'phonological syntax' (Marler 1977).

A poorly explored parallel between human language and birdsong is the importance of the ordering of syllables composing natural songs for their information content. Are birds able to distinguish and react to sequences composed of syllables in an incorrect order, and is the organization of syllables crucial for the behaviourally relevant information carried by songs? There is evidence in primates that adding a particular suffix before an alarm call (Zuberbühler 2002; Ouattara et al. 2009a) or combining calls (Arnold & Zuberbühler 2006; Ouattara et al. 2009b) changes the meaning of the sequence. In swamp sparrows, *Melospiza georgiana*, adding an extra note at a particular position within a sequence of three notes constituting a syllable, which transformed the syllable of a dialect area to the syllable characteristic of another dialect area, triggered more aggressive responses in males than unmodified song (Balaban 1988). However, in birds, as in other species producing complex vocalizations, modifications of the temporal arrangements of existing discrete acoustic units ('broken syntax'; e.g. changing 'a b c d e' into 'b d e c a') have all been shown to induce a similar response to unmodified vocalizations, or to decrease or suppress the response, but never to increase it (Becker 1982; Mitani & Marler 1989; Holland et al. 2000; Nowicki et al. 2001; Clucas et al. 2004; Gentner 2008; Dahlin & Wright 2012). Any change in response indicates that the syntax modification is perceived by the animals. Yet, a suppression of response or decrease in response could potentially indicate that the broken syntax signal is not fully recognized as a conspecific signal, and is perceived as 'unnatural' and therefore irrelevant. By contrast, an increase in response would provide clear evidence that the temporal arrangement modification has triggered a change in information content.

We examined the role of the temporal ordering of syllables within songs in a species with a very large repertoire (>100 syllables) and complex songs, the skylark, *Alauda arvensis*, in its natural habitat. Male skylarks produce one of the longest and most complex songs among oscines (Briefer et al. 2010a). In this species, during the breeding season, pairs settle in adjacent territories of about 1 ha each. Because of the heterogeneity of the habitat, birds are gathered in small groups (range 5e19 males per group; Briefer et al. 2008a, b) separated from other groups by unsuitable habitat. Within each group, males ('neighbours') remain in their territory to breed and forage at short distances from their nest (Donald 2004). They defend these stable territories by producing long and continuous songs to deter intruders. In these songs, species identity is encoded in temporal parameters, and particularly in the rhythm (Aubin & Brémond 1983). A mean duration of 150 s that can be extended to more than 40 min, and a repertoire size of more than 300 different syllables, provides skylark song with a huge potential for variation in temporal arrangement (Aubin 1982; Briefer et al. 2008a, 2010a). When boundaries between territories are stable in the middle of the breeding season, neighbours establish 'dear-enemies' relationships (Temeles 1994), reacting weakly to a territory intrusion by their familiar neighbours and aggressively to an intrusion by unfamiliar individuals established in other distant groups ('strangers'; Briefer et al. 2008b).

Recently, we showed that neighbouring males share many syllables (on average 83% of the syllable repertoire, corresponding to about 269 types of syllable) and many sequences of syllables, named phrases (on average 71% of the sequence repertoire, or 44% of the total song duration, corresponding to about 91 types of sequences), in their songs. Phrases shared by neighbours differ from nonshared sequences by having shorter intersyllable intervals and fewer repetitions of syllables (Briefer et al. 2008a). Males identify and recognize these particular phrases to achieve neighbour-stranger discrimination, indicating that shared phrases support a group signature and serve as a basis for the

dear-enemy effect (Briefer et al. 2008a). These phrases might also allow recognition of distant neighbours within the same dialect area. Indeed, adjacent and distant neighbours (established two to three territories away) within a group share a similar number of syllables and phrases in their songs (Briefer et al. 2010b). Two-by-two comparisons showed that pairs of neighbours share 216 syllable types and 59 different sequences on average, compared to 105 syllable types and only 0.5 sequences for pairs of strangers. Furthermore, phrases shared by at least two neighbours in a group are never found in the songs of strangers. Between two distinct groups, phrases may thus be composed of a few similar syllable types, but not with the same phonology, making the temporal arrangement of syllables constituting one group signature unique (Briefer et al. 2008a).

In this study, we assessed between-individual variation and complexity of sequential arrangements of syllables within shared phrases and, using playback experiments, we tested the hypothesis that these sequential arrangements are crucial for group recognition. We predicted that, if birds rely on single syllables composing the shared phrases to identify group members, they should react in a similar way to phrases containing syllables in their natural order as they do to phrases in which the phonological syntax has been modified. Conversely, if birds perceive shared phrases as 'auditory objects' (i.e. a set of acoustic events that can be perceived as a whole; Gentner 2008), or if they rely on the temporal ordering of syllables composing shared phrases (i.e. on the phonological syntax), they should not be able to recognize shared phrases with syllables presented in a modified order (i.e. with broken syntax) and therefore react as if these sequences were produced by a stranger.

METHODS

Study Area, Subjects and Song Recordings

We carried out our recordings for the song analysis and our playback experiment in the fields surrounding the University of Paris South, France. The songs for the analysis had been recorded, as part of a previous study (Briefer et al. 2010a), in four locations (i.e. four groups) during the breeding season in 2005 from the middle of February to the end of March. In total, the songs of 11 territorial males were analysed. The playback experiment was carried out in three additional locations (i.e. three groups) during two successive breeding seasons in 2007 and 2008. Playbacks were conducted in May, when territory borders are stable and males consider their neighbours as 'dear enemies' (Briefer et al. 2008b). In total, we tested 17 territorial males. Within a group, males are established in adjoining and stable territories of about 1 ha and are referred to as 'neighbours'. The total number of neighbours (group size) within each of the seven groups of our study was 10.43 ± 1.13 (mean \pm SE; range 7-15 territorial males). We refer to males from two different locations situated at least 2 km apart as 'strangers' (Briefer et al. 2008a, 2010a). We estimated the territory boundaries of the studied birds after numerous observations of the birds' movements at different times of day, and recorded GPS coordinates at the centre of each territory (Garmin GPSMAP 76S). We recorded several songs per individual between 0900 and 1200 hours Eastern Daylight Time using a Marantz PMD 690 numeric recorder (sampling rate: 48 kHz) connected to a Sennheiser ME 64 K6 omnidirectional microphone (frequency response: 30 Hz to 20 kHz \pm 1 dB) mounted on a Telinga Universal parabola (diameter: 50 cm). Song files were transferred to a computer and high-pass filtered (cutoff frequency: 1600 Hz) to remove background noise. We used Avisoft SASLab pro version 4.31 software (Avisoft Bioacoustics, Berlin, Germany) and Seewave (Sueur et al. 2008) for subsequent analyses and for the preparation of songs played back.

Song Analysis

We carried out a song analysis on 100.03 ± 0.02 s of continuous song per individual ($N = 11$) to measure between-individual variation in syllable number (i.e. total number of syllables) and in syllable type ordering

(i.e. how the different types of syllable, which are characterized by different overall frequency modulation shapes, are temporally ordered) within shared phrases. Then, we measured the complexity of shared phrases, by calculating the number of different shared phrases per group and the number, the diversity and the organization of syllables within these phrases.

Shared phrases used in this study had been previously identified, as part of another study (Briefer et al. 2010a), by labelling all the syllables in the songs, then using a custom MatLab program (Mathworks, Natick, MA, U.S.A.) to obtain the repertoire of all syllables and sequences of syllables produced by each individual, and subsequently comparing sequence repertoires of individuals of the same group to find shared phrases (Briefer et al. 2008a, 2010a). They were defined as sequences, which were composed of at least two different syllable types, and which were produced by at least two individuals of the same group (Briefer et al. 2008a). We assumed that these phrases were also potentially shared by the rest of the group ('group signature').

First, we quantified the variation between individuals in terms of length of phrases and number of repetitions of syllable types. Between-individual comparisons were carried out on one phrase per individual for each phrase type. For cases of within-individual variation in phrase length, when a phrase was found twice in the song of a given individual, we selected the longest sequence for its shared phrase repertoire. Then, for each group, we selected one shared phrase of each type for further measures of complexity. When some individuals in the group produced only parts of the phrase type (e.g. 'abc' instead of 'abcde'), we selected the longest shared sequence for further measurements (e.g. 'abcde' produced by at least two birds). For cases of between-individual variation in number of consecutive syllable repetitions (e.g. 'abbc' versus 'abbbbbc'; 'abcd' versus 'abcbcbcd' or even 'abcde' versus 'abcdbcdc'), we selected only the shortest sequence of syllables (e.g. 'abbc'; 'abcd' or 'abcde', respectively, for the examples listed above). For each selected shared phrase type, we then calculated the transition versatility (Briefer et al. 2010a), in percentages, by dividing the number of transitions occurring between two syllable types by the total number of transitions in the phrase [i.e. number of syllables in the phrase minus 1; range 0-100%; 100% indicating that all transitions in the phrase are between different syllables; e.g. the transition versatility is 33.3% ($2/6 \times 100$) for 'abbbbbc' and 100% ($7/7 \times 100$) for 'abcbcbcd'].

Song Treatments

We tested the importance of the temporal arrangement of syllables within shared phrases in neighbour-stranger discrimination by comparing the responses of birds to (1) a natural song of a stranger (i.e. a bird established in another distant group; S song), as a control; (2) the same stranger song including 30% of artificially inserted phrases shared within the group of the subjects tested ('chimeric song', C song); (3) the same chimeric song, in which the order of the syllables composing the artificially inserted shared phrases had been modified ('chimeric modified song', Cm song; Table 1, Figs 1, 2 and Supplementary material). In this latter type of song, syllables composing each shared phrase inserted had been mixed up at random.

We prepared the three categories of songs (S, C and Cm) as follows. We selected songs from our recordings and adjusted all of them to the same duration by taking the first 90 s (Fig. 2). This song duration is within the natural range (a few seconds to up to 40 min; Hedenström 1995) and was shown to be sufficient to allow neighbour-stranger and even neighbour-neighbour discrimination (Briefer et al. 2008a, b, 2009, 2010b). Using Goldwave version 5.11 (www.goldwave.com), we rescaled each recorded song to match the root mean square amplitude of the different songs at the same output level. For a given group, we selected good-quality S songs (six different S songs in total for the three locations) recorded in another distant group (5-10 km away). Each selected S song was used to prepare the corresponding C and Cm songs, resulting in six different C and Cm songs for the three locations. For this purpose, 30%

($30.12 \pm 0.21\%$, corresponding to 27.11 ± 0.19 s) of the total duration of the S song was replaced by an equivalent duration of phrases shared by birds of the tested group (Fig. 2). Slightly fewer shared phrases than the number found in natural songs ($43.5 \pm 2.6\%$ of the total song duration) were inserted, because our aim was not to transform the stranger song into a neighbour song, but instead to assess the response of the birds to shared phrases (Briefer et al. 2008a). Our previous analyses revealed very weak individuality in acoustic parameters (frequency and duration) of sequences, indicating that similar types of syllable sequences have the same acoustic characteristics even if they have been produced by different birds (Briefer et al. 2009). As the purpose of the experiment was to reveal the group signature, shared phrases were therefore extracted from good-quality songs of three birds of the group ($N = 12.33 \pm 0.33$ phrases per C or Cm song, range 12-13 different phrases of 2.20 ± 0.21 s each; Fig. 2). This ensured that the responses to C and Cm songs would not be specific to one given group member. They had been previously identified by visually comparing songs produced by at least three birds of each location. Cm songs were prepared in the following way: each syllable composing the artificially inserted shared phrase was removed manually, from the beginning of the syllable to the end of the silence immediately following, and replaced in a random position within the new phrase (Fig. 1 and Supplementary material). In this way, we did not alter the average sound to silence ratio (rhythm) in the phrase, which is a key parameter of the species-specific coding (Aubin & Brémond 1983).

Figure 1. Song treatments. Sound spectrograms of (a) a part of a stranger song, (b) a part of a chimeric song with an artificially inserted shared phrase, and (c) a part of a chimeric modified song with an artificially inserted shared phrase, whose syllable ordering had been modified (total duration of song treatments = 90 s; FFT length = 256; frame = 100%; window = Hanning). Syllables composing the shared phrase are indicated with a number. These sounds are available as audio files in the Supplementary material.

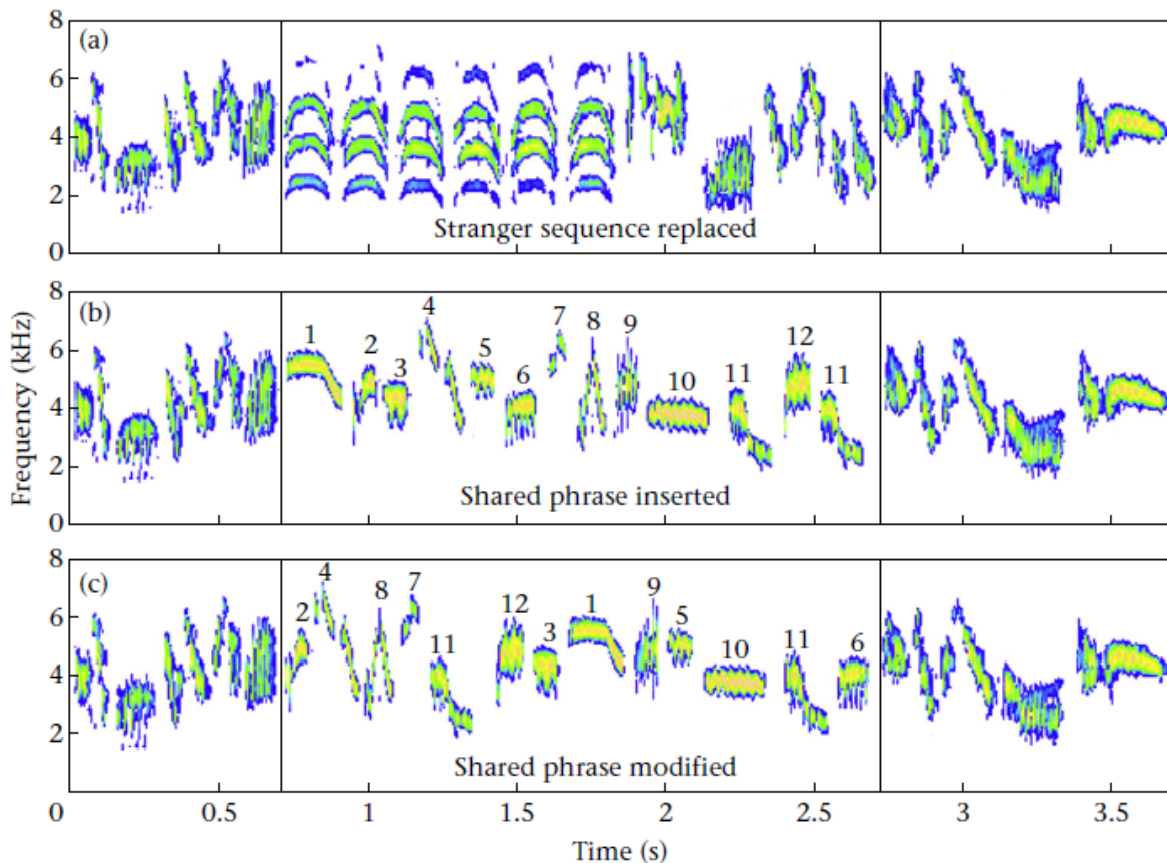


Figure 2. Chimeric song. Sound spectrograms of a stranger song (duration = 90 s), in which 30% of the total duration was replaced by an equivalent duration of phrases shared by birds of the group (grey rectangles). The rest of the song is composed of natural (i.e. unmodified) sequences of the stranger song.

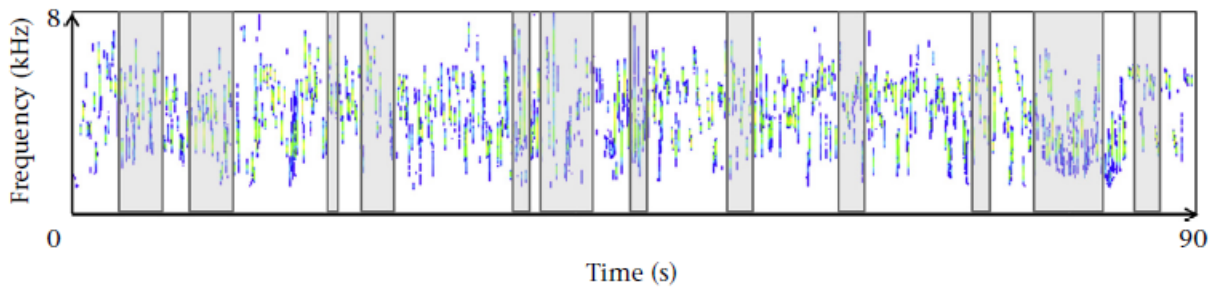


Table 1. Examples of shared phrases

Group	Phrase Type	Individual	Syllable sequence
1	A	1-1	a b c d e f g f g a h h
		1-2	a a b c d e f g f g f g a a
		1-3	a b c d e f g f g a h
		1-Cm	d f e g h a d f d b g
1	B	1-1	a a b b c d d d c d d e f e f g h i j j
		1-2	a a b b c d d d e f g h i j e f e f g
		1-3	c d d d e f e f g h i j j
		1-Cm	i f d c e h j g d f j e d
2	C	2-1	a a a a a b c b c b c a d d d e f e f e f g
		2-2	b c a a b c a d d e f e f g g
		2-3	c c b c a a d d e f e f g g
		2-Cm	a g b c f d g e c f a c e d
3	D	3-1	a b c d e f g h i j k l
		3-2	a b c d e f g h i j k m k l
		3-3	a b c d e f g h i j k m k l
		3-Cm	f l c a g j d k i h b e m k

Four types of shared phrases (A-D) from three different groups (1-3), in which the playback experiments were carried out, along with the corresponding modified sequence inserted in the stranger song to make chimeric modified song (Cm). For a given group, the syllable sequence composing a phrase shared by three individuals is indicated. To simplify the table, syllable types are indicated by letters a-m. According to this coding, the same letters within a type of phrase (A-D), but not between phrase types, indicates the same syllable type. None of the syllable types presented here was shared between two groups.

Playback Procedure

We played back experimental songs between 0900 and 1200 hours Eastern Daylight Time, with a Marantz PMD 690 digital recorder connected via a 20 m cable to a 10W Megavox Pro mega-6000 loudspeaker (frequency response: 400 Hz to 10 kHz \pm 3 dB), at the intensity estimated to be normal for the birds (90.8 ± 0.8 dB sound pressure level measured at 1 m from the loudspeaker with a Brüel & Kjaer 2235 sound level meter, linear setting). The loudspeaker was positioned on the ground at approximately 5 m inside the territory of the subject. The experimenter stood at 20 m from the loudspeaker. The playback was initiated when the subject was standing on the ground inside its territory at more than 10 m from the

loudspeaker (ca. 10e80 m away), just after it had been seen singing and when adjacent neighbours were quiet. The three categories of songs (S, C and Cm) were broadcast to each subject on the same day, to avoid any effect of change in the environment on birds' responses, in a random order of presentation separated by at least 10 min intervals. This time interval allowed the birds to return to normal activities (Briefer et al. 2008a, b, 2009).

Responses Measured and Statistical Analysis

For each trial, the response of the bird was scored during 180 s, corresponding to the broadcast of 90 s of continuous song and 90 s of postplayback observations. The skylark male displays a very strong territorial behaviour with stereotyped patterns, which are easy to observe (Delius 1963). It reacts vigorously against territorial intrusion by flying towards the intruder and by landing in its vicinity or flying low over it. It then takes up a fight posture, head and crest up, and makes threat calls. Thus, we chose to score the five measures of response listed in Table 2 to assess the effects of the different song treatments (Briefer et al. 2008a, b, 2009, 2010b). Only playbacks for which the subject was observed moving were considered, as an absence of reaction could indicate that it did not hear the song being broadcast. When the subject did not approach the loudspeaker at less than 10 m or 5 m during the whole trial, we gave a value of 180 s (corresponding to the total trial duration) to the latency measures of response for the playback. We used a principal components analysis (PCA) based on the correlation matrix to create a composite score with the measures of responses, which are likely to be correlated (McGregor 1992). The scores of the principal components with an eigenvalue greater than 1 (Kaiser's criterion) were compared using linear mixed-effect models (LMM), including as fixed effects the order in which the treatments were played back (1-3) to control for a potential order effect, the S song exemplar used to prepare the C and Cm songs to control for a potential effect of song exemplars on responses, and the treatment (S, C and Cm). Subject identity was included as an error term to account for repeated measurements of the same individuals across treatments. A sequential Bonferroni adjustment was used for two-by-two comparisons and results retained significance when $P < 0.016$ ($0.05/3$). Statistical analyses were carried out using R v.2.15.0 (R Development Core Team 2012). All means are given \pm SE.

Table 2. Factor loadings of the behaviours scored on the first principal component

Statistics and response measures	PC1
Eigenvalue	2.60
Percentage of variance explained	51.98
Total duration of movements	0.851
Duration of movements at less than 10 m	0.740
Latency to approach at less than 5 m	-0.733
Total number of threat calls	0.704
Latency to approach at less than 10 m	-0.541

Eigenvalue, variance explained and loadings of the measures of response on the first principal component (PC1) for playbacks of stranger, chimeric and chimeric modified songs are shown.

RESULTS

Song Analysis

The analysis (100.03 ± 0.02 s of continuous song per individual; $N = 11$), corresponding to a total of 503.64 ± 9.01 syllables per individual, revealed 24.75 ± 4.15 (range 18-35) different shared phrases per

group ($N = 99$ phrase types for the four groups; 80 phrase types shared by two individuals and 19 phrase types shared by three individuals, making a total of 217 phrases).

Variation in syllable number

Some individuals sometimes produced only parts of a phrase (e.g. 'abc' instead of 'abcde'; 9/217 cases). Furthermore, phrases varied in their number of consecutive syllable repetitions between individuals, from one to six successive repetitions of the same syllable type (e.g. 'abc' versus 'abbbbbc'), from one to three repetitions of the same sequence of two syllable types (e.g. 'abcd' versus 'abcbbcdd') and from one to two repetitions of the same sequence of three syllable types (e.g. 'abcde' versus 'abcdbcde'; see Table 1 for examples of variation). These two types of variation (i.e. phrases not produced entirely and number of successive repetitions) resulted in phrases that could contain up to 4.5 more syllables in one individual compared to its neighbour(s) (average minimum-maximum difference = 1.24 ± 0.05 times longer; $N = 99$ phrases). In total, 40.4% of the shared phrases varied in length between individuals.

Variation in syllable type ordering

The order of syllable types within phrases was always consistent between individuals, that is, each syllable type could be followed by only one given other syllable type (e.g. 'a' is always followed by 'b'), except in 12 cases involving nine phrases (9.09% of the phrases, $N = 99$ phrases), for which one syllable type could be followed by one of two different syllable types (e.g. 'a' can be followed by either 'b' or 'c'). This mainly happened (except in one case) when two syllable types were repeated (e.g. in 'abcbcd', 'c' can be followed by either 'b' or 'd').

Shared phrase complexity

Shared phrases selected for complexity analysis (i.e. phrases produced by one of the individual in the group, which contained the largest number of syllable types, but the lowest number of repetitions of similar syllable types) were composed of 2-19 (5.21 ± 0.37) syllables in total. Among these syllables composing a phrase, 2-13 (3.88 ± 0.24) belonged to different syllable types ($N = 99$ phrases, 11 individuals). In total, all the shared phrases in a group corresponded to 80.75 ± 10.22 (range 61-103) different syllable types ($N = 4$ groups). Within phrases, $85.80 \pm 2.44\%$ (range 14.29-100.0%; $N = 99$ phrases, 11 individuals) of the syllable successions were between two different syllable types ('transition versatility'; Briefer et al. 2010a).

Playback Experiment

Only the first principal component (PC1) of the PCA was considered for further analyses (eigenvalue > 1). PC1 explained 52% of the variance in the responses (Table 2). Examination of the loadings of the response measures on PC1 (Table 2) revealed that the duration of movements and the number of calls were correlated positively with PC1, whereas latencies were correlated negatively with PC1. Therefore, more negative PC1 scores indicated less aggressive responses and more positive PC1 scores indicated more aggressive responses; birds spent more time moving, approached closer to the loudspeaker after a shorter latency and produced more calls. A comparison of PC1 scores showed that responses were significantly different depending on the song treatment (LMM: $F_{2,31} = 9.09$, $P = 0.0008$). The effects of treatment order (LMM: $F_{1,31} = 3.65$, $P = 0.07$) and song exemplar (LMM: $F_{4,12} = 0.51$, $P = 0.73$) on PC1 scores were not significant. Two-by-two comparisons revealed that subjects responded significantly more strongly to stranger songs and chimeric modified songs than to chimeric songs (LMM controlled for treatment order and song exemplar: S song versus C song: $F_{1,15} = 8.56$, $P = 0.010$; Cm song versus C song: $F_{1,15} = 14.54$, $P = 0.002$; Fig. 3). Responses to stranger songs and chimeric modified songs did not differ significantly (LMM: $F_{1,15} = 1.96$, $P = 0.18$; Fig. 3).

DISCUSSION

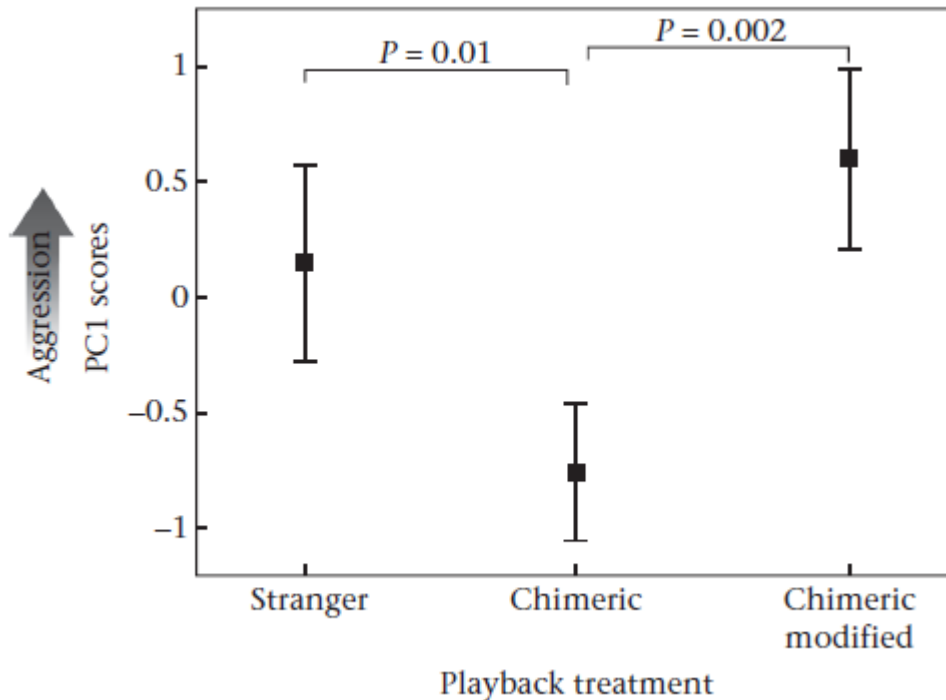
Using playback experiments, we tested the hypothesis that the temporal arrangements of syllables within skylark song are crucial for group recognition. Our results showed an increase in aggressiveness to songs containing shared phrases (dialect) whose syllable order had been artificially modified (chimeric modified song), compared to the same songs with an unmodified syllable order (chimeric song). The responses triggered by chimeric modified songs were similar to the responses triggered by the natural songs of strangers (i.e. unfamiliar individuals established in a distant group). This indicates that these artificially modified songs were perceived by birds as natural songs of conspecifics, and not as interspecific songs, which would have elicited a weaker or null response (Aubin & Brémond 1983). Therefore, modifications of the temporal ordering of syllables composing shared phrases appeared to change the information content from 'group member' (familiar and nonthreatening individual) to 'stranger' (unfamiliar and threatening individual). These results indicate that skylarks do not rely simply on single syllables composing shared phrases to recognize group members, but on syllable temporal ordering. Shared phrases could be perceived as 'auditory objects', that is, a set of acoustic events that can be perceived as a whole (Gentner 2008). Alternatively, birds might identify incorrect phrases using grammatical rules governing the succession of syllable types composing the shared phrases (Abe & Watanabe 2011). Our song analysis revealed that the ordering of syllable types within shared phrases was very consistent between birds. Conversely, the number of successive repetitions of similar syllable types and groups of two or three syllable types varied, resulting in variation in the length of phrases (up to 4.5 times longer). This variation in phrase length associated with the consistency in syllable type ordering between birds suggests that skylarks could rely on simple grammatical rules governing the succession of adjacent syllable types, more than on auditory objects, to recognize group members. This represents a clear link between temporal composition of acoustic units and information content in animal communication studied in a natural environment.

We showed that group identity is encoded in phrases shared by neighbours and that males use the sequential arrangement of syllables within these phrases to discriminate group members from nongroup members. Despite the fact that each phrase lasted 2.2 s on average and was inserted randomly within the stranger song to prepare the chimeric and chimeric modified songs (Fig. 2), skylarks were still able to identify them in the 90 s continuous songs played back, and reacted more aggressively when syllables were not in the order corresponding to their own dialect. This indicates that males can distinguish between correct (familiar and memorized) and incorrect (not familiar and not memorized) syllable order within complex songs composed of a continuous succession of many different syllables.

If skylarks were memorizing only the sets of single syllables composing each shared phrase, they should be able to perceive and react to shared phrases, independently of whether the syllable order has been altered or not. Our results show that this is not the case and that the temporal ordering of syllables composing shared phrases is crucial for skylarks to recognize their neighbours. Shared phrases could be perceived by birds as 'auditory objects' embedded within songs. Auditory objects are defined as 'a set of coincident, or closely coincident, acoustic events that can be perceived as a whole, and that carries with it behaviourally relevant information' (Gentner 2008, page 9). Modifying an auditory object changes/suppresses its information content, in the same way as modifying the ordering of phonemes in human language changes/suppresses the meaning of a word. Similarly, modifying the note order in a musical composition changes the melody. Further playback experiments using chimeric modified songs with shared phrases that have been more or less modified, for example from a permutation or deletion of a few syllables to a modification of the entire phrase, could give indications about the lower bound for the length of component features that skylarks can perceive as relevant. Such an experiment could bring

useful knowledge about how the birds memorize, perceive and use the temporal order of song units in their natural environment.

Figure 3. Responses to playbacks. Mean \pm SE first principal component scores ($N = 17$) for playbacks of stranger, chimeric and chimeric modified songs. Higher positive PC1 scores correspond to higher levels of aggression.



Our song analysis revealed that phrases shared by individuals were composed of syllable types produced in the same order. Indeed, 90.9% of the phrases contained syllable types produced in the exact same order between birds of a group, with exceptions happening when groups of two syllable types were repeated. Conversely, phrases could differ in the number of consecutive repetitions of similar syllable types or groups of syllable types (e.g. 'abbc' versus 'abbbbbc'; 'abcd' versus 'abcdbcdbcd' or even 'abcde' versus 'abcdcbde'). Some individuals sometimes also produced incomplete phrases. These two kinds of variation concerned 40% of the phrases and resulted in phrases that could be up to 4.5 times longer in some individuals than others. If shared phrases are considered as auditory objects, we could hypothesize that skylarks still recognize them despite their variation in length.

Alternatively, because the ordering of syllable types, unlike the phrase length, was very consistent between males, a more likely hypothesis would be that skylarks rely on simple rules governing the succession of pairs of syllables. These rules could be described as a 'finite-state' grammar, a system that is limited to local dependencies between units. Indeed, in most cases, each syllable type was followed by only one given other syllable type. This is consistent with our previous findings, showing that skylark songs follow a first-order Markov chain, with each syllable being especially dependent on the syllable immediately preceding it (Briefer et al. 2010a). Furthermore, shared phrases from two distinct groups can be composed of a few similar syllable types, but never with the same arrangement (Briefer et al. 2008a). The rules governing the arrangement of syllables within shared phrases thus differ considerably between groups. Finite-state grammars are less complicated than the 'context-free' grammars (or 'phrase structure grammars') that can be found in human language (Chomsky 1957). Context-free grammars, unlike finite-

state grammars, include long-distance dependencies and recursive embedding of units of a particular category inside larger units of the same category. The existence of context-free grammars in nonhuman animal communication is still debated and requires further experiments (Hauser et al. 2002; Gentner et al. 2006; van Heijningen et al. 2009; Abe & Watanabe 2011; Bloomfield et al. 2011; Beckers et al. 2012). The ability of skylarks to discriminate syntactical rules (finite-state or context-free grammars) could be tested using operant conditioning techniques, with individuals trained to discriminate artificially synthesized syllable strings (Abe & Watanabe 2011; Beckers et al. 2012).

To conclude, our song analysis revealed an important variety in the production of syllable types within group signatures (86% of syllable successions are between two different syllable types). These signatures represent almost half of the song duration, corresponding to at least 60 s of song (Briefer et al. 2008a). Skylarks might therefore require considerable cognitive skills to learn, memorize and decode the temporal arrangement of such long series of syllables during vocal interactions or, more probably, to learn syntactical rules governing the ordering of syllables. According to two recent studies, the volume of the nucleus HVC (particularly large in the skylark) could be related to the song repertoire size (Moore et al. 2011), and the anterior nidopallium could be involved in the syntactical analysis of syllables (Abe & Watanabe 2011). Thus, the skylark constitutes a highly promising model to investigate how information coded by complex syllable arrangement can be studied in a natural communication context, and to help understand the evolution of nonhuman communication.

Acknowledgments

We are grateful to Alan McElligott, Peter Slater and the anonymous referees for helpful comments on the manuscript. This work was supported by the CNRS-University Paris 11. E.F.B. acknowledges support from the French Minister of Research and Technology during the data collection.

References

- Abe, K. & Watanabe, D. 2011. Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience*, 14, 1067-1074.
- Arnold, K. & Zuberbühler, K. 2006. Language evolution: semantic combinations in primate calls. *Nature*, 441, 303.
- Aubin, T. 1982. Habituation au chant territorial chez l'alouette des champs (*Alauda arvensis* L.). *Biology of Behaviour*, 7, 353-362.
- Aubin, T. & Brémond, J.-C. 1983. The process of species-specific song recognition in the skylark *Alauda arvensis*. An experimental study by means of synthesis. *Zeitschrift für Tierpsychologie*, 61, 141-152.
- Balaban, E. 1988. Bird song syntax: learned intraspecific variation is meaningful. *Proceedings of the National Academy of Sciences, U.S.A.*, 85, 3657-3660.
- Becker, P. H. 1982. The coding of species-specific characteristics in bird sounds. In: *Acoustic Communication in Birds* (Ed. by D. E. Kroodsma & E. Miller), pp. 213-252. New York: Academic Press.
- Beckers, G. J. L., Bolhuis, J. J., Okanoya, K. & Berwick, R. C. 2012. Birdsong neurolinguistics: songbird context-free grammar claim is premature. *Neuroreport*, 23, 139-145.
- Beecher, M. D. & Brenowitz, E. A. 2005. Functional aspects of song learning in songbirds. *Trends in Ecology & Evolution*, 20, 143-149.
- Beecher, M. D. & Burt, J. M. 2004. The role of social interaction in bird song learning. *Current Directions in Psychological Science*, 13, 224-228.
- Bloomfield, T. C., Gentner, T. Q. & Margoliash, D. 2011. What birds have to say about language. *Nature Neuroscience*, 14, 947-948.

- Bolhuis, J. J., Okanoya, K. & Scharff, C. 2010. Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews Neuroscience*, 11, 747-759.
- Briefer, E., Aubin, T., Lehongre, K. & Rybak, F. 2008a. How to identify dear enemies: the group signature in the complex song of the skylark *Alauda arvensis*. *Journal of Experimental Biology*, 211, 317-326.
- Briefer, E., Rybak, F. & Aubin, T. 2008b. When to be a dear enemy: flexible acoustic relationships of neighbouring skylarks, *Alauda arvensis*. *Animal Behaviour*, 76, 1319-1325.
- Briefer, E., Aubin, T. & Rybak, F. 2009. Response to displaced neighbours in a territorial songbird with a large repertoire. *Naturwissenschaften*, 96, 1067-1077.
- Briefer, E., Osiejuk, T. S., Rybak, F. & Aubin, T. 2010a. Are bird song complexity and song sharing shaped by habitat structure? An information theory and statistical approach. *Journal of Theoretical Biology*, 262, 151-164.
- Briefer, E., Rybak, F. & Aubin, T. 2010b. Are unfamiliar neighbours considered to be dear-enemies? *PLoS One*, 5, e12428.
- Catchpole, C. K. & Slater, P. J. B. 1995. *Bird Song, Biological Themes and Variations*. Cambridge: Cambridge University Press.
- Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton.
- Clucas, B. A., Freeberg, T. M. & Lucas, J. R. 2004. Chick-a-dee call syntax, social context, and season affect vocal responses of Carolina chickadees (*Poecile carolinensis*). *Behavioral Ecology and Sociobiology*, 57, 187-196.
- Dahlin, C. & Wright, T. 2012. Does syntax contribute to the function of duets in a parrot, *Amazona auropalliata*? *Animal Cognition*, 15, 647-656.
- Delius, J. D. 1963. Das Verhalten der Feldlerche. *Zeitschrift für Tierpsychologie*, 20, 297-348.
- Donald, P. F. 2004. *The Skylark*. London: T. & A.D. Poyser.
- Doupe, A. J. & Kuhl, P. K. 1999. Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*, 22, 567-631.
- Gentner, T. Q. 2008. Temporal scales of auditory objects underlying birdsong vocal recognition. *Journal of the Acoustical Society of America*, 124, 1350-1359.
- Gentner, T. Q., Fenn, K. M., Margoliash, D. & Nusbaum, H. C. 2006. Recursive syntactic pattern learning by songbirds. *Nature*, 440, 1204-1207.
- Hauser, M. D., Chomsky, N. & Fitch, W. T. 2002. The faculty of language: what is it, who has it, and how did it evolve? *Science*, 298, 1569-1579.
- Hedenström, A. 1995. Song flight performance in the skylark *Alauda arvensis*. *Journal of Avian Biology*, 26, 337-342.
- van Heijningen, C. A. A., De Visser, J., Zuidema, W. & Ten Cate, C. 2009. Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proceedings of the National Academy of Sciences, U.S.A.*, 106, 20538-20543.
- Holland, J., Dabelsteen, T. & Paris, A. L. 2000. Coding in the song of the wren: importance of rhythmicity, syntax and element structure. *Animal Behaviour*, 60, 463-470.
- McGregor, P. (Eds). 1992. *Playback and Studies of Animal Communication*. New York: Plenum Press.
- Marler, P. 1977. The structure of animal communication sounds. In: *Recognition of Complex Acoustic Signals* (Ed. by T. H. Bullock), pp. 17-35. Berlin: Springer.
- Mitani, J. C. & Marler, P. 1989. A phonological analysis of male gibbon singing behavior. *Behaviour*, 109, 20e45.
- Moore, J. M., Székely, T., Büki, J. & DeVoogd, T. J. 2011. Motor pathway convergence predicts syllable repertoire size in oscine birds. *Proceedings of the National Academy of Sciences, U.S.A.*, 108, 16440-16445.
- Nottebohm, F. 2005. The neural basis of birdsong. *PLoS Biology*, 3, e164.

- Nowicki, S., Searcy, W. A., Hughes, M. & Podos, J. 2001. The evolution of bird song: male and female response to song innovation in swamp sparrows. *Animal Behaviour*, 62, 1189-1195.
- Ouattara, K., Lemasson, A. & Zuberbühler, K. 2009a. Campbell's monkeys use affixation to alter call meaning. *PLoS One*, 4, e7808.
- Ouattara, K., Lemasson, A. & Zuberbühler, K. 2009b. Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National Academy of Sciences, U.S.A.*, 106, 22026-22031.
- R Development Core Team 2012. R Foundation for Statistical Computing, Vienna, Austria.
- Sueur, J., Aubin, T. & Simonis, C. 2008. Equipment review: Seewave, a free modular tool for sound analysis and synthesis. *Bioacoustics*, 18, 213-226.
- Temeles, E. J. 1994. The role of neighbours in territorial systems: when are they 'dear enemies'? *Animal Behaviour*, 47, 339-350.
- Yip, M. J. 2006. The search for phonology in other species. *Trends in Cognitive Sciences*, 10, 442-446.
- Zuberbühler, K. 2002. A syntactic rule in forest monkey communication. *Animal Behaviour*, 63, 293-299.