Birds and babies : a comparison of the early development in vocal learners

Haar, S.M. ter

Citation

Haar, S. M. ter. (2013, June 5). *Birds and babies : a comparison of the early development in vocal learners*. Retrieved from https://hdl.handle.net/1887/20932

Author: Haar, Sita Minke ter
Title: Birds and babies : a comparison of the early development in vocal learners
Issue Date: 2013-06-05

4 Zebra finch song phonology and phonological syntax across populations and continents

Sita. M ter Haar[1,2], Caroline A. A. van Heijningen[1,2], Rob F. Lachlan[3] and Clara C. Levelt[2,3], Carel ten Cate[1,2]

[1]Behavioural Biology, Institute of Biology Leiden, Leiden University
Sylvius Laboratory, Sylviusweg 72, P.O. Box 9505, 2300 RA Leiden, The Netherlands

[2]Leiden Institute for Brain and Cognition, Postzone C2-S, P.O. Box 9600, 2300 RC, The Netherlands

[3]Leiden University Centre for Linguistics (LUCL), Leiden University, P.O. box 9515, 2300 RA, Leiden, The Netherlands

[4]Department of Biology, Duke University, P.O. Box 90338 Durham, NC, 27708-0338, USA

CAA van Heijningen and SM ter Haar share first authorship.

Abstract

The zebra finch (*Taeniopygia guttata*) has been a widely used as model-species. Many different neuro- and behavioural biology laboratories maintain a local population of these song birds. In this paper we present an extensive study of the phonology and phonological syntax (restrictions in sequences of elements) from 13 different populations from Europe, North-America and Australia. Our aim was to explore whether, similar to language, in addition to the well-established individual variation in songs there are more or less species-wide 'universals' as well as systematic variation among populations. By using a computational approach, we found 11 element types occurring among all 13 populations. The only commonality in terms of sequences of elements seems to be the position of long flat elements towards the end of the song. Despite these commonalities there is considerable variation between populations in the distribution of element type frequencies. For syntax, 8 out of 13 populations had a significant amount of structural constraints in their songs. However, these constraints were absent on a species-wide and continent level. These results shed light on the apparent variation in song characteristics between populations of a widely used model species for language evolution and vocal learning. These findings may be important for interpretation of results from different populations and other (bird) species used in experimental research. Moreover, they are suggestive of cultural differentiation being at least as important as genetic diversification processes in causing large scale changes in song characteristics.

Introduction

Birdsong and language exhibit many parallels. First of all both communication systems rely on learning. Human infants as well as young songbirds must be exposed to adult vocalizations in order to properly learn the sounds. Second, predispositions for species specific aspects of the sound seem to play a role in the learning process. Third, both song and language learning go through a babbling stage in which auditory feedback shapes vocal development. And as a last parallel, songbirds and humans posses specialized brain mechanisms for vocal production and learning. Together, this evidence suggests that there may be similar mechanisms underlying the learning of birdsong and language (Bolhuis et al, 2010; Doupe & Kuhl, 1999).

Another similarity, and relevant to this paper, is that both songs and language consist of rapid sequences of differently sounding vocal elements, organized according to specific patterns. In this paper we present an extensive study of the phonology and phonological syntax (restrictions in sequences of elements, independent of differences in meaning) of a songbird species, the zebra finch. Our aim is to explore whether, similar to language, underneath the well-established individual variation in songs there are more or less species-wide 'universals' as well as systematic variation among populations.

Studying both universal and non-universal song features may tell us more about the relative roles of experience dependent and independent factors on vocal development and on the way sound patterns are organized. Birds of different populations can have different 'dialects' if birds from the same population have a more similar song (or element repertoire) than birds from neighbouring populations often related to geographic variation. Differences between populations have been found in for instance, white-crowned sparrows (Marler & Tamura, 1964; Petrinovich & Baptista, 1984) and black-capped chickadees (*Poecile atricapillus* (Kroodsma et al, 1999)) showing distinct sound patterns between neighbouring populations, related to the geographical distance and social interaction between individuals. Over generations such populations may diverge even further, changing their vocalizations because of reduced exchange between populations combined with cultural transmission with slight modifications within populations. Cultural transmission processes may, however, also help to conserve sound patterns, and this, combined with biological factors such as perceptual biases and production constraints, may limit the types of patterns that are possible, leading to some song features that may occur in many (but not necessarily all) populations ('universals') whereas others occur in a few.

Zebra finches are an excellent model system for studying how song patterns may start to differentiate between populations, for which parameters, and which mechanisms cause these changes. Since the 1960's, domesticated zebra finches (*Taeniopygia guttata*) have been an increasingly popular model species for neurobiological and behavioral research. Instead of repeatedly acquiring wild individuals directly from Australia to maintain a stock, research laboratories and commercial breeders around the world have successfully bred many generations of domesticated zebra finches. This process thus creates an unprecedented large scale experiment that allows us to examine both the universal song features present across all populations as well as how strongly song features have diverged among populations. Here, we take a first step towards addressing this issue by providing a detailed description of the variation among 13 populations.

The process of domestication had consequences for the genetic and cultural evolution of these (laboratory) 'populations'. The genetic relatedness between 18 of these laboratory populations from North America and Europe and two wild populations from Australia was analyzed by Forstmeier et al. (2007). They found that populations differed significantly according to broad geographic relationships: the most pronounced genetic differences occurred between the three continents. In addition, domesticated birds were found to differ phenotypically from wild zebra finches, being physically larger than wild zebra finches. In conjunction with this genetic differentiation over the years, cultural evolution may have led to differentiation in learned song characteristics. Repeated (and at least partly isolated) breeding at research laboratories might have lead to the formation of dialects comparable to dialects in isolated populations in other bird species. Comparing the genetic analyses with the variation in songs among different populations and continents can provide insight in whether songs in genetically more similar populations are also more similar in song features or not.

Furthermore, it is yet unknown whether the song of domesticated zebra finches still reflects the natural variation of wild populations or whether domestication has lead to differences in song characteristics compared to those of wild birds. For instance, it is known for Bengalese finches (*Lonchura striata* var. *domestica*) that the process of domestication of white-backed Munia's (*Lonchura striata)* originating from China into Bengalese finches by Japanese breeders of (starting approximately 260 years ago) resulted in more variable intra-individual element transitions and hence much more variable songs (Honda & Okanoya, 1999).

In the current study we examine the variation and universal patterns of phonology and phonological syntax in zebra finch song from wild and laboratory populations and whether the differentiation among populations has parallels with the genetic differences found between continents.

Phonology

Many studies on song variability have focused on 'phonology'. By phonology we mean the study of the smallest units of which a song consists: 'elements' or 'notes', which are usually categorized by visual inspection of spectrograms (fig. 4.1 See Method section for an overview of the different categorization schemes in the zebra finch song literature). In our study, we follow the usual convention of studies on song bird phonology (Lachlan et al, 2010; Prather et al, 2009; Sturdy et al, 1999a; Sturdy et al, 1999b). See also Ten Cate , Lachlan & Zuidema, in press.) by focusing on the signal itself instead of the phonological/categorical perception that has been demonstrated for some species (Nelson & Marler, 1989). Phonological variation between local populations in the type of elements has been shown in several species, for instance in white crowned sparrows (Nelson, 1998). For the zebra finch, some comparative element type studies have been done (Lachlan et al, 2010; Slater & Clayton, 1991; Sturdy et al, 1999b; Zann, 1993a; Zann, 1993b; Zann, 1996). These studies were limited in scope, however. In the present study we examine the presence and distribution of element types on a much wider scale and rather than relying on visual inspection and categorization, we use an advanced computerized element categorization algorithm.

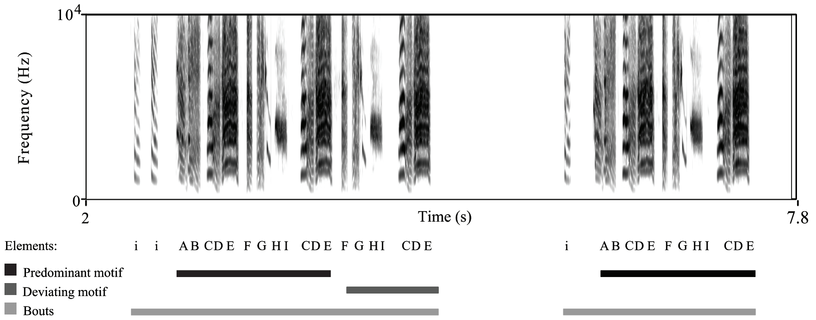

Fig. 4.1. Example of a zebra finch song. A song consists of bouts (light grey) of which the predominant motif (black) is selected for further analyses. The predominant motif is the version occurring most often in an individual's repertoire, compared to a less often occurring one like the deviating motif presented here (medium grey). Introductory elements (i) were not taken into account in this study.

Phonological syntax

Syntax in the context of bird research can be defined in general as restrictions in the sequencing or position of structural units (e.g Lachlan et al., 2010). The degree of variability in the sequence of song elements differs greatly between bird species. For example, Chipping sparrows (*Spizella passerina*) have a single song, consisting of one repeated element (Albrecht & Oring, 1995) while on the other side of the complexity spectrum, the nightingale is a wide known for it's relatively high individual sequence variability (Todt & Geberzahn, 2003; Todt & Hultsch, 1998). Various methods have been employed to describe and analyze the patterns in variability, such as transition diagrams (Honda & Okanoya, 1999) and (higher order) Markov Models (Hailman & Ficken, 1986; Lemon & Chatfiel.C, 1973); see also ten Cate and Okanoya (2012) for an overview). But most such methods have been concerned with investigating variability within individual repertoires, rather than variation between different populations.

In zebra finches, individual males were previously thought to sing in bouts of largely invariant motifs (Glaze & Troyer, 2006; Immelmann, 1969; Zann, 1996). But the current general opinion is that although males sing relatively stereotyped songs, they can vary on this pattern by adding, deleting or sometimes altering elements in the motif (Helekar et al, 2003; Helekar et al, 2000; Scharff & Nottebohm, 1991; Sossinka & Bohner, 1980; Sturdy et al, 1999b; Volman & Khanna, 1995). Additionally, Helekar et al., (2003; 2000) found relatively many repeated elements in songs from their lab population, a phenomenon that has been observed in isolate song (raised without a tutor) as well (Helekar et al, 2000; Jones & Slater, 1996).

On the inter-individual (or population) level, there is far more variability, more specifically in where element types can occur in the song. Lachlan et al., (2010) found high variability in the syntactic patterns as all element types can be followed by each other. This is in contrast to an earlier study by Zann (1993 a, b & 1996), who found patterns for element types relative to the position in the song (start, middle or end). Zann suggested that song elements are not equally distributed over the song: the first section consists mostly of introductory elements, the middle section consists of a 'high' element and 'stacks' and 'distance calls' occur more towards the end of the song (Zann, 1996). Thus, previous studies have shown that both variation and common patterns can be found in zebra finch song.

In the present study we provide an extensive species-wide analysis including both phonology and phonological syntax on data from 12 laboratory populations and one wild population across three continents (populations grouped according to geographical proximity. Therefore, we grouped the Auckland population from New Zealand with the wild population sample from Australia). Computational bioacoustic analyses are performed on zebra finch song to answer the following questions for both phonology and syntax:
1) What are the species-wide element types and are they present in all populations (i.e. are they universal) or do populations differ in this?
2) What is the degree of species-wide phonological syntax and how consistent is this across populations (i.e. are they universal) or do populations differ in this?
3) Are there differences between continents and how do these relate to population differences?

Methods

Song analysis by computer software: Luscinia

Zebra finch song is known for it's relatively difficult segmentation and annotation in element types and/or syllables. Except for Lachlan et al., (2010) all previously mentioned articles have been based on visual inspection for assessing element categorization. Even though the use of several observers can reduce the variability of visual classification (Jones et al, 2001), such studies are limited in the sample size of songs they can analyze, and the degree to which quantitative measures of inter-element similarity can be provided (as opposed to same/different binary scoring). It is therefore difficult to statistically address questions of phonological and syntactical organization without computational methods.

Using the software package Luscinia (http://luscinia.sourceforge.net) for acoustic and statistical analysis Lachlan et al., (2010) compared zebra finch song from four laboratory populations as part of a larger comparison between several songbird species. Because we had relatively few suitable songs from Australian regions (Alice Springs and Northern Victoria) we decided to merge these to one population. Moreover, because of the high level of dispersion in wild zebra finches, there are no clear dialects found in wild zebra finch song (Zann, 1996), nor large genetic differences between wild populations (Forstmeier et al, 2007). See table 4.1 for a list of all populations in this analysis.

Sound Analysis in Luscinia

Recently, several software packages (e.g. Sound Analysis Pro, Tschernikovski et al., 2000, Luscinia, Lachlan et al., 2010) have been developed that allow the acoustic analysis of large datasets of birdsongs. By using computational algorithms to compare songs, quantitative comparisons between large numbers of elements can be made, with little or no observer bias. In Luscinia, measures of several acoustic parameters are taken at each spectrogram time slice throughout each element. Luscinia then uses an implementation of Dynamic Time Warping (DTW) to compare elements based on these time-series vectors. (See the SI for the DTW parameters chosen for this study). The output of this comparison is a matrix of the dissimilarities between each pair of elements in the dataset. This matrix can be ordinated into Euclidean space using nonmetric multidimensional scaling. The resulting principal components can then serve as the basis for cluster analyses in statistical software packages like R (and open-source package for statistical analysis available from www.r-project.org). With this species-wide clustering solution it is possible to identify element clusters specific to populations or universal to all populations.

Motif selection

As shown in fig 2.1, zebra finch songs consist of motifs, preceded by a number of introductory elements. Repeated motifs are called a song bout. Motifs can vary slightly within birds. We selected the predominant motif of each bird for analysis (i.e. the motif with the highest probability for that bird (Helekar et al, 2000; Sturdy et al, 1999a). Introduction notes were excluded unless they recurred in motifs later in the bout. Motifs were selected by two observers (StH and CvH). Each observer selected half of the motifs for each population reducing the possible effect of observer bias. For a subset of the songs both observers selected the motif in order to test reliability. Since this reliability between the observers was relatively low (0.60 (see SI for more details)), the songs that were ambiguous were re-examined by both observers and the motif for which consensus was reached was selected.

Element segmentation

Throughout the years of zebra finch song research, multiple methods of element segmentation have been published (Lachlan et al, 2010; Scharff & Nottebohm, 1991; Sturdy et al, 1999a; Sturdy et al, 1999b; Zann, 1993a; Zann, 1993b). These are not all in agreement. In the present study we segmented motives into elements based on the criteria of silent gaps and/or abrupt changes in frequency and amplitude. The segmentation was performed manually by StH and CvH, using Luscinia, after several years of experience with zebra finch song analysis and multiple training rounds on other songs. Each observer performed element segmentation on half of the motifs from each population reducing the effect of observer bias. For an estimate of reliability a subset of the data was segmented by both observers (see SI for more details). Average agreement between observers was 86 %.

Cluster analysis

A *cluster analysis* was performed with Mclust (package 'mclust' was built under R version 2.12.2) (Fraley & Raftery, 2002) within R (version 2.13.0, available at http://www.r-project.org) on the output from Luscinia (the first ten PCO values which explained more than 92% of the variation in the data set). Mclust clusters data using Gaussian Mixture Modeling. This approach allows for clusters of varying size and shape to be discovered (Fraley & Raftery, 1998). Mclust then uses the Bayesian Information Criterion approach to determine which model best fits the data, allowing a statistical assessment of the number of clusters present.

Having placed elements into clusters, further statistical phonological analyses were performed on the zebra finch songs annotated according to the element types calculated by Mclust.

Relative position

Some element types may be restricted to certain positions in the song. For instance the songs of chaffinches (*Fringilla coelebs*) show relatively strict patterns (Slater & Ince, 1979). The relative position of elements in a motive was estimated by ranking each element according to its position in the song (low numbers for elements early in the song and higher numbers for those at the end). Because songs differ in total number of elements, a relative measure of position was calculated by dividing the position number by the total number of elements per song. For example: the 8th element in a 10-element song has a relative position of '0.8'. A relative position was determined for each element in each song. The mean relative position was calculated for each cluster of elements (as classified by the Mclust algorithm in R). This thus reveals if there is a higher chance for certain types of elements to in a specific region of the song. This does not mean that the elements are restricted to that area and cannot occur in other areas, but rather that a certain area (the last part of the song for instance) contains more elements of a certain type (for instance cluster 1) than another area.

Species-wide syntax analysis

The amount of structure, or syntax, was measured with a Sliding Window Match Length (SWML) entropy estimator that analyses the sequences of elements in the zebra finch songs (Suzuki et al, 2006). 'Entropy', in this context, is a measure developed in the field of information theory to describe the unpredictability in a signal. This theory has been applied before to find patterns in animal vocalizations, for instance by Lemon and Chatfield (1973) in rose-breasted grosbeaks (*Pheucticus ludovicianus*), by Suzuki et al., (2006) in humpback whales (*Megaptera noveangliae*) and by Briefer at al (2010) in skylarks (*Alauda arvensis*).

The higher the information entropy, the lower the degree of structure or constraints in the data is. For instance, a random sequence of different element transitions results in a very high entropy value. In contrast, messages with lexicographical, grammatical and contextual constraints result in a lower entropy value as not all combinations of elements are present in the message (Suzuki et al, 2006).

One limitation with entropy-based measures of syntax is that they require constituent elements to be placed into discrete categories. Different categorizations of elements may result in very different entropy estimates. A solution to this problem was proposed by Lachlan et al., (in prep). This method searches for the partitioning of elements that minimizes entropy (that is maximizes syntax). There are two steps in this method. In the first, elements are assigned to clusters based on their acoustic similarity (as determined from the DTW comparison, above) to a predetermined number of *k* cluster centroids. In the second step, entropy is estimated using the first order Markov estimator for the clustered data. The algorithm proceeds by searching for centroids that minimize the entropy estimator in a similar way to the k-medoids algorithm (from which this algorithm is adapted).

Once the partitioning with most structure had been identified, we recalculated entropy for that partitioning using the SWML estimator. The benefits of this type of entropy estimator are that it is applicable to a relatively broad range of data, as it has less a-prior data assumptions than other syntax inducing algorithms. Also, it is relatively robust to departures from these assumptions. And lastly, even if the sample sequence is not long enough to achieve asymptotic convergence, the resulting entropy estimates are an upper bound on the source entropy (Suzuki et al, 2006). Therefore, even with relatively small sample sizes this method can estimate the amount of structure present in zebra finch song reliably. It is, however, too slow to use during the clustering part of the algorithm (above). We set the window length for this algorithm to be the length of the individual motifs. We then rescaled entropy as 'redundancy', calculated as $(H_{baseline}-H_{swml})/H_{baseline}$. We used $H_{baseline}$, the entropy calculated from randomly sequenced sequences instead of the maximum entropy (H_{max}) that Suzuki et al. (2006) used as $H_{baseline}$ turned out to be a more stable parameter. Our redundancy estimates then served as our measures of syntactic structure. Redundancy is a way of scaling entropy were '0' means that element transitions are random (like randomly shuffling the sequences of elements in each song) and '1' means that sequences are entirely predictable and non-variable. We estimated confidence limits around our estimates using a bootstrapping technique (Suzuki et al, 2006). We estimated redundancy for values of *k* between 2 and 10 and selected the highest redundancy value as our estimate of syntactical structure for the particular level of analysis (species-wide, per laboratory or per continent).

Population level syntax analysis

The individual redundancy scores per population were calculated and bootstrapped to arrive at confidence intervals around these scores to indicate presence or absence of structural constraints, or redundancy (ρ) per population.

The redundancy analysis returns one value per population, so in order to compare the populations with each other for in their degree of syntactical structure, a deviation scores for each individual in each population was calculated and analyzed in an ANOVA. The deviation (calculated from the SWML distances, converted to PCO values) is a measure for how different song syntax is between one song sequence and it's population average (centroid), an adaptation of the method described for variation tests by Anderson (2006) and Anderson et al., (2006). If songs are on average further away from the centroid, they also on average vary more from each other, and hence have a higher entropy (and are therefore less structured or constrained in terms of syntax).

Continent level syntax analysis

To calculate the amount of syntax in each continent, pairwise ρ scores for within-continent populations were calculated instead of ρ scores per continent. The latter scores are less suitable in this context because the Australian continent contained data of only two populations (one wild population from Australia and 1 laboratory population from Auckland, New Zealand). The other populations contain 6 (North America) and 5 populations (Europe) and the ρ is influenced by the number of populations in the comparison. We therefore calculated pairwise ρ scores between populations within continents to estimate the amount of syntax within continents and bootstrapped the results to calculate confidence intervals. We then averaged over the pairs within continents to arrive at an average and confidence interval per continent.

Results

Phonology

To cluster all elements according to acoustic parameters we measured them in Luscinia

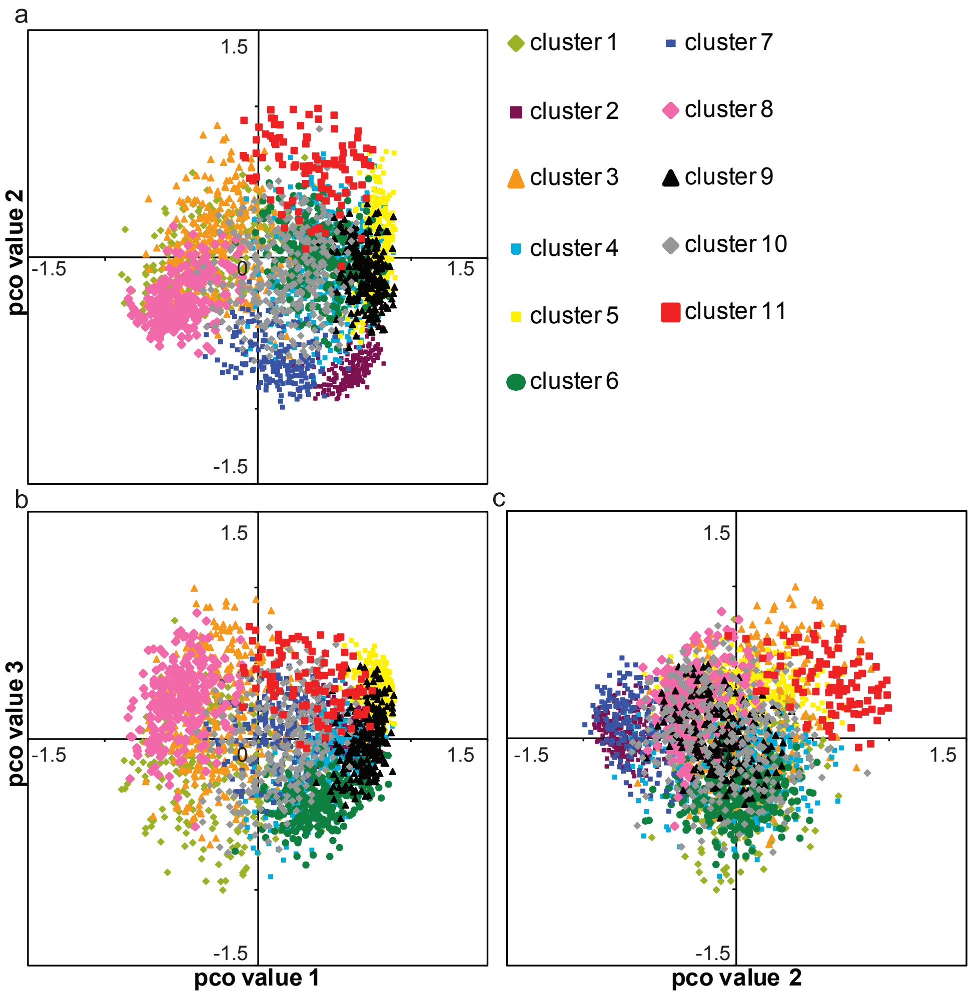

Fig. 4.2. Clustering graph 11 clusters all populations together. A) PCO value 1 vs 2. B) PCO value 1 vs 3. C) PCO value 2 vs 3. Each data point is an element. Different colors (and shapes) represent the different element types (clusters as modeled by Mclust).

and performed a DTW analysis which was followed by an NMDS and subsequent PCO analysis based on 10 components. These 10 components explained 92% of the variation (Kruskal stress value of 0.051, indicating a good fit (Ozturk et al, 2009).

The Mclust algorithm calculated an optimal model of 11 clusters (see fig. 4.2a, b and c) in a VEV configuration (Ellipsoidal, variable volumed, equal shaped and variably oriented) and a BIC value of 8821.2. The next best fitting models had BIC values of 8820.8 with 14 clusters and 8777.1 with 13 clusters respectively. We chose the 11-clus-

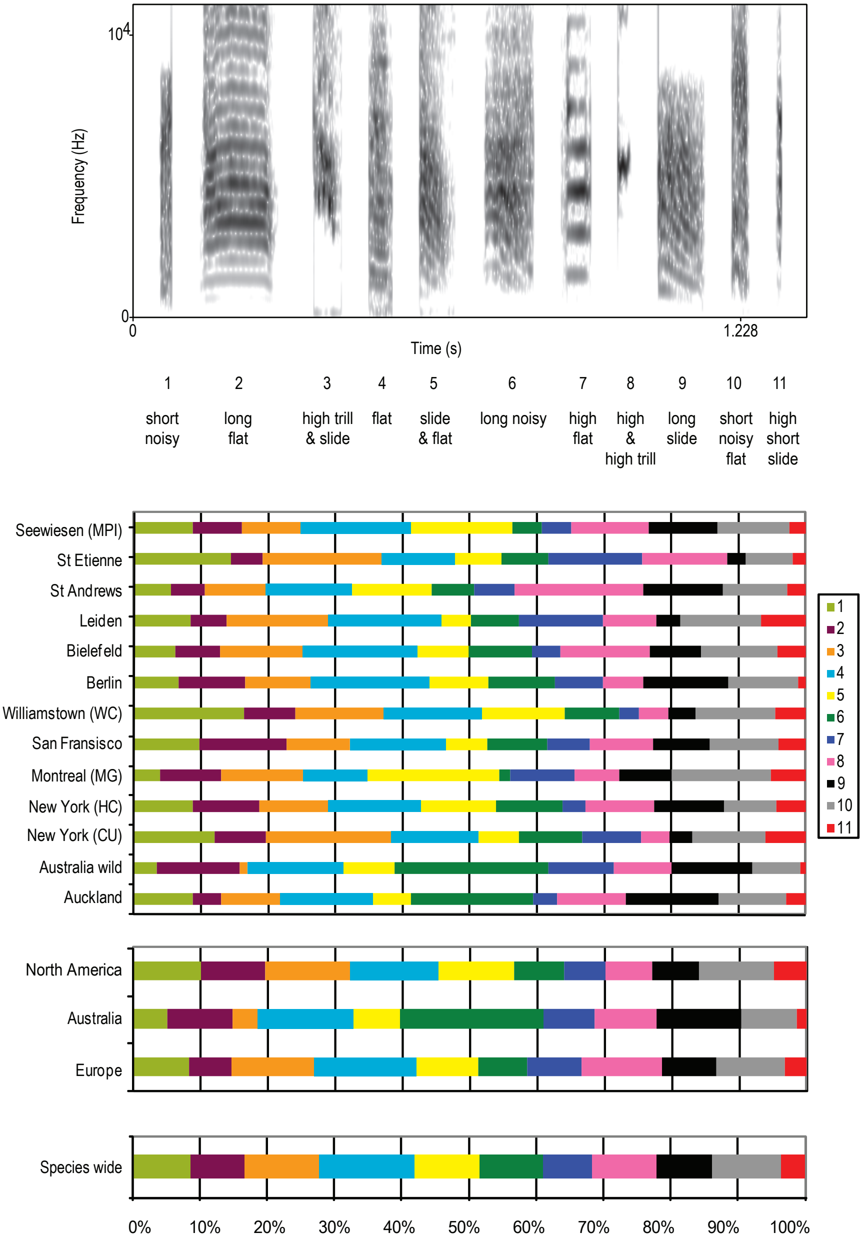

Fig. 4.3. Element types and their distributions per population, continent and species-wide level. Top panel provides representative examples of the element types found in the Mclust clustering analysis. Numbers indicate individual element types (clusters). Lower panels show element type distributions in percentages per population, continent and species-wide.

ter model as this one had the highest BIC and it is the most parsimonious model where the number of clusters is concerned. To estimate the overall tendency for the data to cluster, we compared the one cluster solution BIC value (-2283.4) with the value Mclust gives for the optimal clustering model. The large difference between the two indicates a very high propensity of the data to fall into clusters.

To give an example for each of our clusters, we selected the first three elements from the densest area of each cluster to characterize the element type. If the first three elements would not give a conclusive characterization, we analyzed four elements (in the case of cluster 3 and 8). For labeling of the element clusters, see fig. 4.3.

The distribution of the 11 clusters per laboratory, per continent and species-wide is shown in fig. 4.3. All clusters occur in every population, indicating that all element types occur in all populations and therefore that there are no population-specific element types. Yet, we found that the elements from the different clusters are not distributed evenly over the populations using a X^2 permutation test in which individuals were permuted between populations (with 10000 permutations, $X^2 = 441.8$ $p < 0.001$, Cramér's effect size 0.13). We also found that continents differed in element type distribution by X^2 permutation test in which individuals were permuted between continents (with 10000 permutations, $X^2=162.9$, $p <0.01$, Cramér's effect size 0.17).

Number of elements per song

There was variation in the number of elements per song between the different populations, see fig. 4.4. (F=2.344, deletion p = 0.008, based on 188 df, with a glm model with quasi-poisson link). However, when aggregated over continents, there were no differences in element number per song (F=1.736, deletion p > 0.05, based on 188 df, with a glm model with quasi-poisson link).

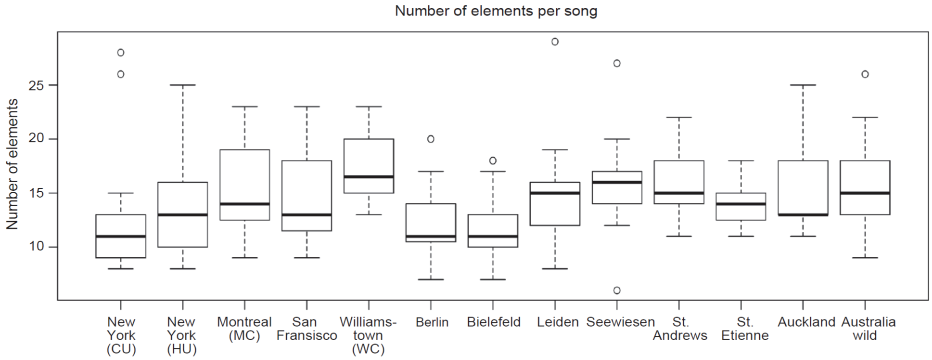

Fig. 4.4. Boxplot distribution of the number of elements per song, per population.

Relative position of elements within a song

The mean relative position of each of the 11 element types is presented in fig. 4.5. The overall Kruskal-Wallis test was significant (K-W X^2=166.84, df=10, p <0.001 and therefore the element types are not occurring evenly over the positions in the song. Further testing showed that on a species-wide level clusters 2 (two-sample Wilcoxon test, W=158762, p <0.001, and 7 (two-sample Wilcoxon test W=198835, p <0.001) occurred in the end region of the song (see fig. 4.5). (All tests were still significant after Bonferroni correction for multiple testing). Roughly the same distribution pattern can be observed on continent level: For all continents, the overall K-W tests were significant: X^2= 32.4, df = 10, p < 0.001 (Aus), K-W test X^2= 125.3, df = 10, p < 0.001 (N-Am); K-W test X^2= 52.8, df = 10, p < 0.001. After post-hoc tests, we found similar patterns compared to the species-wide pattern, except for Australia, where cluster 9 was also occurred relatively to the end (K-W X^2= 6847, p = 0.006, still significant after Bonferroni correction) and cluster 7 was not.

Phonological syntax

Species-wide level syntax

After analyzing the songs with the k-medoids/Markov chain modeling and SWML estimator algorithm in Luscinia, we found a mean redundancy (ρ, amount of structural constraints) that did not significantly differ from 0 (-0.016 with maximum redundancy at k= 8 syntactical clusters CI 95%: -0.065 to 0.030). This implies that there is no evidence for syntactic structure, in terms of sequences of specific element types, on a species-wide level in zebra finch song.

Per population syntax

When the data from each individual population were analyzed separately for the amount of structure (redundancy), we found that the bootstrapped estimate of ρ, averaged over all populations was 0.109. This is higher than when all populations were analyzed together on the species-wide level. For eight of the populations, there was significant evidence of syntactic constraints (see table 4.1 with pop mean ρ's, k=nr of clusters and bootstrap p value). Estimates of ρ were not influenced by sample size in any detectable way when comparing between populations, indicating that these sample sizes were adequate (see fig. SI 4.1).

We compared the syntactical structuring between the different populations by calculating the deviations, but there were no significant differences found (ANOVA, df 188, p-value 0.19), so there were no significant differences between the lab populations in the amount and variability of syntax in the songs of that population.

Table 4.1. List of populations with amount of syntax scores (rho).

lab population	continent*	sample size	rho	highest k	bts p-value	sign
ALL	-	189	-0,016	8	p> 0.5	
Auckland	Australia	9	0,108	4	p<0.025	*
Bielefeld	Europe	18	0,082	4	p>0.025	
Berlin	Europe	15	0,169	6	p<0.0025	*
New York (CU)	N. America	14	0,018	3	p>0.025	
New York (HC)	N. America	15	0,146	10	p<0.025	*
Leiden	Europe	15	0,050	6	p>0.025	
Australia wild	Australia	17	0,128	4	p<0.025	*
Montreal (MG)	N. America	15	0,121	10	p<0.025	*
St Andrews	Europe	14	0,102	9	p>0.025	
St Etienne	Europe	15	0,138	2	p<0.025	*
San Fransisco	N. America	15	0,178	2	p<0.0025	*
Seewiesen (MPI)	Europe	13	0,061	7	p>0.025	
Williamstown (WC)	N. America	14	0,096	7	p<0.0025	*

* 'continent' refers to grouped in terms of geographical distances. CU=Columbia University,

HC=Hunter College, MG=McGill University, MPI=Max Plank Institute for Ornithology, WC=Williams College. Bts: bootstrap

Per continent syntax

Estimating continent-wide syntactical structure was complicated by the relatively low number of Australian populations (two: 1 wild, 1 lab from Auckland) compared to the others (five and six). Based on the population level analysis (above), this would be expected to artificially boost the estimate from Australia compared to the other continents. To avoid this effect we therefore made pairwise comparisons between all possible pairs populations within each continent and calculated deviation scores. The mean deviation values averaged over the pairs of populations in these continents were: 0.047 (North America, with CI 95%: -0.053 to 0.140) 0.03 (Australia with CI 95%;-0.012 to 0.175) and 0.036 (Europe with CI 95%: -0.052 to 0.122). However, none of these scores were significantly different from 0, indicating that there is no evidence for structure on a continent level. Therefore there is also no variation between continents.

General discussion

Our computational analysis of zebra finch song from 13 populations indicates that in terms of phonology, there is evidence for clustering of elements into 11 types on a species-wide scale. All element types occur in all populations but the distribution of element types differs significantly between populations as well as continents. In terms of syntax, or structural constraints found in the sequences of elements in zebra finch song, there is variation in the amount of syntax on population level, but these effects are absent when looking at continent and species-wide scale (all results from this study are summarized in table 4.2).

Table 4.2. Results overview

	phonology	**phonological syntax**		**relative position**	**tot nr of elements**
level of analysis	**element types**	**amount of syntax (ρ)**	**deviation**		
species-wide	11 clusters	0	-	2 and 7 in end region	-
continent	variation in distribution of these 11 clusters	0	-	2 (Aus,Eur, N-Am), 7 (Eur, N-Am), 9 (Aus) in end region	no variation
population	variation in distribution of these 11 clusters	8/13>0	not sign	-	variation

Phonology

The 11 element categories as obtained by the clustering algorithm analysis of all 13 zebra finch populations together is different from earlier findings using a comparable analysis, but data from three populations only. This resulted in just two categories; high notes and all low frequency elements together (Lachlan et al, 2010). The difference may be explained by the number of populations used. The difference in clustering method however, may also account for some of the differences, since the Mclust method used here also takes into account different cluster shapes and sizes, whereas the simple agglomerative clustering method used in the previous study assumed only perfect spherical clusters. Other studies have described different note types as well, using visual inspection of spectrograms of laboratory birds' songs (Price, 1979; Sturdy et al, 1999b; Williams & Staples, 1992), or from wild individuals only (Zann, 1993a; Zann, 1993b). Zann reported 14 element types, Sturdy 5 types and Price also 5 types. There is some overlap in categories described previously and the ones found in the present study. For instance high notes and slides are describes as separate categories by Zann (1993a and b) and Sturdy et al. (1999b). This is also valid for the present study, however multiple categories of slides and stacks are found. Sturdy et al.'s study including four laboratory populations also found differences in distribution of element types between populations but every category was found in all populations (Sturdy et al, 1999b). This seemingly corroborates our present findings. It should be noted however, that the classification by Sturdy at al. was aimed at providing generalizable information about zebra finch song. Thus Sturdy et al. probably deliberately classified elements (or note types) into more universally applicable categories. Differences in classification methods between studies make a more direct comparison with the present results difficult. For instance some of the previous authors used silence as a segmentation criterion, and therefore their classification includes larger units such as combination elements. In our study these combination elements are segmented into multiple elements as we also included abrupt changes in frequency as a segmentation criterion.

This brings us to an important point of consideration for interpreting the current findings: the analysis is directed at the level of elements instead of syllables, since we also use abrupt frequency change as a criterion for segmentation. At this point it is not entirely clear what level of representation is that birds use themselves while processing songs. Ten Cate and Slater (1991) as well as Williams and Staples (1992) reported that songs are copied mostly as chunks separated by silent intervals, which may include several elements and/or syllables. The same studies, however also demonstrated that elements and syllables within chunks could also be copied separately, indicating a more detailed representation. Therefore, in the present study we chose an analysis on the element level.

None of the 11 element clusters found in the present study is specific to any of the populations (see fig. 2.3) and specifically each cluster appeared in each population. The species-wide distribution of element clusters does show significant variation between populations, however. For instance, cluster 11 seems more consistent in presence among populations compared to cluster 3, which appears to show much more variation in abundance. Although perception and production might go in parallel, it is yet unknown whether the zebra finches would perceive the same 11 clusters we found as separate perceptual categories. Since distribution of categories is different between populations in the present study, it is also possible that category boundaries are different across populations (Prather et al, 2009). This might be the case for zebra finches as well and therefore the clustering itself might differ between populations. Using clustering methods similar to those in the present study but rather on population level, could give insight in where these boundaries may lay (however for the present study, this was not possible due to sample size constraints). But ultimately, perceptual experiments with zebra finches are required to support the conclusions of computational and statistical studies like this one.

Variation in distribution of element clusters is also found when populations are grouped per continent, implying that there is some continent specific phonology (see fig. 2.3). The effect sizes from both the population and continent analysis are rather similar (0.13 and 0.17 respectively). The study by Forstmeier et al. demonstrates genetic differentiation between zebra finches in different continents, but nevertheless considerable variation within continents exists (Forstmeier et al, 2007). When we compare our phonology results to the genetic relatedness in the unrooted dendrogram from Forstmeier et al, (2007), we found a slightly different pattern, namely that the variation between continents was more or less equal to the variation between populations. In the Forstmeier study (2007) instead, continent variation was larger than population variation. Nonetheless, the roughly equal variation in song elements between populations and continents suggests that cultural evolution might have an equally important role in the diversity of birdsong. To analyze this in more detail, one would also need information on the cross-laboratory exchange history of birds. On the other hand, the fact that similar element types exist among populations that have been separated for several generations suggests that the cultural variation might to a certain extent be constrained by perceptual or learning biases. That certain constraints shaping the song into a certain direction are present, has also been indicated by an intriguing experiment by Feher et al., (2009). They used birds that were raised without exposure to song, and therefore developed aberrant isolate song and used these birds as tutors for the next generation. After approximately three generations the song features were already similar to wild-type song. Moreover, a recent study in our lab suggests that biases for acoustic features common to zebra finch song may be present in juvenile birds before they learn song (Chapter 3, this thesis).

Total number of elements per song

There is significant variation in the total number of elements per song between the different lab populations (see fig. 2.4), but this was not the case when aggregated over continents. Slater and Clayton (1991) did not find such a difference when analyzing three populations. However, they analyzed the number of *different* elements per individual song. It would therefore be interesting to re-analyze the total number of element data according to element type in a follow-up study. It is important though to keep in mind the difference between within song element differences (Slater & Clayton, 1991) and analysis based on element *types*: two elements in a song may belong to the same basic type, but still be noticeably different. We also did not find clear evidence for the occurrence of repeats, as was found by Helekar (2000 and 2003). It would be interesting to further analyze the dataset for within song immediate repeats and subsequently whether these can be linked to specific element types.

Relative position of element types in song

First, we examined whether there are associations between element type and relative position in the song. When all lab populations are averaged, the element type from cluster 2 (long flat), and to a lesser extend also cluster 7 (short noisy flat), are positioned towards the end of the song (see fig. 4.5). The same pattern can be observed for North America and Europe. Interestingly, in Australia, also cluster 9 (and not cluster 7) has a tendency to occur in the end region of the song (see fig. SI 4.2). An explanation why this did not lead to a higher ρ in the redundancy syntax analysis, might be that the long flat element do not necessarily occur only exactly as the last element of the song, but on average in the end region (mean position 0.7) and can be flanked by any type of other element. The link between long flat notes and the end region of the song has been observed by other researchers as well (Zann 1993 a, b and 1996).

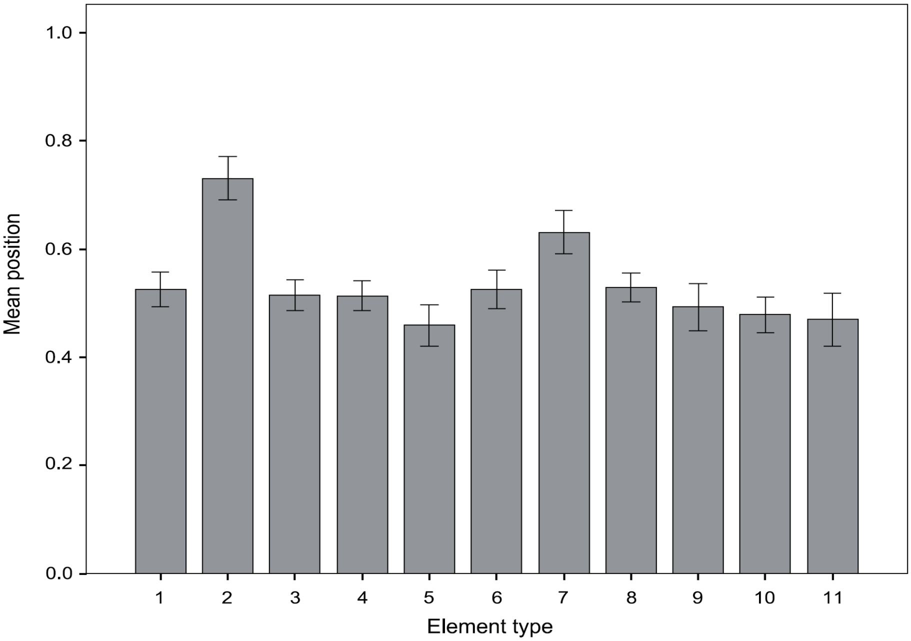

Fig. 4.5. Relative element position on species-wide level. A mean position of 0.5 indicates that the element type does not particularly occur at the beginning or end of a song. A mean position of 0.7 means that that element type occurs relatively often at in towards the end of the song. Error bars represent 95% CI.

Phonological syntax

Species-wide, continent and population level syntax

On a species-wide and continent level, we did not find evidence for species-wide constraints in the sequences of zebra finch song. The estimates of redundancy found for zebra finches in this study were, for example, much lower than a recent study of chaffinches, which found population-level redundancy values of approximately 0.5 (Lachlan et al., in prep). Lachlan et al., (2010) found no evidence for clear structural constraints based on data combined from three lab populations (from different continents), similar to our results.

Nevertheless, on population level we did find evidence for the presence of syntactical constraints in 8 of the 13 populations. So in a number of populations we found restrictions in what kind of song sequences the local zebra finches sing, while in others there is no evidence for such restrictions or syntax.

At first sight this might seem to conflict with the finding that on a species-wide scale, we found no evidence for such constraints in zebra finch song. This is not due to small sample sizes, which potentially might affect the SWML estimator (Suzuki et al., 2006). As shown in fig. SI 4.1, the estimates of ρ were not influenced by sample size in any detectable way when comparing between populations. On the other hand, we found a significant difference in the deviations when aggregated for all populations compared to the individual deviations calculated for each population separately (paired t test, t= 13.35, p <0.001). Therefore, although there was syntactical structure within populations, it varied between populations. Thus when the populations were aggregated and analyzed together, there was no evidence for species-wide structure.

The most likely explanation is that there might be an effect of tutor sharing leading to frequently occurring sequences in particular populations. Zebra finches can incorporate chunks (short sequences of elements) from their tutor(s) in their song (ten Cate and Slater 1991) and can also from learn other males nearby (Williams and Staples 1992). The most obvious way that two individuals could share elements was if they were siblings. Although the aim was to keep the number of brothers in the laboratory datasets as minimal as possible, this was not always possible to avoid, especially when birds are allowed to breed in aviaries and the paternity cannot be established easily. We could not control whether males shared ancestors in earlier generations, such as grandfathers. Nevertheless, such a process would resemble the formation of dialects known from other song birds and human literature (Marler & Tamura, 1964 and Tserdanalis & Wong, 2004 respectively).

What is important to keep in mind here is that is hard to quantify syntactical structure; different methods make different assumptions, and consequently capture different aspects of syntactical structure. It is possible that we missed patterns on a higher level such as syllable- and chunk sequences and nonadjacent dependencies. Furthermore, as a first effort, we chose to include one motif per individual, but it might be interesting to take into account the intra-individual variation in future studies as well.

Domestication

The differences we have observed in the relative frequencies of element types and syntax between different populations might be in part an effect of domestication. Effects of domestication in songbirds have been observed in a closely related species, the Bengalese finch (Honda & Okanoya, 1999) that has had a longer history of domestication (approximately 260 years, compared to approximately 150 years in zebra finches (Rogers, 1979; Sossinka, 1970). It has been shown that song of domesticated Bengalese finches is less constrained and has different acoustic characteristics compared to those of wild strain birds (Honda & Okanoya, 1999). It might be that after a longer period of domestication, zebra finch populations show more phonological diversity than they do now and that the commonalities are partly a result of cultural transmission as such. However, the history of zebra finch domestication, if any, longer than 150 years ago is mostly unknown (Sossinka, 1970) so we cannot be certain if this comparison is realistic.

Domestication effects can be caused by different factors, such as founder/common tutor effects in small populations (more random drift effects) or human/breeder selection choices such as the number and size of generations and whether specific traits were selected for, such as bird size and clutch size. Such effects probably differ between populations. In case of aviary breeding female choice for specific male (song) traits may play a role (see Riebel 2009) for a review on female preference in zebra finches). Also for disentangling the effects of domestication or drift, more should be known about the breeding history of several populations.

Taken together, this may have consequences for how representative data from one population is for zebra finches in general, since different research laboratories may not always be able to replicate each other's results. Domestication may have different effects on different populations, depending on the size of populations; influx of new birds from other populations and breeding schedules. It remains difficult to disentangle the relative effects between cultural transmission, biological evolution and domestication on the evolution of song. Nevertheless, the methods in the current study present a useful set of tools to investigate such issues in more detail.

Comparison with human language

Songbirds are often used as a model for studying vocal learning and language or speech. The current findings may offer some insight in how biological and cultural evolution may relate to difference in sound patterns across populations worldwide. A direct comparison with language variation among populations of humans is difficult. In terms of phonology usually one language or dialect has a different phoneme inventory then another, but all individuals within a population make use of the same inventory. This is different in zebra finch populations first of all because individual differences between birds in element use are abundant (thus only a subset of the elements of 11 clusters, are present in all individual birds, but all element clusters are present in all populations). Second, the present findings suggest that all phoneme categories are present in all populations. However the different distribution of elements may still tell something about cultural evolution, since some categories seem more flexible than others. For instance studies on language change suggest that some phonemes are more prone to change than others, although the cause of such change is a subject of debate (Moreton, 2008). Additionally, language and speech research shows that some aspects of languages are (initially) universal in both perception and production, indicating biological constraints on the possible variation of sounds (Jakobson, 1941). It would be interesting to see in a longitudinal study whether the more flexible element categories found in the present study indeed show more change over time than the less flexible ones.

Conclusion

By using a computational approach, we found 11 element types occurring among all 13 populations. The only commonality in terms of sequences of elements seems to be the position of long flat elements towards the end of the song. Despite these commonalities there is considerable variation between populations in the distribution of element type frequencies. For syntax, 8 out of 13 populations had a significant amount of structural constraints in their songs. However, these constraints were absent on a species-wide and continent level. These results shed light on the apparent variation in song characteristics between populations of a widely used model species for language evolution and vocal learning. These findings may be important for interpretation of results from different populations and other (bird) species used in experimental research. Moreover, they are suggestive of cultural differentiation being at least as important as genetic diversification processes in causing large scale changes in song characteristics.

Acknowledgements
We would like to thank Peter Snelderwaard for help with the software and database installation.

SUPPLEMENTARY INFORMATION

Luscinia DTW Settings list
Settings and parameters we included in the DTW analysis were: weight by amplitude; log transform frequencies; time, fundamental frequency, fundamental frequency change and harmonicity; compression factor 0.25; SD ratio 0.5.

Motif Selection Reliability
For reliability estimate, a subset of the motifs was selected by two observers. Percentage of motifs that were the same for both observers was used as a measure of reliability and is reported in the manuscript

Element Selection Reliability
For reliability a subset of the data was segmented by both observers. Agreement was calculated as the percentages of elements that were segmented similarly by the two observers. Since the total amount of elements was slightly different for each observer, two percentages were calculated; one for each observers' total number of elements per motive (83 % and 89 %). The average of these two numbers was used as average percentage agreement and is reported in the manuscript

Fig. SI 4.1. No relationship between population size and estimated ρ.

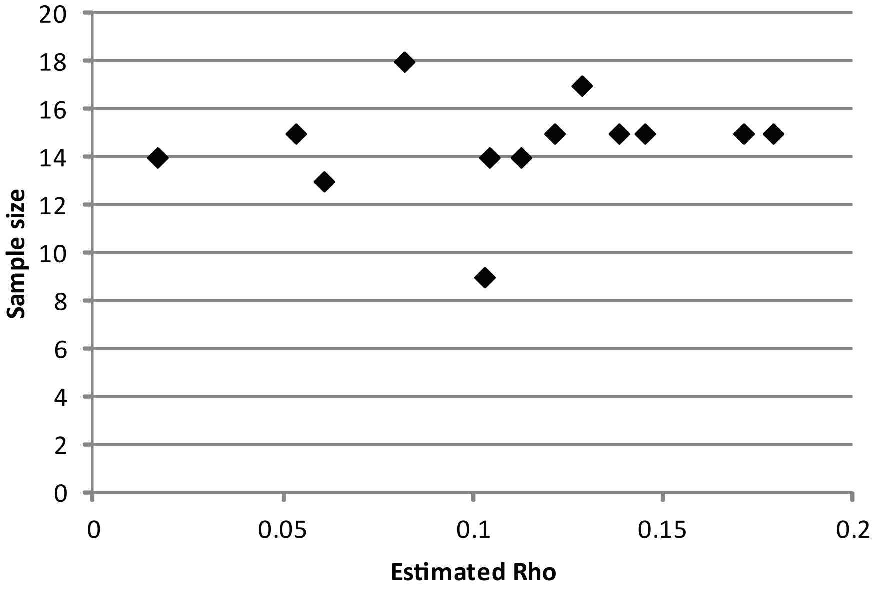

Fig. SI 4.2. Relative element position in the song on continent level. Error bars represent 95% CI. A) North America. B) Australia. C) Europe.