Author for correspondence:
Sita M. ter Haar
e-mail: smhaar@ulg.ac.be

The interplay of within-species perceptual predispositions and experience during song ontogeny in zebra finches (*Taeniopygia guttata*)

Sita M. ter Haar[1,2,3], Wiebke Kaemper[1], Koen Stam[1], Clara C. Levelt[2,3] and Carel ten Cate[1,3]

[1]Institute of Biology Leiden (IBL), Leiden University, PO Box 9505, Leiden 2300 RA, The Netherlands
[2]Leiden University Centre for Linguistics (LUCL), Leiden University, PO Box 9515, Leiden 2300 RA, The Netherlands
[3]Leiden Institute for Brain and Cognition (LIBC), Leiden University, c/o LUMC, Postzone C2-S, PO Box 9600, Leiden 2300 RC, The Netherlands

Vocal acquisition in songbirds and humans shows many similarities, one of which is that both involve a combination of experience and perceptual predispositions. Among languages some speech sounds are shared, while others are not. This could reflect a predisposition in young infants for learning some speech sounds over others, which combines with exposure-based learning. Similarly, in songbirds, some sounds are common across populations, while others are more specific to populations or individuals. We examine whether this is also due to perceptual preferences for certain within-species element types in naive juvenile male birds, and how such preferences interact with exposure to guide subsequent song learning. We show that young zebra finches lacking previous song exposure perceptually prefer songs with more common zebra finch song element types over songs with less common elements. Next, we demonstrate that after subsequent tutoring, birds prefer tutor songs regardless of whether these contain more common or less common elements. In adulthood, birds tutored with more common elements showed a higher song similarity to their tutor song, indicating that the early bias influenced song learning. Our findings help to understand the maintenance of similarities and the presence of differences among birds' songs, their dialects and human languages.

1. Introduction

Vocal learning is essential for spoken language as well as for birdsong, and the learning processes involved show many parallels [1,2]. One interesting parallel is that both processes are guided by perceptual predispositions (i.e. perceptual biases independent of perceptual experience) that interact with experience to guide vocal development. While presence of this interaction is broadly accepted, ongoing debates and discussions concern the nature of the predispositions and of the interaction in shaping vocal production [3,4]. In this paper, we address these questions for a songbird species, the zebra finch (*Taeniopygia guttata*).

The presence of predispositions in vocal learning is suggested by the distribution of sound patterns within and between populations. In human languages, some sounds are more or less universally shared, suggesting a possible predisposition for such sounds, while others are more language-specific. Similarly, different populations of the same songbird species can share elements but also sing different 'dialects' and song elements ('notes') [5–7]. In addition to geographical variation, song can differ between individuals in the same population. Zebra finch song, for instance, consists of different types of elements (figure 1), and individual birds may vary in which elements are used and how they are combined. Some element types are more common between individuals than others. A recent
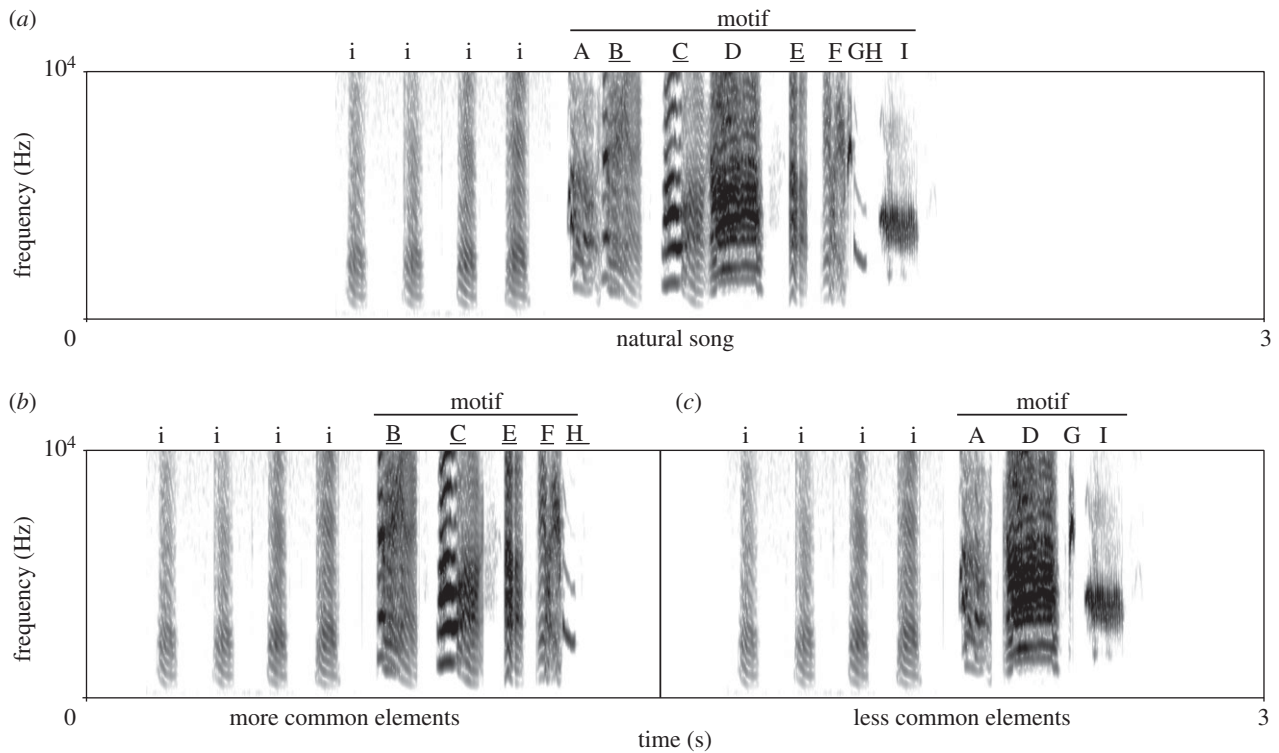
Figure 1. Examples of one pair of stimuli, constructed from (*a*) one original song. From the original song, more common (MC) element types were selected (indicated by underlined letters) and combined into (*b*) an artificial 'common song' stimulus and similarly (*c*) 'uncommon' song stimuli were constructed using less common (LC) elements of the same original song. Both stimulus types started with four introductory notes from the original song (indicated by 'i').

study of song elements present in 13 different zebra finch populations showed that, although all element types occur in all populations, there is variation in the proportion of some elements and no variation in the proportion of other elements between populations. In addition, individuals within a population can differ substantially in which elements they share [8]. How can these individual- and population-level differences be explained by variation during the vocal learning process? And can commonalities that are also found among populations be related to the presence or the absence of perceptual predispositions that guide within-species vocal learning? Patterns or elements that are common across individuals and populations might indicate species' general predispositions facilitating selective learning and constraining vocal variation. The population- or individual-specific elements might not be based upon such predispositions but might instead arise from plasticity in the learning process, allowing deviating elements to develop and be learned and maintained by cultural transmission. If both predispositions and learning by experience can affect the sounds in a population, how are these processes entwined during development?

Evidence for perceptual predispositions in songbirds so far has mainly come from experiments showing a preference to learn conspecific sounds over heterospecific ones in studies involving isolate rearing and tape tutoring. In most studies, adult song production is used as a measure of learning or selective preference. Only a few studies have examined perceptual predispositions in naive birds. An experiment in which juvenile zebra finches (*T. guttata*) could elicit exposure to either conspecific or heterospecific song by hopping on a perch showed that birds hopped more on the perch generating conspecific song than on the one generating heterospecific song [9,10]. In the white-crowned sparrow (*Zonotrichia leucophrys*), fledglings produced more begging calls in response to conspecific song than to heterospecific song [11,12]. Preference for the

birds' own subspecies over other subspecies was not confirmed, but exposure to songs of the birds' own subspecies led to better discrimination than experience with another subspecies's song. This outcome suggests that the perceptual system is more attuned to acoustic features within a subspecies [13]. Moreover, while white-crowned sparrows were found to respond equally to different conspecific phrase types prior to song exposure [14], tutoring with each of these phrase types showed that the universal white-crowned sparrow's introductory whistle functions as a cue for song learning. Songs (even heterospecific ones) are better copied when they contain these universal whistles [15], and whistles are preferably copied [16]. This suggests that in addition to a preference for conspecific versus heterospecific vocalizations, there are perceptual predispositions for certain *within-species* element types. This has also been shown in a study on grasshopper sparrows, where naive female fledglings responded more to one conspecific song type ('buzz': simple structure and uniform across individuals) than another conspecific one ('warble': complex and possibly individually specific [17]). While the presence of *within-species* perceptual predispositions in male songbirds is suggested by the above-mentioned studies, clear demonstrations of their presence are lacking, especially their role in song development and their relation to the abundance of elements in a population.

Selective song production is usually considered as an indication for the presence of perceptual preferences in song learning. Yet the few studies that attempted to measure the perceptual preferences before song exposure show no direct relationship between perception and production. How relevant are perceptual predispositions for guiding song learning? Can experience change the perceptual preferences? How do these predispositions and experience interact and affect final song production? In the present study, we address these questions. Perceptual preferences are tested in

juvenile male zebra finches at different stages of development. By testing the birds before hearing song and by manipulating subsequent exposure, we disentangle the effects of possible predispositions and auditory song experience. We also examine the similarity of the acquired songs to those heard during exposure, and discuss how predispositions and vocal learning relate to the distribution of song elements over populations and individuals.

We use the zebra finch in our study, which is the most prominent model species in the world for studying vocal learning. It is also a model for comparative studies on song development in birds, and language and speech development in humans [1,2]. Our study is also relevant from a comparative perspective, as the role of experience-independent and -dependent processes and their influence on typological patterns in human infants are strongly debated. In infants, the relative influences of these processes are hard to disentangle. The zebra finch provides an excellent model to examine the influence of predispositions and auditory experience on vocal learning in a controlled, experimental way.

2. Material and methods

The birds were reared by their mothers and were tested for their auditory preference at age 37 ($\pm$2), 47 ($\pm$2) and 57 ($\pm$2) days post hatch (dph). They were tutored for 30 days and songs were recorded afterwards, according to the following procedures.

(a) Subjects and housing

For the experiments, 16 male wild-morph domesticated zebra finches from an out-bred breeding colony at Leiden University, The Netherlands, were used. The birds were kept at 20–22°C and 55–65% humidity on a 13.5 L : 10.5 D schedule. Food, water and a cuttlebone were available ad libitum.

At the age of 8 dph ($\pm$2), the young birds and their mother were moved into a room where no adult males were present. At the age of 37 dph ($\pm$2), the birds received the first preference test. Each bird was moved to the preference cage the day before the test in order to acclimatize to the new cage and to isolation. After the test, the birds were isolated in sound-attenuated chambers for long-term song exposure. In the sound-attenuated chambers, food, water and cuttlebone were available ad libitum. The light : dark schedule in the sound attenuation chambers was 13.5 : 10.5 h, with a temperature of 21–24°C and a humidity of 50–55%.

(b) Stimuli

Unlike the songs of some other species, zebra finch song elements do not show fully discrete element types, although different clusters of element types can be identified on both visual and calculated similarities [8,18]. The classification of elements, as common or not, was based on the literature describing the element types found across several populations [19–23] or reporting rare elements [21]. These data are summarized in the electronic supplementary material, table S1, which was used to estimate which elements were more or less common among populations or individuals. As can be seen in the electronic supplementary material, table S1, the frequencies are a continuum rather than a discrete distribution of common and uncommon elements. However, in general, we can consider stacks, slides, short slides and tones to be more common, and high notes (especially inspiratory ones), trills, high sweeps, noisy elements and elements that do not clearly fall into a category to be less common elements. Using this distinction, we constructed 'common' and 'uncommon' songs from natural songs produced by normally reared birds in the Leiden University zebra finch colony. From each of eight natural songs, two versions of a motif were created: one 'common' version, using the common elements of the song and one 'uncommon' version using the uncommon elements from the same original song (figure 1). In this way, individual factors like voice characteristics cannot cause a difference in preference between common and uncommon stimuli. Each stimulus song consisted of four introductory notes followed by five motifs. Rearranging the elements does, of course, disrupt the original element sequences, but zebra finches have no strict transition rules between different elements [8,18], so we did not consider this a problem.

The tutor songs used for exposure were selected from the same set of stimuli used for the preference test and each bird was tutored with a different song. All stimuli were modified using PRAAT sound analysis software (v. 5.1.41 for Windows) and had a mean motif duration of 0.385 s (range 0.284–0.519 s) for common songs and 0.379 s (range 0.276–0.548 s) for uncommon songs. Also the number of syllables and elements (we refer to elements as the smallest units of the song, separated by abrupt changes in frequency or amplitude or silent intervals; syllables are defined as within motif units separated by silent intervals) were similar (mean 7.0 with range 4.0–10.0 for common elements and mean 8.3 with range 6.0–11.0 for uncommon elements; mean 3.7 with range 3.0–4.0 for common syllables and mean 4.2 with range 2–5 for uncommon syllables). All 16 stimuli were RMS equalized.

(c) Exposure

The birds remained in isolation in sound-attenuated rooms while tutored with either 'common' or 'uncommon' song via a speaker. The amount of exposure for all birds was the same (approx. 180 bouts per day; five motifs per bout).

After the first day of preference testing, the birds were moved to a sound-attenuated isolation chamber where exposure (tutoring) started the next day. Every bird was tutored approximately 20 times per hour (random timing) from 7.15 to 13.15 h and 10 times per hour from 13.15 to 19.15 h. Thus, the amount of exposure for all birds was the same: (6 h × 20 =) 120 + (6 h × 10 =) 60) = 180 bouts per day, with five motifs per bout. Birds were tutored each day from age 37 to 67 dph, with the exception of the days of preference testing.

(d) Preference tests

Birds were tested for their preference at 37 dph (before tutoring started), 47 dph and 57 dph, each time using the same four sets of stimuli (i.e. the same blocks, see below), including the stimuli to which the birds were exposed during the tutor phase.

Preferences were measured using a phonotaxis set-up [24]; a cage with one speaker on each side, alternating song playbacks with more common element types from one speaker and less common element types from the other. The time spent on the left or the right side of the cage was used as a measure of preference (measured from the first response after playback). When the birds were in the centre of the cage (a neutral zone), this was not included in the response time.

Each test consisted of four blocks on one day (always in the morning when birds were most active), each block with a different pair of common and uncommon stimuli. Two sets of eight stimuli (four pairs of common and uncommon song) were used; eight birds were tested with the first set and eight birds with the second set. The common and uncommon songs of a pair were derived from the same natural song (figure 1). One block consisted of 14 min alternating each minute between common song played from one speaker and uncommon song from the other speaker. Each minute contained seven identical songs. The order of the type of stimulus and side from which they were played back

was counterbalanced between blocks and between subjects. Songs were broadcast at approximately 70 db. After each block, the bird had a break of 45 min before the next block started.

Video recordings of each test were analysed while blind to the stimuli using ELAN software (v. 3.8.1, http://www.lat-mpi.eu/tools/elan; Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands [25]). When birds did not show any response during a given block, this block was excluded from further analysis. Owing to the lack in response in all four blocks, four birds had to be excluded from the 37 dph analysis, thus the analysis was based on 12 birds.

(e) Song analysis

Birds' songs were recorded at age 120 dph or older (when they did not sing at 120 dph). Of the 16 birds in the experiment, six birds were housed in isolation until adulthood, the day of recording. Three of these six birds were tutored with common elements, the other three with uncommon ones. For logistical reasons, the others were housed in isolation until day 66–75 and afterwards in cages grouped with birds from the same experiment. Birds in these cages were in auditory but not visual contact with the rest of the colony.

From each bird, the predominant motif was selected, and similarity measures (%) between the subject (hereafter 'tutee') song, and the artificial tutor song were measured using SOUND ANALYSIS PRO (SAP) 2011 [26]. Similarity for the different acoustic features—pitch, frequency modulation (FM), amplitude modulation (AM), entropy and goodness of pitch—for each song comparison was also analysed separately in Euclidian distances using SAP. Smaller Euclidian distances indicate higher similarity.

For each song and feature analysis, two comparisons were made: (i) the difference between tutor–tutee similarity and the similarity between the tutee and a random control song of the same type (common or uncommon); and (ii) the difference between tutor–tutee similarity and the similarity between the tutee and the tutor song's counterpart, meaning the song originating from the same natural song (figure 1a) but constructed of the elements of the other category (figure 1b).

(f) Statistics

All statistical analyses were performed using R v. 2.11.0. Linear mixed effect models were performed for preference test data using the *nlme* package for R, v. 3.1–96 [27]. 'Subject' was included as a random factor, with 'block' (the four songs tested per bird per age) nested within bird. Deletion *p*-values were accomplished by comparing models with and without the variables of interest using the ANOVA method in R. Model assumptions (normally distributed errors and lack of heteroscedasticity) were always verified after model selection.

Statistical analyses for song similarity measures were performed using Wilcoxon signed-rank tests and Kruskal–Wallis rank-sum tests.

3. Results

(a) Perceptual preference for more common element types in male birds naive to song

Juvenile male zebra finches were reared only by their mother, from approximately 8 dph, well before the start of the sensitive phase for song learning [28–30]. As female zebra finches do not sing, we thus created relatively natural rearing conditions where zebra finches are not exposed to song. The birds were tested at 37($\pm$2) dph, when they were naive to
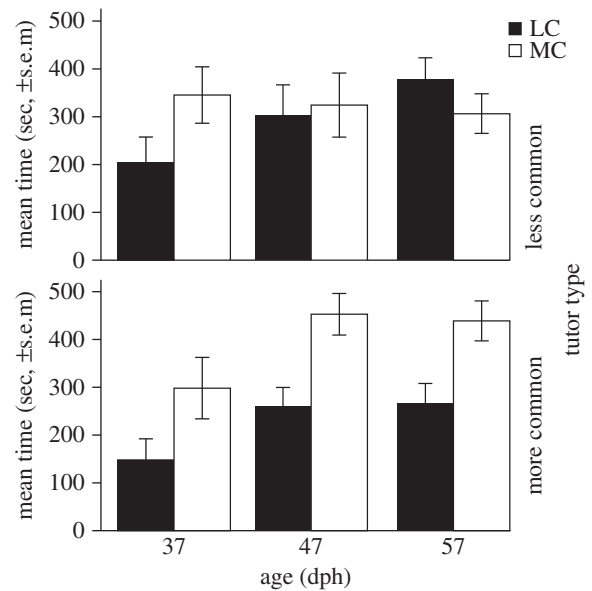

Figure 2. Preferences independent of and dependent on song exposure. The preference was measured as time in seconds ($\pm$ s.e.m.) near the speaker broadcasting songs constructed with MC elements or songs constructed with LC elements. For zebra finches at 37 dph (before tutoring; see Material and methods) this preference is higher for more common elements (white bar) than for less common elements (black bar). At 57 dph (after tutoring), birds prefer the type of sounds they have been tutored with; birds tutored with MC elements (lower panel) prefer songs with more common elements and birds tutored with less common elements (top panel, LC) preferred less common element types.

song, to see if they have a preference for more common or less common elements (see Material and methods for details).

The results show that juvenile males significantly prefer songs containing common elements (hereafter 'common songs') over those containing uncommon elements (hereafter 'uncommon songs' ($n = 12$, deletion $p < 0.01$; electronic supplementary material, S2; figure 2). So, initially, before song exposure, males have a bias for more common zebra finch song element types.

(b) Preference for tutor song at 57 dph

After the first preference tests, the birds were exposed to (tutored with) either a common or an uncommon song until approximately 67 dph, and preference tests were repeated at 47 ($\pm$2) dph and 57 ($\pm$2) dph in order to test the effect of the subsequent exposure on the birds' preferences. Each preference test again consisted of four blocks (four pairs of common versus uncommon song), one of which included the tutor song. By comparing preference for the pair including the tutor song to the other three pairs, we could test whether the preference at 57 dph was specific for the tutor song only or other songs of the same type (common/uncommon) are preferred.

A significant four-way interaction was found between preference (common/uncommon song), age (37, 47, 57 dph), tutor type (common/uncommon tutor song) and block (tutor/non-tutor). This indicates that the preference changes with age, depending on the type of tutoring and whether it is the tutor song or not ($n = 16$, deletion $p < 0.05$; figure 2; electronic supplementary material).

Because four-way interactions can be hard to interpret, and in order to confirm the interactions at lower levels,
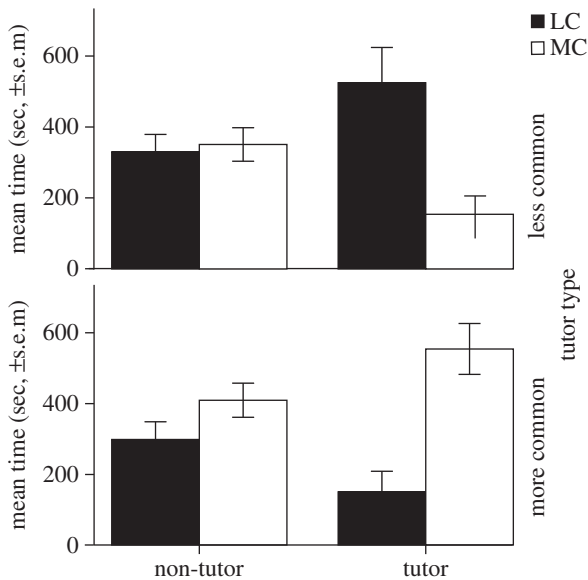
Figure 3. Specific preference for tutor song at 57 dph. Birds tutored with songs containing less common elements (top panel, LC) as well as birds tutored with songs containing more common elements (lower panel MC) prefer their tutor song. Birds lost their initial preference, thus they do not show a preference for common or uncommon songs that were not their tutor song.

separate analyses were done for the 57 dph data. This revealed a significant interaction between preference and tutor type at 57 dph for the blocks including tutor songs ($n = 13$, deletion $p < 0.01$), but not for the other three blocks ($n = 16$, deletion $p > 0.05$), suggesting that the later preference was specific to the tutor song. When we look at the two tutor groups separately, the interaction between preference and block (tutor/non-tutor) was significant for both the birds tutored with common songs ($n = 8$, deletion $p < 0.05$) and birds tutored with uncommon songs ($n = 8$, deletion $p < 0.01$; figure 3). These findings indicate that birds specifically prefer their tutor song (and thus song exposure) at 57 dph, in line with previous findings in adult birds [31,32].

(c) Experience affects song learning

After the last preference test at 57 dph, the birds were tutored in isolation for 10 more days. Adult birds' songs were recorded (approx. 120 dph or older [28]), and similarity between the tutee song and the tutor song was measured using SAP [26]. The similarity scores for both groups were not very high, as expected for tape tutoring with unnatural songs (see figure 4 and electronic supplementary material S3 for spectrograms). Nevertheless, these measurements revealed that the tutee song similarity to the tutor song was higher than similarity to a control song of the same type (common/uncommon song, paired Wilcoxon: $n = 16$, $p < 0.05$). We also compared similarity to the tutor song with similarity to its counterpart, derived from the same original natural song but belonging to the other song category (figure 1). Thus, for a song from a tutee tutored with common songs (figure 1b), a comparison was made for similarity to the common tutor song (figure 1b) and the uncommon counterpart of that song (figure 1c), and vice versa for the tutee tutored with uncommon song. This comparison revealed an overall tendency for higher similarity to the tutor song than to its counterpart (paired Wilcoxon: $n = 16$, $p = 0.06$). Most notable, however, was the tutor group

difference found for this latter comparison. Tutees tutored with common songs showed higher similarity to the tutors than to their counterparts (paired Wilcoxon: $n = 8$, $p = 0.02$; average SAP score tutor–tutee comparison: 48%), whereas this difference was not significant for the group tutored with uncommon songs ($n = 8$, $p > 0.05$; average SAP score tutor–tutee comparison: 27%; group difference: Kruskal–Wallis: $n = 16$, $p = 0.02$; see figure 4 and electronic supplementary material S3 for examples of spectrograms). In other words, evidence for similarity to the tutor song is stronger for tutees tutored with common song.

For logistical reasons, 10 of the 16 tutees were housed in isolation until day 66–75 and afterwards in cages grouped with birds from the same experiment. Birds in these cages were in auditory but not visual contact with the rest of the colony, so learning from other birds in the colony or from each other cannot be completely excluded for this group. No significant differences were found, however, between tutor–tutee similarities of songs from the birds housed in isolation until 120 days and those housed in isolation until 66–75 days (Kurskal–Wallis: $p > 0.05$).

Visual inspection of the sonograms for similarity of song elements ('notes') did not reveal significant differences between birds exposed to common and those exposed to less common songs, as was shown by SAP similarity (Kruskal–Wallis: $p > 0.05$). Thus, the similarity between tutor and tutee song might be due to differences in acoustic features rather than from carefully copying elements or syllables. We also tested whether the higher similarity measures in birds tutored with common songs are due to specific acoustic features being copied better by the birds tutored with common songs. SAP similarity measures for acoustic features separately (Euclidian distances for pitch, FM, AM entropy and goodness of pitch) revealed no significant differences between songs of the two groups of tutees (Kruskal–Wallis: $p > 0.05$ for all five features) when comparing the similarity to the tutor with similarity to another song of the same type. If similarity to the tutor is compared with the similarity to the counterpart (from the same natural song but with elements of the other type), we do find significant differences between tutor groups. The difference between Euclidian distances of the tutor–tutee comparison (difference = counterpart–tutee distance – tutor–tutee distance) was higher for the birds tutored with common songs than those tutored with less common ones for pitch, FM and entropy (Kruskal–Wallis: $n = 16$, $p < 0.001$, $p < 0.05$ $p < 0.005$, respectively). In other words, songs from birds tutored with common songs resembled their tutor songs more for these parameters than the songs of birds tutored with uncommon songs did. Goodness of pitch also showed a significant difference between tutor groups, but in the other direction: birds tutored with less common songs had higher difference in Euclidian distance than birds tutored with common songs (Kruskal–Wallis: $n = 16$, $p < 0.01$). Thus, only for goodness of pitch did songs from birds tutored with less common songs show higher similarity to the tutor song than songs from birds tutored with common song.

4. Discussion

Our findings suggest that juvenile birds naive to song have perceptual predispositions that make some elements of conspecific song more attractive than others. In addition to this, we show
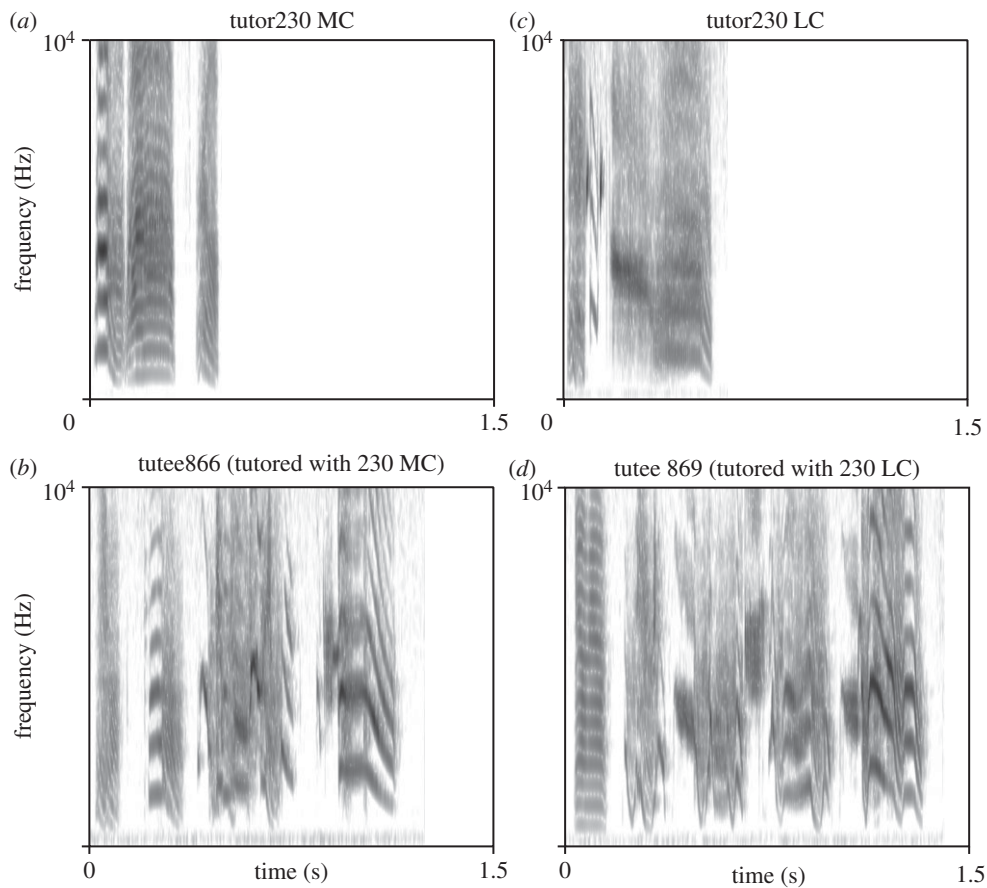
Figure 4. Spectrograms of the tutor–tutee combination. Spectrograms of the highest similarity for the group tutored with MC songs are represented for (*a*) tutor and (*b*) tutee, and for the groups tutored with LC songs, for (*c*) tutor and (*d*) tutee. All other spectrograms are presented in the electronic supplementary material S3. All songs in this figure and in the electronic supplementary material S3 are presented without introductory notes. Tutees' songs were generally longer than tutors' songs, probably because our artificial tutors' songs have approximately half as many elements as the tutors' songs and are thus unnaturally short.

that these perceptual preferences can be modified by exposure, even resulting in a preference for initially non-preferred elements. Furthermore, while young birds can incorporate both common and uncommon song elements in their later songs, common elements are more likely to be copied.

Altogether, these observations provide evidence of a mechanism that may explain the species-wide presence and maintenance of particular types of elements in a vocal learning species, as follows. The perceptual bias present in naive juvenile males can guide the learning process, directing the learner's attention towards particular conspecific vocal elements. This is likely to result in including these element types in the bird's later song production. The result of this process will be that these elements are more likely to be maintained in a population and hence become, or stay, more common. Over generations this process is likely to cause stabilization of vocal patterns containing these common features. There is an interesting parallel here with a mechanism that has been proposed for language, which has been described in terms of markedness, suggesting unmarked ('universal') sounds are acquired early in development and marked ones later. Unmarked sounds are more likely to (re)occur and to be maintained in languages. Acquisition of unmarked sounds [33], similar to acquisition of common song elements, might be driven by processes independent of linguistic input [34].

In addition to processes resulting in song conformity, if a young bird is exposed to elements for which there is initially no perceptual bias, this exposure can still result in copying

such elements. This may facilitate song variation, and the appearance and maintenance of these elements in a population might depend on local factors and chance (drift) affecting cultural transmission.

Interestingly, our study provides empirical support for a mechanism suggested by a study by Feher *et al.* that elegantly showed a process of vocal convergence on more common species-specific song features over generations [35]. Zebra finch males reared without song exposure, which produce aberrant song, were used as tutors for a second 'tutor generation'. The latter birds were again used as tutors for a third tutor generation, and so on. Within three to four generations, songs evolved towards songs with wild-type characteristics. The strongest change was already in the first generation. The tutees copied most of the elements of the aberrant song, but also induced alterations to their tutor song. It was suggested that selective or biased imitation resulted in accumulation of these alterations over tutor generations [35,36]. As a result, the songs in the later generations became more similar to wild-type songs. Although the songs in the experiments by Feher *et al.* [35] contained both common and less common elements, and were not analysed to examine whether some elements were more likely than others to be present in the final songs, we suggest that the biases we demonstrated could also have affected the direction of element changes, driving the elements towards becoming more similar to preferred (and also more normal and common) elements. We do not want to suggest that the existence of uncommon elements would be eliminated

by the bias for common elements, however. In fact, experience seems able to partly override the initial perceptual biases for common elements (figure 2, 47 and 57 dph), thereby allowing both common and uncommon elements to persist. Nevertheless, the stronger bias for common elements would shift the frequency of occurrence within populations, or possibly even within birds, towards more common elements.

While the observed bias is independent of song exposure, we cannot fully exclude that it is independent of any acoustic exposure, as it may have been affected by the mothers' vocalizations during rearing. Even though females do not sing, they do produce calls that may affect auditory preferences. There is some evidence for perceptual preferences being formed before 35 dph [37,38]. The female's call has some features in common with some of the more common elements (stacks), but certainly not all of them (for instance slides). Future research could elucidate this issue by examining in more detail which specific acoustic features make certain element types attractive and by using muted females to rear the subjects. Regardless of the cause of the sensitivity, the effect of the bias is independent of *song* exposure, and thus remains relevant in terms of development and evolution, and may result in maintenance of such sounds in a species.

While the present experiment can provide insight into the evolutionary consequences of the perceptual biases and the developmental processes involved, less can be said about the evolutionary origin of the perceptual biases. Nevertheless, our findings may be interesting for future research on mate attraction. In zebra finches, songs are supposed to function primarily in mate choice and pair bonding. There is ample evidence that female zebra finches prefer specific songs or song features over others [39]. These preferences may concern the presence of certain general features of songs. For example, naive and normally reared female zebra finches prefer conspecific song [9,40], which might help to maintain the species specificity of male songs. On the other hand, females prefer tutored song (normal quality) to untutored song (abnormal quality) [40] and larger element repertoires over smaller ones [21]. This latter preference may drive the use of additional uncommon song elements, and thus male song plasticity, since larger element repertoires are more likely to include more uncommon elements in addition to the common ones. However, little is known about specific element types being preferred by females and how the presence of these contributes to attractiveness of the song as a whole. It is known that females do not systematically prefer songs with expiratory elements (classified as common) over songs with inspiratory high notes (here classified as uncommon) [21], but it would require more specific tests to examine whether female zebra finches differentially prefer (songs with) common or uncommon elements to get insight in the evolutionary dynamics from which the current male songs have arisen.

Possibly some perceptual or acoustic features are more relevant than others for the distinction between common and uncommon songs. Measuring similarity by visually counting the number of similar elements or syllables between tutor and tutee did not reveal a significant difference between tutoring with common or uncommon songs. Interestingly, an acoustic feature-based similarity analysis in SAP showed higher tutor–tutee similarity for birds tutored with common song elements. Tutor–tutee similarity measures for acoustic features separately showed that entropy, pitch and FM are more similar for birds tutored with common songs, whereas goodness of pitch was more similar for birds tutored with uncommon songs. This might indicate that different acoustic features are learnt in different ways: FM, pitch and entropy might be more strongly involved in predispositions and biased during early perception, whereas goodness of pitch might be acquired primarily by exposure-based learning. More research specifically disentangling the features would be useful to further explore these questions.

Our similarity scores (48% for tutor–tutee comparisons in the group tutored with more common songs and 27% for the birds tutored with less common songs) for both groups are lower than usual for SAP analyses of zebra finch songs. For comparison, natural tutor–tutee comparisons reveal a similarity score of over 60%, whereas random paired natural songs show similarities of 40% or lower [26]. The reason for this is likely to be the artificial structure of the tutoring songs in our study. Additionally, syntactic differences may play a role in both preference and song copying. Although zebra finches do not have a strict element sequence at the species level [18], it might be that common song elements are more naturally or easily combined with each other or with uncommon elements than are uncommon elements with each other. Future research could further explore whether certain combinations of elements or syllables are easier to learn than (or perceptually preferred over) others.

The finding that both an initial bias for more common elements and later experience affect song learning has a striking parallel in human infants. In infants (and possibly in songbirds [41]), early phonetic discrimination is universal and becomes more language-specific later on [42]. A similar developmental change can be observed for acquisition of syllable structure. In early language productions, the first syllables are of the consonant–vowel type, which is common across different languages. Subsequent development of novel syllable types is influenced by frequency of occurrence and may therefore also be experience-dependent [43]. Although there is a clear parallel, the distinction between more and less common elements is somewhat different from that between universal and non-universal speech sounds in humans. In zebra finches, there are clear individual differences within populations, while differences in sound inventories between populations are less clear [8]. By contrast, human speech sound inventories differ between languages and people speaking the same dialect usually make use of approximately the same phoneme inventory. It should be noted, however, that the use of different analytical methods for human language and birdsong makes a direct comparison difficult. Thus, the developmental mechanism may be similar (attention changing from more common to less common, i.e. from internal biases to external influences), but the eventual effect of the developmental plasticity due to the vocal learning may differ between humans and songbirds. If the developmental mechanism is indeed the same for birds and humans, then the implications described above may also hold for language evolution. Initial biases could maintain the universals in languages, whereas additional plasticity allows for learning language-specific patterns and facilitates cultural evolution.

References

1. Doupe AJ, Kuhl PK. 1999 Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* **22**, 567–631. (doi:10.1146/annurev.neuro.22.1.567)

2. Bolhuis JJ, Okanoya K, Scharff C. 2010 Twitter evolution: converging mechanisms in birdsong and human speech. *Nat. Rev. Neurosci.* **11**, 747–759. (doi:10.1038/nrn2931)

3. Adret P. 2004 In search of the song template. *Ann. NY Acad. Sci.* **1016**, 303–324.

4. Woolley SMN. 2012 Early experience shapes vocal neural coding and perception in songbirds. *Dev. Psychobiol.* **54**, 612–631. (doi:10.1002/dev.21014)

5. Marler P, Tamura M. 1964 Culturally transmitted patterns of vocal behavior in sparrows. *Science* **146**, 1483–1486. (doi:10.1126/science.146.3650.1483)

6. Petrinovich L, Baptista LF. 1984 Song dialects, mate selection, and breeding success in white-crowned sparrows. *Anim. Behav.* **32**, 1078–1088. (doi:10.1016/S0003-3472(84)80224-9)

7. Kroodsma DE. et al. 1999 Geographic variation in Black-capped Chickadee songs and singing behavior. *Auk* **116**, 387–402. (doi:10.2307/4089373)

8. Ter Haar SM. 2013 Birds and babies: a comparison of the early development in vocal learners. PhD thesis, Leiden University, The Netherlands.

9. Braaten RF, Reynolds K. 1999 Auditory preference for conspecific song in isolation-reared zebra finches. *Anim. Behav.* **58**, 105–111. (doi:10.1006/anbe.1999.1134)

10. Braaten RF, Petzoldt M, Cybenko AK. 2007 Recognition memory for conspecific and heterospecific song in juvenile zebra finches, *Taeniopygia guttata*. *Anim. Behav.* **73**, 403–413. (doi:10.1016/j.anbehav.2006.08.009)

11. Nelson DA, Marler P. 1993 Innate recognition of song in white-crowned sparrows—a role in selective vocal learning. *Anim. Behav.* **46**, 806–808. (doi:10.1006/anbe.1993.1258)

12. Soha JA, Marler P. 2001 Cues for early discrimination of conspecific song in the white-crowned sparrow (*Zonotrichia leucophrys*). *Ethology* **107**, 813–826.

13. Nelson DA. 2000 A preference for own-subspecies' song guides vocal learning in a song bird. *Proc. Natl Acad. Sci. USA* **97**, 13 348–13 353. (doi:10.1073/pnas.240457797)

14. Whaling CS, Solis MM, Doupe AJ, Soha JA, Marler P. 1997 Acoustic and neural bases for innate recognition of song. *Proc. Natl Acad. Sci. USA* **94**, 12 694–12 698. (doi:10.1073/pnas.94.23.12694)

15. Soha JA, Marler P. 2000 A species-specific acoustic cue for selective song learning in the white-crowned sparrow. *Anim. Behav.* **60**, 297–306. (doi:10.1006/anbe.2000.1499)

16. Soha JA, Marler P. 2001 Vocal syntax development in the white-crowned sparrow (*Zonotrichia leucophrys*). *J. Comp. Psychol.* **115**, 172–180.

17. Soha JA, Lohr B, Gill DE. 2009 Song development in the grasshopper sparrow, *Ammodramus savannarum*. *Anim. Behav.* **77**, 1479–1489. (doi:10.1016/j.anbehav.2009.02.021)

18. Lachlan RF, Peters S, Verhagen L, ten Cate C. 2010 Are there species-universal categories in bird song phonology and syntax? A comparative study of chaffinches (*Fringilla coelebs*), zebra finches (*Taeniopygia guttata*), and swamp sparrows (*Melospiza georgiana*). *J. Comp. Psychol.* **124**, 92–108. (doi:10.1037/a0016996)

19. Holveck MJ, de Castro ACV, Lachlan RF, ten Cate C, Riebel K. 2008 Accuracy of song syntax learning and singing consistency signal early condition in zebra finches. *Behav. Ecol.* **19**, 1267–1281. (doi:10.1093/beheco/arn078)

20. Sturdy CB, Phillmore LS, Weisman RG. 1999 Note types, harmonic structure, and note order in the songs of zebra finches (*Taeniopygia guttata*). *J. Comp. Psychol.* **113**, 194–203. (doi:10.1037/0735-7036.113.2.194)

21. Leadbeater E, Goller F, Riebel K. 2005 Unusual phonation, covarying song characteristics and song preferences in female zebra finches. *Anim. Behav.* **70**, 909–919. (doi:10.1016/j.anbehav.2005.02.007)

22. Zann R. 1993 Structure, sequence and evolution of song elements in wild Australian zebra finches. *Auk* **110**, 702–715. (doi:10.2307/4088626)

23. Price PH. 1979 Developmental determinants of structure in zebra finch song. *J. Comp. Physiol. Psych.* **93**, 260–277. (doi:10.1037/h0077553)

24. Holveck M-J, Riebel K. 2007 Preferred songs predict preferred males: consistency and repeatability of zebra finch females across three test contexts. *Anim. Behav.* **74**, 297–309. (doi:10.1016/j.anbehav.2006.08.016)

25. Sloetjes H, Wittenburg P. 2008 Annotation by category—ELAN and ISO DCR. In *Proceedings of the 6th International Conference on Language Resources and Evaluation* (*LREC 2008*). See www.lrec-conf.org/proceedings/lrec2008.

26. Tchernichovski O, Nottebohm F, Ho CE, Pesaran B, Mitra PP. 2000 A procedure for an automated measurement of song similarity. *Anim. Behav.* **59**, 1167–1176. (doi:10.1006/anbe.1999.1416)

27. Pinheiro JC, Bates DM, DebRoy S, Sarkar D, Team a.t.R.C. 2009 nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-96, http://www.R-project.org.

28. Jones AE, ten Cate C, Slater PJB. 1996 Early experience and plasticity of song in adult male zebra finches (*Taeniopygia guttata*). *J. Comp. Psychol.* **110**, 354–369. (doi:10.1037/0735-7036.110.4.354)

29. Eales LA. 1987 Song learning in female-raised zebra finches: another look at the sensitive phase. *Anim. Behav.* **35**, 1356–1365. (doi:10.1016/S0003-3472(87)80008-8)

30. Eales LA. 1985 Song learning in zebra finches—some effects of song model availability on what is learnt and when. *Anim. Behav.* **33**, 1293–1300. (doi:10.1016/S0003-3472(85)80189-5)

31. Houx BB, ten Cate C. 1999 Song learning from playback in zebra finches: is there an effect of operant contingency? *Anim. Behav.* **57**, 837–845. (doi:10.1006/anbe.1998.1046)

32. Houx BB, ten Cate C. 1999 Do stimulus-stimulus contingencies affect song learning in zebra finches (*Taeniopygia guttata*)? *J. Comp. Psychol.* **113**, 235–242. (doi:10.1037/0735-7036.113.3.235)

33. De Lacy P. 2006 *Markedness: reduction and preservation in phonology*. Cambridge, UK: Cambridge University Press.

34. Jakobson R. 1941 *Child language, aphasia and phonological universals*, The Haque, The Netherlands: Mouton.

35. Feher O, Wang HB, Saar S, Mitra PP, Tchernichovski O. 2009 *De novo* establishment of wild-type song culture in the zebra finch. *Nature* **459**, 564–568. (doi:10.1038/nature07994)

36. Feher O, Tchernichovski O. 2013 Vocal culture in songbirds: an experimental approach to cultural evolution. In *Birdsong, speech, and language: exploring the evolution of mind and brain* (eds. JJEM Bolhuis) pp. 143–156. Cambridge, MA: MIT Press.

37. Amin N, Doupe A, Theunissen FE. 2007 Development of selectivity for natural sounds in the songbird auditory forebrain. *J. Neurophysiol.* **97**, 3517–3531. (doi:10.1152/jn.01066.2006)

38. Roper A, Zann R. 2006 The onset of song learning and song tutor selection in fledgling zebra finches. *Ethology* **112**, 458–470. (doi:10.1111/j.1439-0310.2005.01169.x)

39. Riebel K. 2009 Song and female mate choice in zebra finches: a review. *Adv. Study Behav.* **40**, 197–238. (doi:10.1016/S0065-3454(09)40006-8)

40. Lauay C, Gerlach NM, Adkins-Regan E, Devoogd TJ. 2004 Female zebra finches require early song exposure to prefer high-quality song as adults. *Anim. Behav.* **68**, 1249–1255. (doi:10.1016/j.anbehav.2003.12.025)

41. Miller-Sims VC, Bottjer SW. 2012 Development of auditory-vocal perceptual skills in songbirds. *PLoS ONE* **7**, e52365. (doi:10.1371/journal.pone.0052365)

42. Tsao FM, Liu HM, Kuhl PK. 2006 Perception of native and non-native affricate-fricative contrasts: cross-language tests on adults and infants. *J. Acoust. Soc. Am.* **120**, 2285–2294. (doi:10.1121/1.2338290)

43. Levelt CC, Schiller NO, Levelt WJM. 2000 The acquisition of syllable types. *Lang. Acquisit.* **8**, 237–264. (doi:10.1207/S15327817LA0803_2)