Chapter 3. Evolution of ISO song over multiple generations

Background & Rationale

In the previous chapter we showed that the imitation of ISO songs by young zebra finches is more similar to WT songs than the song of the ISO tutor. We observed a shift towards WT song features at three timescales of song structure: in spectral features (frequency modulation, amplitude modulation and goodness of pitch), note-level features (such as note & syllable length and duration of acoustic state), and bout-level features (syllable abundance and song rhythm). We saw that rather than preferentially imitating WT-like syllables, pupils changed the ISO-like syllables to be more WT-like, and we call this effect imitation biases to indicate that these biases must occur during the process of imitation (as opposed to production biases, or selective imitation). The biases that drove the changes were quite similar across tutors and across pupils.

The next question that arises is what happens to the song beyond the first generation. Perhaps, the first generation birds have revealed the full extent of innate biases, so the ISO song has been changed as much as it can be towards WT. In this case, the changes would have been a reaction to an overwhelmingly abnormal song stimulus, but pupils of first generation birds would imitate faithfully. However, we saw that the songs of the first generation birds are still as different from WT songs as from ISO songs, so it is possible that the progression towards WT-like features will continue if we let later generations of pupils learn the songs. In the present chapter, we will follow the subsequent evolution/development of the songs as they are passed down over generations of song learners and try to answer the question whether the progression towards WT-like features continues over multiple learning generations. As noted in the Introduction, cumulative cultural evolution is thought to occur in birds, but there is no direct empirical evidence. If it is true that the progression from ISO to WT takes multiple generations, then, by definition, song culture is cumulative – namely, the WT singing culture we observe cannot be achieved in a single generation, but it can be achieved by recursive learning.

Methods

Upon reaching adulthood and song crystallization, we used 4 of our pupils from the ISO tutor/pupil experiment to train young males individually. When adult, these first generation pupils become the tutors of the next generation pupils, and so on, recursively (Fig. 3.1). As before, pupils were assigned randomly to their tutors (minimizing genetic relatedness). The isolate tutor and his pupil were kept together for 90 days in a sound isolated chamber. Subsequently, we removed the tutor and recorded the pupil's song and compared it to that of his tutor. Following this, we placed a 30-day old juvenile male in the sound chamber housing the now mature pupil who became song tutors for the next generation pupils.

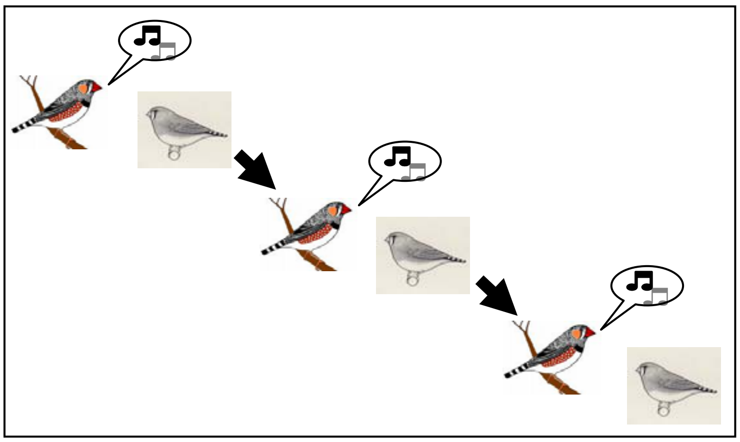

Figure 3.1 | Schematic diagram of multigenerational tutoring paradigm. The pupil (grey birds), upon reaching adulthood, becomes the tutor (colorful birds) of another juvenile. The tutoring continues recursively over several learning generations.

Table 3.1 presents the tutor and pupil identities and generation numbers for each bird in the multi-generation experiment. Note that the lineages and generations indicate recursive learning generations and not genetic relations.

	Tutoring Lineage 1	Tutoring Lineage 2	Tutoring Lineage 3	Tutoring Lineage 4
ISO Tutor	19 (Tutor 1)	1211 (Tutor 2)	1238 (Tutor 3)	1249 (Tutor 5)
1st generation pupil	1248	1402	1342	1439
2nd generation pupil	1326	1514	1571	1558
3rd generation pupil	1374	1606		
4th generation pupil	1535			
5th generation pupil	1621			

Table 3.1 | Tutoring lineages in multigenerational tutoring experiment.

Results

3.1 Visual assessment of imitation across tutoring generations

As before, we begin by examining the outcome of multi-generational training by inspecting the sonograms and comparing tutor songs to all of their successive pupils' songs.

Lineage 1 - Tutor 1 (Bird 19)

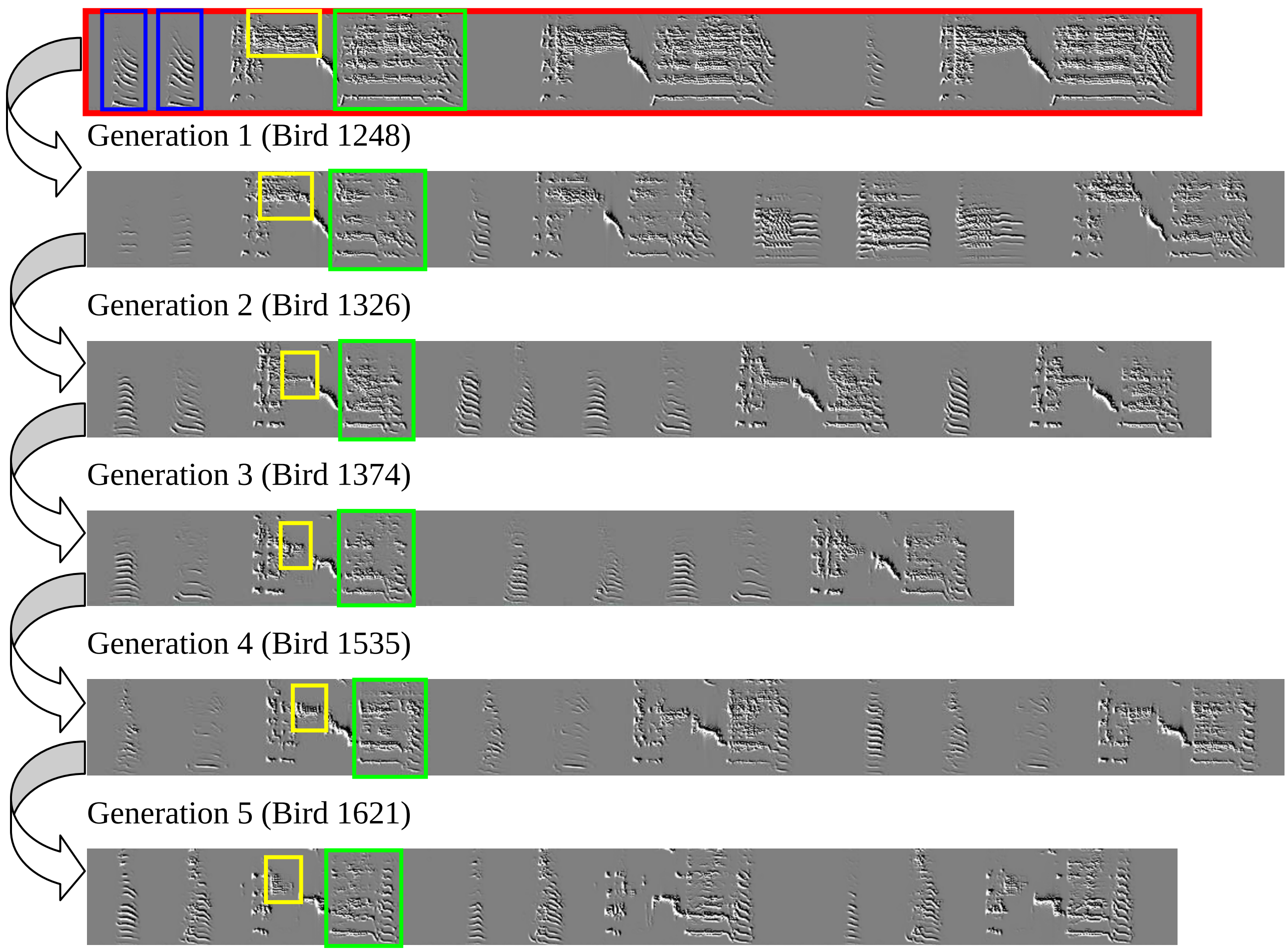

Figure 3.2 | Tutoring lineage of ISO Tutor 19. Notes that were altered are marked with different colored rectangles (blue, yellow and green in order in ISO song).

Following the song of Lineage 1 (Fig. 3.2) over a few generations reveals that much of the structure of the original tutor song was preserved by the pupils. Interestingly, the direction of vocal changes that the first generation pupil made persisted over generations and therefore, changes accumulated in the succeeding generations. For example, the copies of the long and monotonic ISO note (green rectangle) kept decreasing in duration in later generation pupils and its spectral structure was modified and differentiated gradually, so that by the fifth generation pupil, the note became shorter, but also differentiated into two distinct notes, one broadband and the second with pure harmonic structure. Similarly, the yellow note became much reduced in duration, bandwidth and in amplitude over generations of learners, and by the fifth generation, there was only a trace of the original syllable.

There were bout-level changes that were made to this song, as well. The bout got stretched out by the introduction of other, short (introductory-like) syllables between the renditions of the long syllable. The second generation pupil sang some medium duration modulated call-like syllables between the motifs. These syllables were not improvised, as they can be found in the isolate tutor's repertoire, although at much lower frequencies. His pupil sang them in nearly every bout. These syllables didn't become prevalent in the later generations pupils' songs.

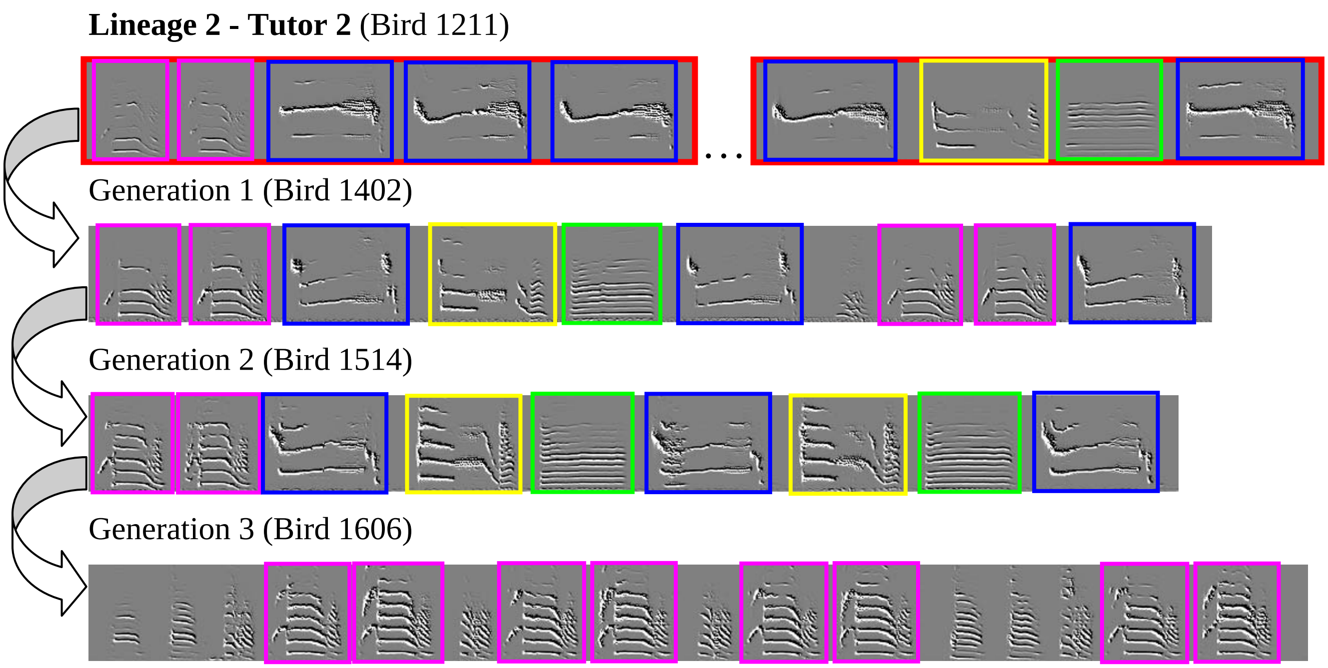

Figure 3.3 | Tutoring lineage of ISO Tutor 1211. Important syllables are marked in different colored rectangles (pink, blue, yellow and green in order).

In Lineage 2 (Fig. 3.3), the isolate tutor sang a very long, high-pitched syllable as the dominant element in his motif (blue rectangle). This song was very unusual both because of the spectral features of the dominant syllable and because it was repeated consecutively (15 times on average) by the ISO tutor. Two other syllable types of similar duration (in yellow and green), sung only once, were nested among these repetitions. Interestingly, the first generation pupil did not imitate the repetitions and constructed his motif out of the three long syllable types, singing them serially once and ending with the first one. Another syntactical reorganization happened in the second generation, when the song went from one rendition introduced by two renditions of the pink syllable, ppABCA, to ppABCABCA. This motif repetition and bout lengthening is reminiscent of the syntax changes that take place during development in an individual bird. The generation 3 pupil completely omitted the long syllables and sang a simple song whose spectral features were WT-like. This was one of the very few instances of selective imitation, where the pupil only imitated the more WT-like syllables and did not imitate the abnormal syllables at all.

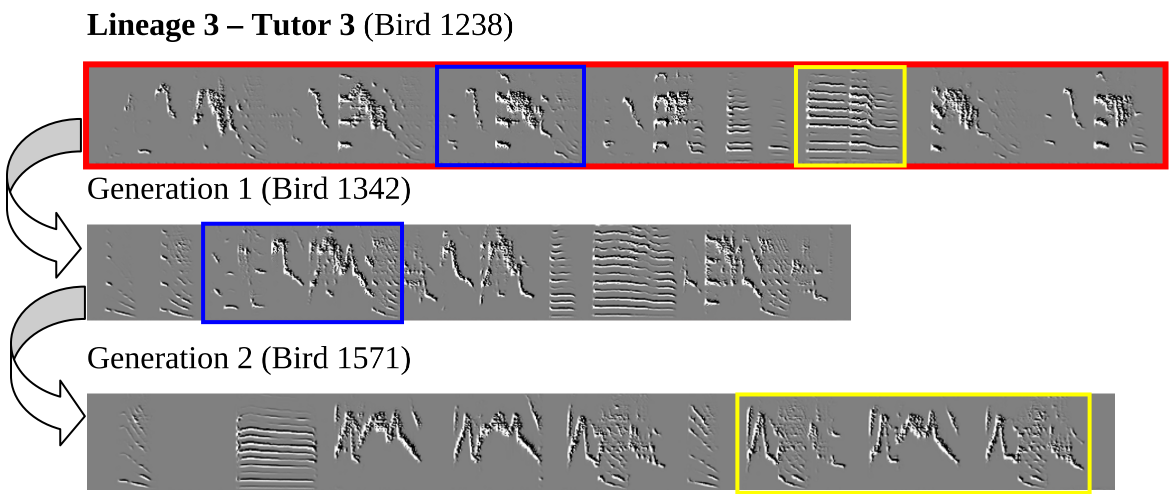

Figure 3.4 | Tutoring lineage of ISO Tutor 1238. Important syllables are marked in different colored rectangles (blue and yellow).

Tutor 3 (Fig. 3.4) sang a complex syllable (blue rectangle) that, based on its acoustic features and fast transitions, could be classified as a WT song. However, the syllable repetition is highly unusual and the stuttered harmonics (longest in yellow) that follow these repetitions are also abnormal. The repetition rate decreases in the song of the first generation pupil and the blue syllable becomes even more complex and differentiated so that it is not repeated exactly the same way each time. The second generation pupil progresses with the differentiation and breaks up the long syllable into two independent types. We still see some syllable repetition, but most of the time alternating syllables are sung in serial order (yellow).

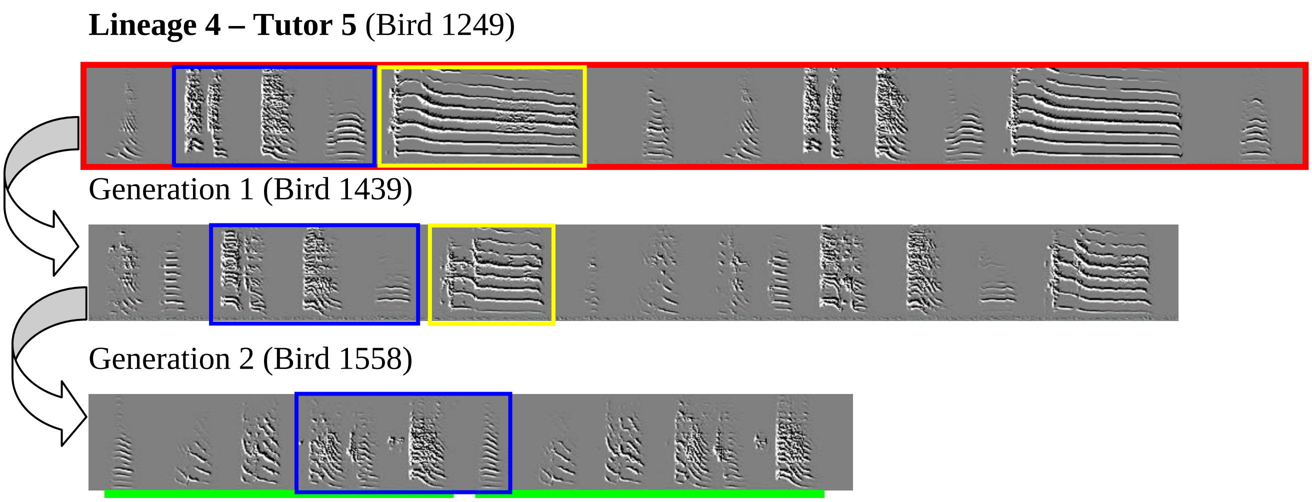

Figure 3.5 | Tutoring lineage of ISO Tutor 1249. Important syllables (yellow) or groups of syllables (blue) are marked in different colored rectangles.

Tutor 4 (Fig. 3.5) sang a simple song with some short, introductory-like notes (blue rectangle) followed by a short and a very long harmonic syllable (yellow rectangle). The first generation pupil imitated everything with high accuracy except he shortened the harmonic, but the second generation pupil omitted the harmonic and changed the structure of some of the short notes (first and last note in blue rectangle). In addition, he introduced a short, high-pitched note in between the short syllables. Due to the shortening of silence intervals, the song motif in the 2nd generation pupil's song (underlined in green) is short and stable like WT motifs.

Summary of subjective inspection

Overall, the biases that surfaced in the first generation pupils were continued by successive pupils. Long ISO syllables were shortened further, spectral complexity and stability increased further and differentiation into notes continued over multiple learning generations.

Next, we present a numerical summary of the imitation accuracy and syllable identities in tutoring lineages.

3.2 Imitation statistics

Tutor	19	1248	1326	1374	1535
Pupil	1248	1326	1374	1535	1621
Tutor: complex	1	1	1	1	1
Pupil: complex	1	1	1	1	1
% copied	100	100	100	100	100
% invented	0	0	0	0	0
Accuracy	81	86	93	81	73
Tutor: call-syll	0	1	1	0	1
Pupil: call-syll	1	1	0	1	1
Tutor: rare syll	2	0	0	0	1
#calls copied	3	1	0	0	1
#calls invented	0	0	0	1	0

Tutor	1211	1402	1514	1238	1342	249	1439
Pupil	1402	1514	1606	1342	1571	1439	1558
Tutor: complex	3	3	3	4	3	2	2
Pupil: complex	3	3	2	3	2	2	2
% copied	100	100	25	75	66	100	100
% invented	0	0	0	0	0	0	0
Accuracy	88	88	92	81	66	90	84
Tutor: call-syll	1	1	2	3	3	3	3
Pupil: call-syll	2	2	1	3	2	3	1
Tutor: rare syll	2	0	0	0	0	0	0
#calls copied	2	2	1	0	2	3	1
#calls invented	0	0	0	0	0	0	0

Table 3.2 | Imitation of isolate tutors over multiple generations of pupils. Columns show the individual statistics for every tutor and his pupil.

Table 3.2 shows that imitation level was generally high, with fairly high accuracy. Innovation was low and complex syllables were imitated in almost all cases. More birds invented call-like syllables than complex syllables. Sometimes the accuracy decreased again in later generations, for example in Tutor 19's lineage, because at first the birds made changes in duration but when that stabilized, spectral modifications were made. This implies that the changes at different timescales are not necessarily occurring simultaneously.

3.3 Progression toward WT continues in multiple learning generations

We tested for additional multigenerational progression towards WT song features using the first Principal Component, as in the previous chapter. Mean PC1 values were compared across groups for spectral features, duration of acoustic state and rhythm. Our groups are now ISO birds, 1^{st} generation pupils, later generation pupils and WT birds.

We first tested if later generation birds progressed toward WT features beyond the level achieved in the first generation pupils. Direct comparisons across first and later generation pupils reached significance only for DAS ($p=0.02$), but multigenerational comparisons suggest further progression towards WT for all song traits.

For spectral frame features, we found that the PC1 of spectral features changes monotonically towards WT over generations. Its mean values for ISO, first generation, later generations, and WT songs were 1.3, 0.3, 0.03, -0.4 respectively. First PC values for later generation songs were significantly different from ISO song ($p<0.005$, t-test, $n=8$ for later generations) but not from WT songs ($p=0.17$). In other words, later generation birds were already within the WT cluster, whereas first generation birds were still significantly different from WT. A close inspection of the result suggests that the lack of significance in direct comparison is because some of the first generation pupils reached WT levels in spectral features, namely, there is a saturation effect.

For DAS, direct comparison reached significance, and indirect assessment gave similar results to that of spectral frame features, namely, the PC1 values also decreased monotonically with generations: 1.1, 0.3, 0.02, -0.3. However, higher generation songs were significantly different ($p<0.01$) from both WT and ISO, indicating that WT approximation was not complete even in the later generations.

For rhythm, results were not significant even in the first generation, and yet, indirect assessment suggest further progression over generation. PC1 values also decreased monotonically with generations: 4.1, 2.2, 1.4, -2, and differences from WT and ISO were marginally significant ($p=0.02$, 0.056 respectively). Statistical tests are further described in Appendix III.

After showing that progression towards WT song features over multiple generations continued in later generation pupils, we attempt to judge qualitatively if similar biases to those we found in first generation pupils can also be seen in the later generation pupils. Again, we investigate these changes at three different timescales of song structure: spectral features, duration of acoustic state and song rhythm.

3.4 Multigenerational transitions in spectral frame features

As noted earlier, three spectral features showed significant difference in distribution across WT and ISO: AM, FM and goodness of pitch. Figure 3.6 shows the probability distributions of these features for all of our first and multigenerational pupils. Red curves represent ISO birds (n=17), blue curves represent WT birds (n=52) and green curves pupils ($n_{Gen1}=13$, $n_{Gen2}=4$, $n_{Gen3}=2$, $n_{Gen4}=1$, $n_{Gen5}=1$). The top panels include WT birds as well for reference. The other panels only include the ISO Tutors (red, n=6) and the respective pupils. These plots demonstrate that in all three spectral features there is a shift towards WT distributions in first generation birds and that this shift is maintained or continued in later generation pupils. For example, for AM (Fig. 3.6a), we see in the top panel that WT birds (blue curves) have a lower peak than ISO birds (red curves). Green curves have a lower peak. Then in later generations this peak remains consistently lower than the tutors' peak, and up until the 5th generation, there is even a slight decrease in the height of the peak. There is a small backwards tendency in AM and FM in the 5th generation, but this should not be taken too seriously as the sample size in the 5th generation is 1. In FM (Fig. 3.6b), similarly to AM, there is a downwards shift in very low values in pupils towards the WT distribution. We mentioned that ISO birds have lower FM typically, so this shift indicated that pupils' songs do not contain so many low FM sounds. In goodness of pitch (Fig. 3.6c), it seems that most or all of the effect takes place in the first generation, but the curves of later generation birds (green curves) tend to be lower and to the left of ISO tutors (red curves).

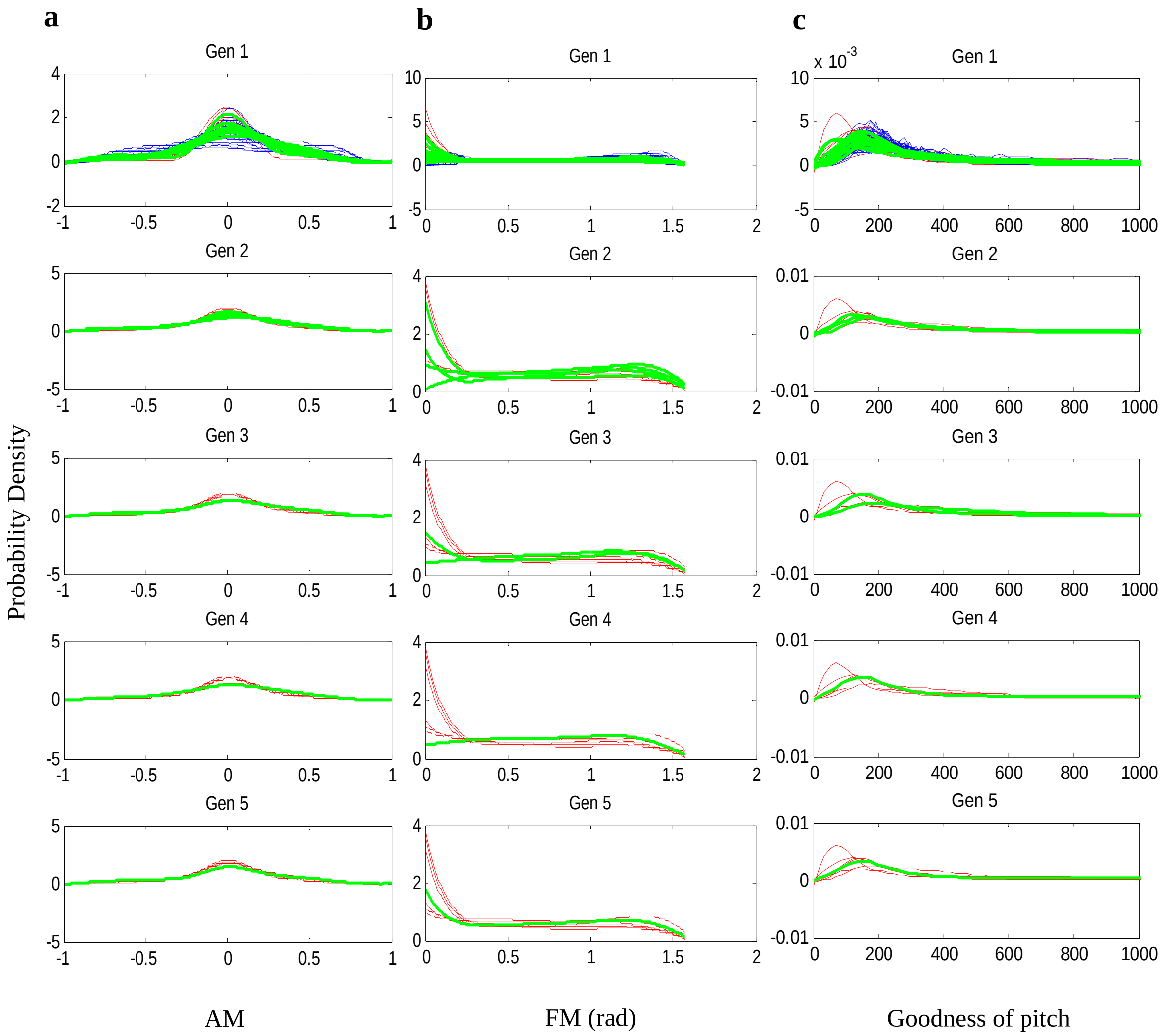

Figure 3.6 | Probability distributions AM (a), FM (b) and goodness of pitch (c) in songs of multiple learning generations. Top panels contain all ISO (red, n=17), WT (blue, n=52) and first generation birds (green, n=13). Lower panels contain ISO Tutors (red, n=6) and later generation birds (green, n_{Gen1}=13, n_{Gen2}=4, n_{Gen3}=2, n_{Gen4}=1, n_{Gen5}=1).

Considering all the spectral features combined, we get a more complete picture of the multigenerational progression towards WT song. From the probability distributions, we calculated cumulative distributions and from those, we performed PCA. PCA plots provide a good visual tool to investigate multigenerational progression. Figure 3.7 shows the first two principal components (PC1 and PC2) with all of our ISO, WT and tutored birds, and the arrows mark the first two generation pupils. As before, arrows point from the tutor to the pupil. The numbers near the origin of the first arrow is the ISO Tutor's name. The second generation pupils in all tutoring lineages continued to shift away form the original ISO tutors in spectral features.

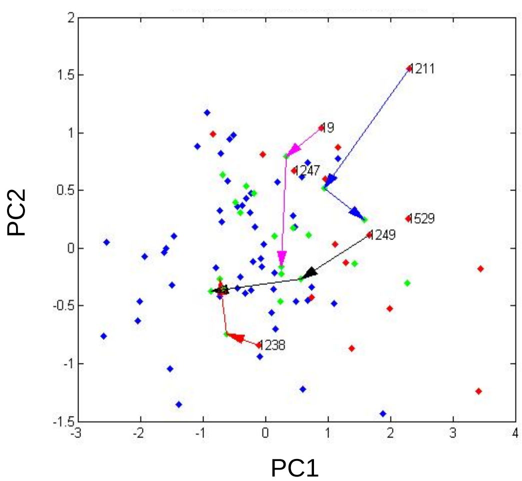

Figure 3.7 | PCA distribution of combined spectral features of first and second generation pupils of isolates. Arrows show that second generation pupil continue the progress away from the ISO tutors.

The representation above is somewhat crowded, which makes it difficult to visualize progression over several generations. To simplify the image and enhance visualization, in addition to individual WT birds, we present the WT density distribution as a cloud. We can now see clearly how the WT distribution is shaped and where it is centered without showing individual birds (Fig 3.8). We can now look at all the multigenerational pupils with reference to WT distribution to see if the continued shift away from ISO tutors also represents a progression towards WT. Fig. 3.8 shows that the progression towards WT indeed continues and does not stop until the pupils' are well within the WT distribution. In the case of our first tutoring lineage (ISO Tutor 19, pink arrows), the progression stops after the second generation, but by this time, the song is nearly in the middle of the WT cloud. Although the arrows still show movement, it seems to hover randomly around the WT distribution center. For the other songs, progression continues for as long as we have data. Based on PC1 mean values, first generation birds are significantly different from both ISO and WT ($p=0.0126$ and 0.018 respectively), but later generation birds only differ significantly from ISO ($p=0.0045$) but not from WT.

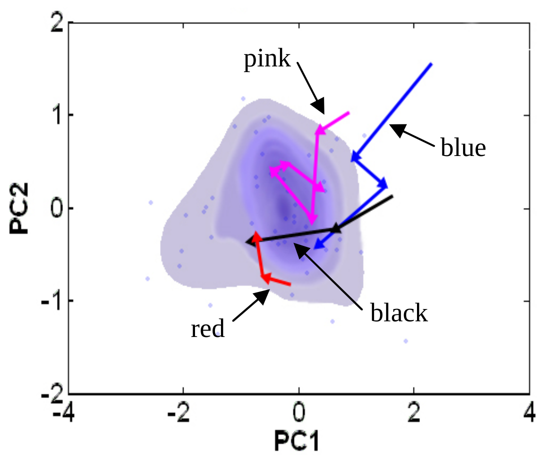

Figure 3.8 | PCA distribution of combined spectral features of multigenerational tutoring lineages. Arrows penetrate deeply in the WT distribution cloud (in grey). Lineage 1: pink arrows, Lineage 2: blue arrows, Lineage 3: red arrows, Lineage 4: black arrows.

3.5 Multigenerational transition towards WT in note-level features

We next investigate alterations in syllable durations over multiple generations of pupils. We will use the durations of acoustic state and note duration ratios to assess multigenerational tendencies.

A. Duration of acoustic state

The sonograms presented in section 3.1 suggested to us that the copies of long ISO syllables were shorter in first generation pupils, and that this trend continued over multiple generation pupils. This suggests that either the pupils compressed (time-warped) stationary syllables only, or shifted acoustic states more rapidly regardless of the copied syllable being stationary or not (namely it would apply to "complex" syllables, too). We now look at the progression of the duration of acoustic states to test two hypotheses:

1. Does DAS decrease over generations toward WT?
2. Is the change in DAS due to selective shortening of longer states?

We start by plotting cumulative histograms averaged across birds in each learning generation (Fig. 3.9b). As a reference, cumulative histograms for all ISO (red curves) and WT (blue curves) are shown in Fig. 3.9a. The gap between the ISO Tutors and the pupils increases in every generation, which indicates that histograms shift in the WT direction and that duration of acoustic state values are lower in later generation pupils.

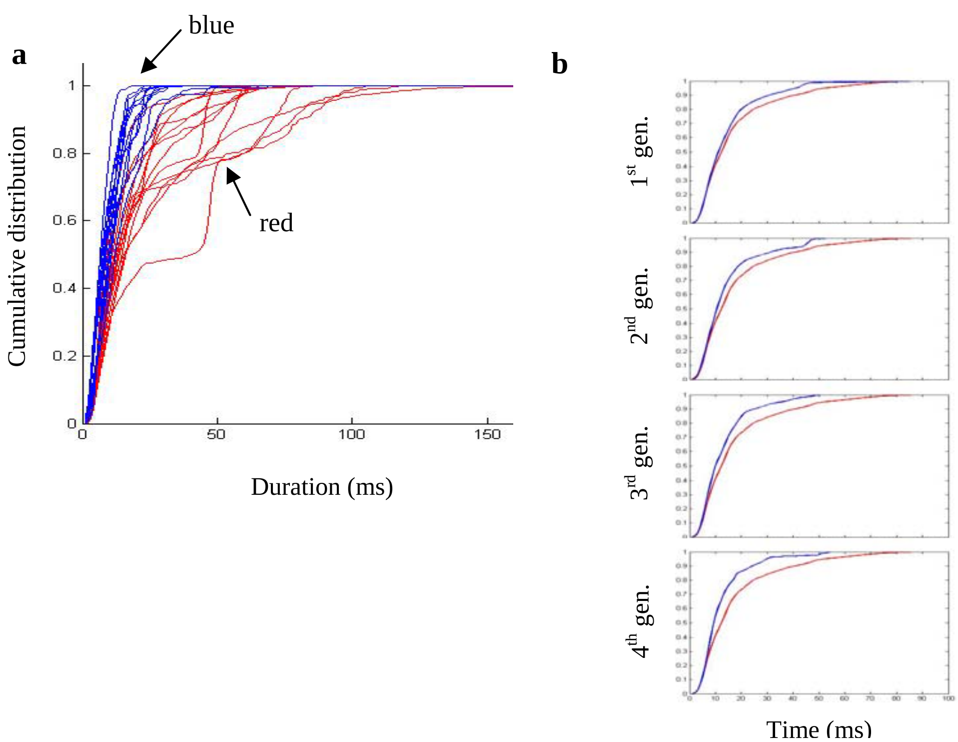

Figure 3.9 | Cumulative histogram of duration of acoustic state in multiple generation birds. a. Cumulative histogram of all ISO (red, n=17) and WT (blue, n=52) birds. **b.** ISO tutors (n=6) are shown in red in every panel (always on bottom), pupils in blue (always on top); panels from top to bottom: first generation pupils (n=6), second generation pupils (n=4), third generation pupils (n=2), fourth generation pupil (n=1).

We see the upward shift in Fig. 3.9, but the magnitude of this event is difficult to judge visually. The Kolmogorov-Smirnov (KS) statistic is an estimate of the difference between any two distributions, and it is sensitive to differences in means, in variances, and in slopes. We calculated the KS-statistic for the cumulative histogram of every bird to estimate their distance from the WT histogram. Figure 3.10 shows these values for every bird in every tutoring lineage (blue lines) and for the mean across lineages (black line). As shown, with the exception of one lineage, where there is further decrease in the KS distance from WT in generation 2, and the distance decreases even more in later generations, where it seems to asymptote. The exception is bird 1558 in Lineage 4. The original ISO Tutor was bird 1249 (Fig. 3.5), who sang very short, fairly simple notes followed by a very long call. Bird 1558 did not imitate the call-like syllable and as a result, his song only contains very short notes. Because WT zebra finch songs tend to contain medium-length harmonic notes, it is likely that this bird falls farther from the WT distribution due to its notes being shorter than the typical WT range.

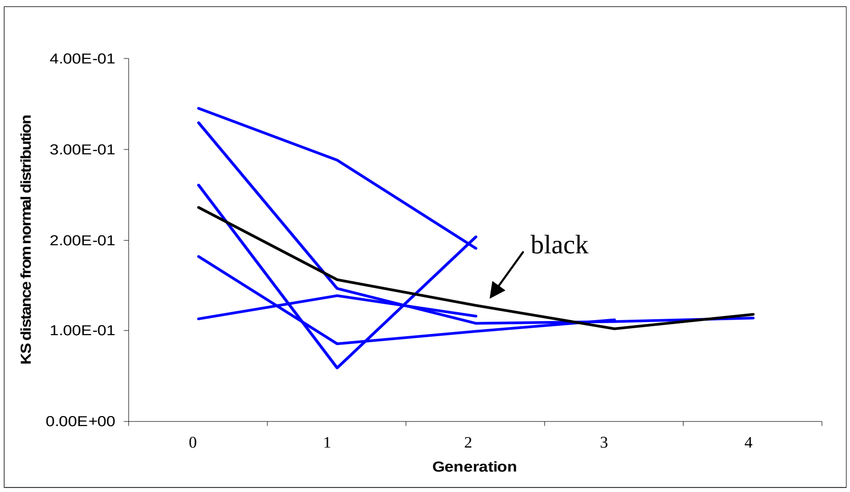

Figure 3.10 | KS statistic for WT and multigenerational pupil cumulative distributions of DAS. Blue lines (pupils in individual tutoring lineages) indicate a decreasing distance between the pupils and WT birds as generation numbers increase. Black line represents mean across lineages.

Finally, we examine the PCA for the duration of acoustic state. Figure 3.11 shows the PC1 and PC2 values for all the WT (blue, n=52) and ISO (red, n=17) birds and pupils (green, n=19) for the 4 multigenerational tutoring lineages. The first two generation pupils are connected by arrows to their tutors. The original ISO Tutor's name is written next to its data point. We can see that for Tutor 19, 1211 and 1249, the progression toward WT is fairly strong as the pupils shift away from the tutors in a general direction of the blue dots. The pupils of 1238 do not seem to be moving in the direction of the WT distribution.

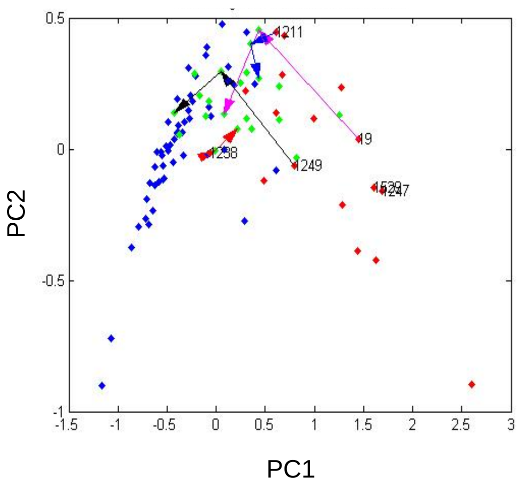

Figure 3.11 | PCA distribution of DAS of first and second generation pupils of isolates. Arrows show that second generation pupil continue the progress away from the ISO tutors.

Figure 3.12 shows all the pupils of all four tutoring lineages using the blurring visualization method discussed earlier. Over multiple generations, we see clear progression towards the center of the WT distribution. In Lineage 1 (ISO Tutor: 19, pink arrows), after generation 2 the trajectory hovers randomly within the WT area, similar to how it happened in the spectral features PCA. Overall, we see that arrows are shorter as they approach the WT center. The only lineage that didn't approximate WT distribution is Lineage 3 (red) where ISO Tutor 1238 sang a complex song with lots of fast transitions, so that the starting point was already close to the WT center. Although the second generation pupil sang more stable syllables, there seems to be less of a variety of acoustic state durations, such as short notes and medium-length notes. We calculated the mean values for PC1 in ISO, first generation, later generation and WT birds. Both first and later generation birds differed significantly from ISO ($p=0.0013$ and $p<0.0001$, respectively) and WT ($p<0.0001$ and $p=0.0089$, for first and later generations birds, respectively), but later generations birds were also significantly different from first generation birds ($p=0.0467$).

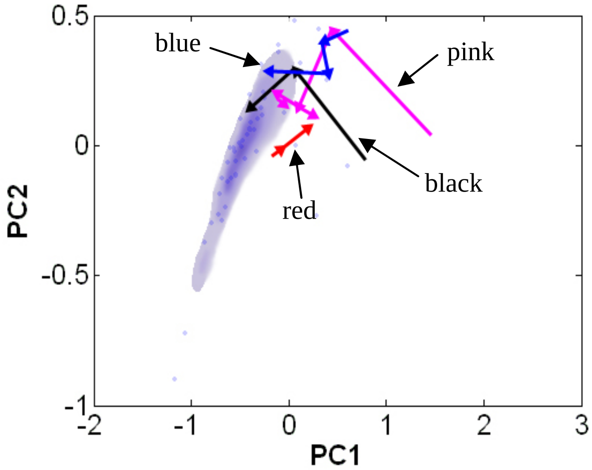

Figure 3.12 | PCA distribution of DAS of multigenerational tutoring lineages. Arrows penetrate in the WT distribution cloud (in grey) in 3 out of 4 cases. Lineage 1: pink arrows, Lineage 2: blue arrows, Lineage 3: red arrows, Lineage 4: black arrows.

We have shown that DAS values do indeed decrease over multiple generations. We now turn to measures of note duration to test our second hypothesis, that is, whether the decreases in DAS were due to the shortening of long notes.

B. Note duration ratios

Our other note-level measure to investigate the multigenerational progression of songs is note duration ratios. We calculated them for each bird and plotted the mean ratios in Figure 3.13. The multigenerational pupils (green) are shown between the ISO (red, n=17) and the WT (blue, n=11) values. In increasing generations, the note duration ratios decrease and become more similar to WT values. This suggests that DAS decreases were due to the shortening of long notes.

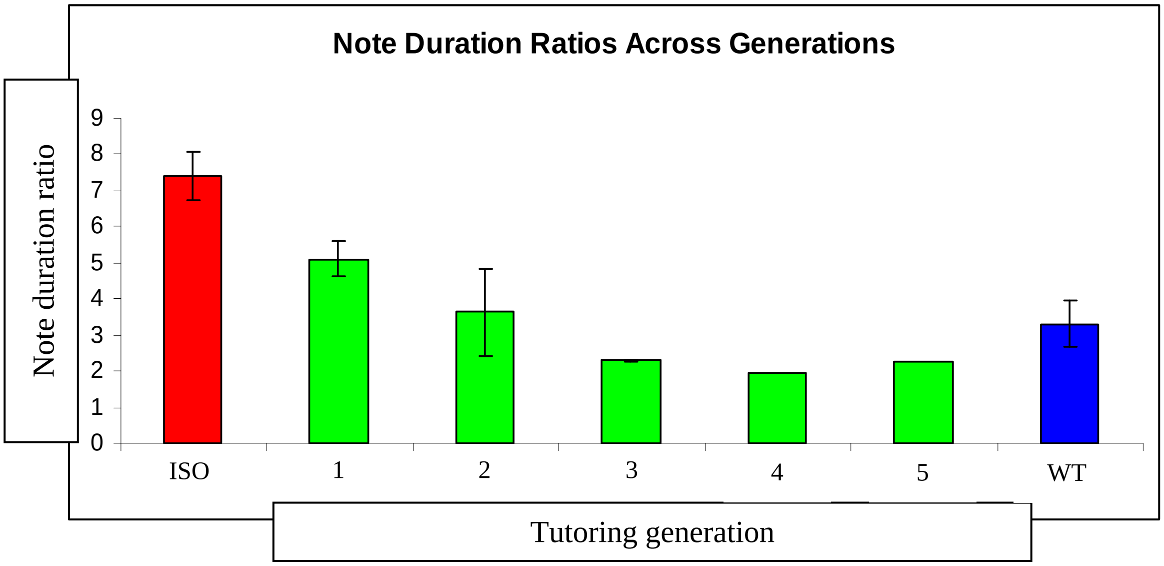

Figure 3.13 | Note duration ratios in multiple generations. ISO birds (red, n=17) and WT birds (blue, n=11) flank the pupils (green bars, n_{Gen1}=13, n_{Gen2}=4, n_{Gen3}=2, n_{Gen4}=1, n_{Gen5}=1). Error bars indicate SEM. In the 4th and 5th generation no error bars are shown, because the sample size is 1.

3.6 Transitions in song rhythm over multiple generations

In Chapter 2, we saw that the song rhythm of pupils of isolates showed a tendency to be more WT-like, but there were some exceptions. Some of these exceptions were likely due to a shortcoming in the method rather than the pupil rhythms becoming less structured. Looking at the multigenerational trajectories (Fig. 3.14), we see similar inconsistencies. The rhythms of multigenerational pupils in three out of the four lineages (Lineage 1, pink; Lineage 2, blue; Lineage 3, red) shift back towards WT, and only one lineage (Lineage 4, black) progresses robustly towards the WT distribution, although it doesn't penetrate by the second generation. PC1 means did not differ from ISO for either first or later generation birds, but both were significantly different from WT (p=0.0033 and 0.017, respectively).

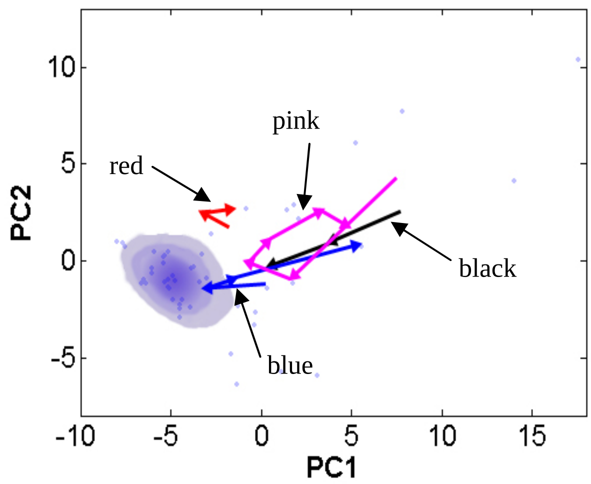

Figure 3.14 | PCA distribution of rhythm spectra in multigenerational tutoring lineages. Arrows penetrate in the WT distribution cloud (in grey) in 3 out of 4 cases. Lineage 1: pink arrows, Lineage 2: blue arrows, Lineage 3: red arrows, Lineage 4: black arrows.

3.7 Conclusions

In this chapter we showed that the progression towards WT features continues over multiple tutoring generations in a recursive learning paradigm. The strategies that the first generation birds employed, namely, imitating WT-like and ISO-like syllables both, but modifying the latter during imitation, were also used by multigenerational birds. The changes were to a lesser degree, but this was due to the fact that first generation pupils sometimes reached WT range feature distributions. In cases where this did not happen, pupils of later learning generations modified more. Overall, within 3-4 generations songs became WT-like at three timescales of song structure: spectral features, acoustic state duration and note length, and song rhythm. We found the weakest effect in song rhythm, and the greatest effect in spectral features.