

Selection, domestication, and the emergence of learned communication systems

Graham Ritchie^{*†}

^{*}Centre for Intelligent Systems and their Applications
University of Edinburgh
graham.ritchie@ed.ac.uk

Simon Kirby[†]

[†]Language Evolution and Computation Research Unit
University of Edinburgh
simon@ling.ed.ac.uk

Abstract

One of the most distinctive characteristics of human language is the extent to which it relies on learned vocal signals. Communication systems are ubiquitous in the natural world but vocal learning is a comparatively rare evolutionary development (Jarvis, 2004). In this paper we take one example of this phenomena, bird song, which displays some remarkable parallels with human language (Doupe and Kuhl, 1999), and we focus on one particular case study, that of the Bengalese finch (*Lonchura striata* var. *domestica*), a domesticated species whose song behaviour differs strikingly from its feral ancestor in that it has complex syntax and is heavily influenced by early learning (Okanoya, 2002). We present a computational model of the evolutionary history of the Bengalese finch which demonstrates how an increase in song complexity and increased influence from early learning could evolve spontaneously as a result of domestication. We argue that this may provide an insight into how increased reliance on vocal learning could evolve in other communication systems, including human language.

1 Introduction

The human capacity for language is one of our most distinctive characteristics. While communication systems abound in the natural world, human language distinguishes itself in terms of its communicative power, flexibility and complexity. One of the most unusual features of human language, when compared to the communication systems of other species, is the degree to which it involves learning. Just how much of language is innate and how much is learned is an ongoing controversy, but it is undeniable that the specific details of any particular language must be learned anew every generation. We do, of course, bring a great deal of innate resources to bear on our language learning process, and the results these innate biases have on the development of languages may explain a great deal about the structure of the languages we see today. But still every child in every new generation must go through a lengthy process of language acquisition if they are to become normal language users.

Once in place, this inter-generational process of language acquisition and use, or *iterated learning* (Kirby and Hurford, 2002), can give rise to cultural evolution, which studies have shown may explain many prominent phenomena of human language, including the emergence of dialects and, by extension,

separate languages (Livingstone, 2002), regular and irregular word forms (Kirby, 2001) and compositional syntax (e.g. Brighton, 2002).

The emergence of learning can therefore be seen as a major transition in the evolution of language and we would like to better understand the evolutionary pressures and factors which caused this transition. A natural point at which to start such an investigation is to look at the communication systems of other animals to see if there are any parallels which might illuminate the relevant ecological factors. Much comparative research has been carried out with the non-human primates, but despite some fascinating results, it seems that their natural communication systems are very different to language, including the fact that learning plays a much less prominent role. In fact it appears that vocal learning systems have evolved in only three groups of mammals: humans, bats and cetaceans, and three groups of birds: songbirds, hummingbirds, and parrots (Jarvis, 2004).

In this paper we concentrate on bird song as it has many striking parallels with language, particularly the way in which it is learned, as Darwin noted in *The Descent of Man*:

The sounds uttered by birds offer in several respects the nearest analogy to language, for all the members of the same species ut-

ter the same instinctive cries expressive of their emotions; and all the kinds that sing, exert their power instinctively; but the actual song, and even the call notes, are learnt from their parents or foster-parents. (Darwin, 1879, p. 108)

Since Darwin's day much research has been carried out into bird song and, to take Tinbergen's four perspectives of ethology, we now know a great deal about its mechanism, development, function and evolution. However, despite much research, in general the evolutionary function of song learning remains unclear (Slater, 2003). The parallels between bird song and human language have also been further elaborated as modern techniques have allowed us to establish the neural mechanisms of both song and language (Doupe and Kuhl, 1999).

2 A case study

Recent studies by Kazuo Okanoya of a domesticated species of finch, the Bengalese finch (*Lonchura striata* var. *domestica*), and its feral ancestor, the white-backed munia (*Lonchura striata*), provide an interesting case study of the interaction of learning and evolution in bird song. The Bengalese finch sings a song with complex finite state syntax which is heavily influenced by early auditory experience. Surprisingly, the munia sings a strikingly simpler, more linear song which is less influenced by early learning. In other words, in a relatively short period of domestication, there have been radical changes in song behaviour. This has happened even though the domesticated species has been artificially bred for plumage rather than song.

Okanoya (2004) has identified the neural mechanism underlying this difference in behaviour and has shown that while Bengalese chicks are able to learn the songs of munia tutors, munia chicks are not able to learn the more complex Bengalese song, clearly demonstrating that there is a physiological basis for this difference.

2.1 Okanoya's hypothesis

As experiments have shown that both female muniyas and female Bengalese finches prefer the more complex song, Okanoya (2002) argues that it is sexual selection which drove this increase in complexity. He argues that domestication freed the Bengalese finch from the pressure of predation and other pressures associated with life in the wild which had previously held song complexity in check. According to

Okanoya, the more complex song of the Bengalese finch may therefore be seen as an honest signal of fitness (Zahavi, 1975); a fitter bird can afford a more complex song. Sasahara and Ikegami (2004) show with a computational model of the finch data that song complexity could indeed increase as a result of sexual selection.

2.2 Deacon's hypothesis

Reviewing the same data, Deacon (p.c.) agrees that domestication masked the natural selection pressure keeping the munia's song simple, but argues that the increase in complexity happened *without* direct selection on the trait. Essentially, he posits that domestication shielded the trait from selection which allowed random genetic drift to erode innate song biases in the munia. This allows previously minor influences, such as mnemonic biases and early auditory experience, to have more of an effect on song structure and learning, which results in the various neural modules involved in song production and learning becoming increasingly *de-differentiated*. Deacon goes on to argue that this process of masking and subsequent de-differentiation is a potential explanation for the evolution of complex functional synergies such as the neural mechanisms for song production now present in the Bengalese finch, and, he argues, in the human capacity for language. The concept of selective masking and its effect on the evolution of language are explored in more detail in (Deacon, 2003).

3 A computational model

In order to evaluate Deacon's hypothesis and to try to establish if such behaviour could evolve spontaneously as a result of domestication, we have developed a computational model of the finch data. The model is designed to be reasonably biologically plausible, and also general enough that it could be extended to other species. The model works with an evolving population of agents, or birds, and the main stages in the simulation are listed here, details of each stage are given below:

Birth The bird's song filter is built up from its genotype as described in section 3.1.

Development The bird is exposed to e songs from its environment, and, using its filter, selects t songs from which it will learn (its training set) as described in section 3.2. The bird then uses the learning algorithm described in section 3.3 to learn the song grammar it will use to sing throughout its life.

Adulthood The bird is tested in f fitness trials, as described in section 3.4 to see how many times, using its filter, it can correctly recognise a bird of its own species and how many times it is correctly recognised by a bird of its own species. These values are added to give a bird’s fitness score.

Reproduction Parents of the same species are selected probabilistically according to their fitness score and their chromosomes are crossed over using one-point crossover with probability pCO (set to 0.7 for all results provided here), to give a new child. Individual genes are mutated with probability $pMut$ (set to 0.05 for all results provided here). The mutation operator used is the ‘Reflect’ operator described in (Bullock, 1999).

Death Each bird in the population is sampled s times and the resulting songs are stored for the next generation to learn from. All of the current birds in the population are removed and their children become the new population.

3.1 The song filter

A bird is modelled as having a genetically coded note¹ transition matrix, which specifies a transition probability from each note to every other note in the used in the simulation, including a probability for the first and final notes. The total number of notes is a parameter of the simulation, $numNotes$, but in all results provided here this was set to 8, i.e. the notes from a alphabetically through to h , this value was chosen as it appears to be the number of unique notes identifiable in both the Bengalese finch and munia’s songs (Okanoya, 2002, p. 56). The matrix is coded for by a chromosome which has one real valued locus for each entry in the matrix which can vary between 0 and 1. This chromosome will thus have $(numNotes + 1)^2$ loci, the 1 is added to include the transitions at the beginning and end of the song. To construct a matrix from the chromosome we look at each $numNotes + 1$ loci of the chromosome in turn, and normalise the values to give a probability distribution for each row of the matrix. An example matrix, and the chromosome that codes for it is shown in table 3.1. Note that this scheme allows different genotypes to code for the same phenotype.

The transition matrix serves one main purpose; to establish the probability that a given song is one of the

¹It should be noted that while we use the term ‘note’ throughout this paper, this is not intended to refer to a particular acoustic note, rather we simply use it to denote an atomic song element that can be reliably differentiated from other elements which appear in the song.

	a	b	c	E
S	0.08	0.15	0.62	0.15
a	0.11	0.89	0.00	0.00
b	0.05	0.10	0.40	0.45
c	0.82	0.09	0.00	0.09

0.1	0.2	0.8	0.2	0.1	0.8	0.0	0.0	0.1	0.2	0.8	0.9	0.9	0.1	0.0	0.1
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Table 1: An example note transition matrix and the chromosome that codes for it. The S indicates the start of the song, and the E indicates the end of the song.

bird’s own species song. This is done by establishing the average probability of each note transition in the song, as shown in equation 1 which defines the preference a given matrix m_x has for a particular song s_y , in this equation n is the number of note transitions in s_y and $m_x(t_i)$ is the entry in m_x for the i th transition of s_y . For example the preference value the matrix in table 3.1 gives for the song cab , which has the transitions $S-c$, $c-a$, $a-b$ and $b-E$, is $\frac{0.62+0.82+0.89+0.45}{4} = 0.695$, while the preference for the song acb is 0.043. Note that we always include the transition to the first note and from the last note, so the empty song ‘’ has a single transition $S-E$, for which this matrix has a preference value of 0.15.

$$preference(m_x, s_y) = \frac{\sum_{i=0}^n m_x(t_i)}{n} \quad (1)$$

The matrix can be thought of as a song ‘filter’. A song with a high probability will be more likely to pass through the filter than one with a lower probability, in our example cab would be much more likely to pass through the filter than acb . If the matrix has a single high probability transition for each note this can be thought of as a strong filter, as it will only accept songs which contain these transitions. If the matrix has even probabilities for each transition it is considered a weak filter as it accepts all songs equally.

We can measure the strength of the filter explicitly by calculating the entropy for each transition distribution (i.e. each row in the matrix), using Shannon’s (1948) measure. This will result in a value which ranges from 0 to $\log(nValues)$, where $nValues$ is the number of probabilities in row r_x (i.e. the number of columns in the matrix). We then normalise this value into the range 0 to 1, as shown in equation 2, which defines the normalised entropy for a given row r_x , in this equation p_i is the probability of the i th transition in r_x . The overall strength of a matrix m_x is then calculated as the average entropy

of each row r in the matrix, as shown in equation 3. A filter strength of 0 means that the filter will only accept one song while an strength of 1 means that the filter will accept all songs equally. As an example, the matrix in table 3.1 has a strength value of 0.56.

$$entropy(r_x) = \frac{-\sum_{i=0}^{nValues} p_i \log(p_i)}{\log(nValues)} \quad (2)$$

$$strength(m_x) = \frac{\sum_{i=0}^{nRows} entropy(r_i)}{nRows} \quad (3)$$

This filter is intended to model the preferences many songbirds have for their species specific song (Catchpole and Slater, 1995). In the model a bird uses its filter for two purposes:

1. To select its training set (the songs it will later use to learn from) from the songs it is exposed to during infancy.
2. To judge whether another bird is a member of the same species for mating or territorial defense.

In this respect, this model is similar to those used in Lachlan’s models of the ‘cultural trap’ in bird song (Lachlan and Slater, 1999; Lachlan and Feldman, 2003). This seems a reasonably plausible assumption, as it is known that some songbirds do have an innate preference for conspecific song both when learning songs as a nestling and also for later mate selection (Catchpole and Slater, 1995).

3.2 Selecting the training set

The infant bird is exposed to e environmental songs to select its t training songs from, both e and t are parameters of the simulation, but were set to 50 and 5 respectively for all results provided here. 5 seems a rather low value of t , but the learning algorithm is very computationally intensive and so a low value is used to speed up the simulation. The e environmental songs are randomly selected from the songs sampled from the previous generation, to compose this set each bird is sampled s times, another parameter which is set to 5 here, so for a population size *popSize* of 100, as used here, this will contain 500 songs.

The infant bird is exposed to each of the e songs in turn and uses its filter to compute the probability it will be accepted. During experimental runs it was determined that checking that the song is accepted once

did not impose enough of a pressure for the bird to correctly select conspecific song and so a song is only added if it is accepted by the filter twice successively. If the bird has not picked t songs after being exposed to all e songs, the process is repeated until t songs have been selected. The training songs are then fed into the learning algorithm described below.

3.3 Song learning

Song learning is modelled as minimum description length (MDL) induction of a probabilistic finite state machine (PFSM), closely following the algorithm described in (Teal and Taylor, 2000). Induction of finite-state machines was chosen to model learning as Okanoya (2002) argues that the songs of both munias and Bengalese finches can be usefully described by a finite-state syntax. The algorithm works by firstly establishing the maximal PFSM that explicitly represents each song in the training set, the prefix tree. The algorithm then searches for nodes which can be merged which will reduce the MDL of the overall machine, whilst also ensuring that the PFSM remains deterministic. The MDL measure takes into account the amount of information (measured by the number of bits) required to code for the machine itself, and also to code for each of the training songs in terms of the machine. Essentially the algorithm searches for the most parsimonious machine in terms of the data. This approach allows us the bird to generalise from its training set, whilst also always being able to reproduce each of the songs it learned from. The reader is referred to (Teal and Taylor, 2000) for a more detailed description of the algorithm used. The only difference between Teal and Taylor’s and our approach is that we also take into account the probability of each note transition, given the probabilities of each transition in the training set.

3.4 Calculating a bird’s fitness

To establish a bird’s fitness we want to check both that its filter allows it to correctly identify its own species, and that its song is correctly identified by other birds of its species. This seems a reasonable model of the pressures acting on song in the wild (Catchpole and Slater, 1995).

To calculate an individual bird b_1 ’s fitness we perform f fitness trials, a parameter set to 250 for the results provided here. In each fitness trail we get the b_1 to produce a song and we then randomly select another member of the population, b_2 and check that b_2 correctly recognises the song using its filter. We also get b_2 to produce a song and check that b_1 cor-

rectly recognises the song with its filter. Every correct recognition means that b_1 's fitness is incremented by 1. With f set to 250, this means that the maximum fitness achievable is 500, or generally $2f$.

3.5 Modelling the finch data

This is a fairly general model of bird song, and so we need set it up to match the data available on the Bengalese finch and the munia as closely as possible. The simulation passes through 3 main phases, each of which runs for 500 generations. The phases are described below.

Phase 1 We know that the white-backed munia has a very stereotyped song and that it seems to only be able to learn songs that match its species-specific song fairly closely (a munia cross-fostered with Bengalese parents is not able to learn its tutor's song). In our model this corresponds to the munia having a strong filter. To simulate this state we seed the environmental songs with a single song type, e.g. *abcdef*. We then run the simulation for 500 generations using the fitness function and learning algorithm described above. As the environment songs are entirely identical the songs that any bird will learn from are always the same, and so they will always induce the same PFISM. This is not meant to be biologically plausible, we simply want the population to develop strong filters for a particular simple song type.

Phase 2 At the end of phase 1 we have a population of birds who sing a stereotypical song and produce offspring with a strong genetic bias to learn that song. To test if the filter can indeed help young birds recognise the appropriate song to learn from in the second phase of the run we start introducing random songs into the bird's environment, this is intended to model hetero-specific song in the environment. We model this by replacing 10% of the s sampled songs with randomly generated songs which use the same notes as the current population and which are constrained to within the same length as the munia songs.

Phase 3 We model domestication of the population simply by ceasing to calculate fitness, but we continue to perform the crossover and mutation operations. This seems a reasonable model of domestication, as in captivity the birds no longer have to recognise their own species to successfully mate or defend their territory as the mating is now controlled by humans and they are kept in aviaries. Domestication can thus be seen to *mask* the selection pressure on these functions. We continue to introduce 10% of random songs into the environment each generation, as it seems a reasonable assumption that the birds will

still be exposed to hetero-specific song, or at least other extraneous sounds, in captivity. Experimental results of this setup are described in the next section.

4 Results

The graph in figure 1 shows several measures taken over the course of each of the three phases described above.

The first measure, in red, is how much variation there is, on average, in an individual's repertoire. This is calculated simply as the total edit (or Levenshtein) distance between a number of an individual's songs (set to 10 for all runs shown here). The second measure, in green, is the average filter strength of the population, calculated as described in equation 3. The third measure, in blue, is the average fitness of the population. We do not calculate this measure in phase 3 and so it does not appear for this phase. The fourth measure we include, in blue, is the average grammar encoding length (GEL) of the population's PFISMs, this is a measure of the size, in bits, it would take to encode a PFISM using the measure defined in (Teal and Taylor, 2000). The fifth measure, in brown, is the average song linearity of the songs tested in measure 1, defined as the number of unique notes in each song divided by the number of unique note to note transitions. The sixth and final measure, in yellow, is the average linearity of the population's PFISMs calculated simply as the number of states divided by the number of transitions. A completely linear PFISM would thus have a linearity of 1, while a maximally non-linear PFISM would have a linearity equal to 1 over the number of transitions in the PFISM.

Two example PFISMs taken from the population at the end of phase 2 are shown in figure 2, and two PFISMs from the end of phase 3 are shown figure 3. The GEL and PFISM linearity values for each machine is also given.

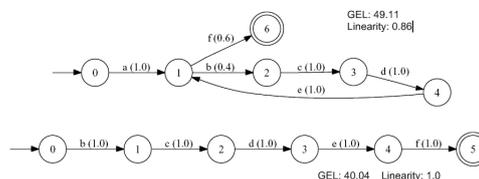


Figure 2: Two example PFISMs from the population at the end of phase 2.

4.1 Analysis

These results demonstrate that the strong filters built up in phase 1, as shown by the increase in filter

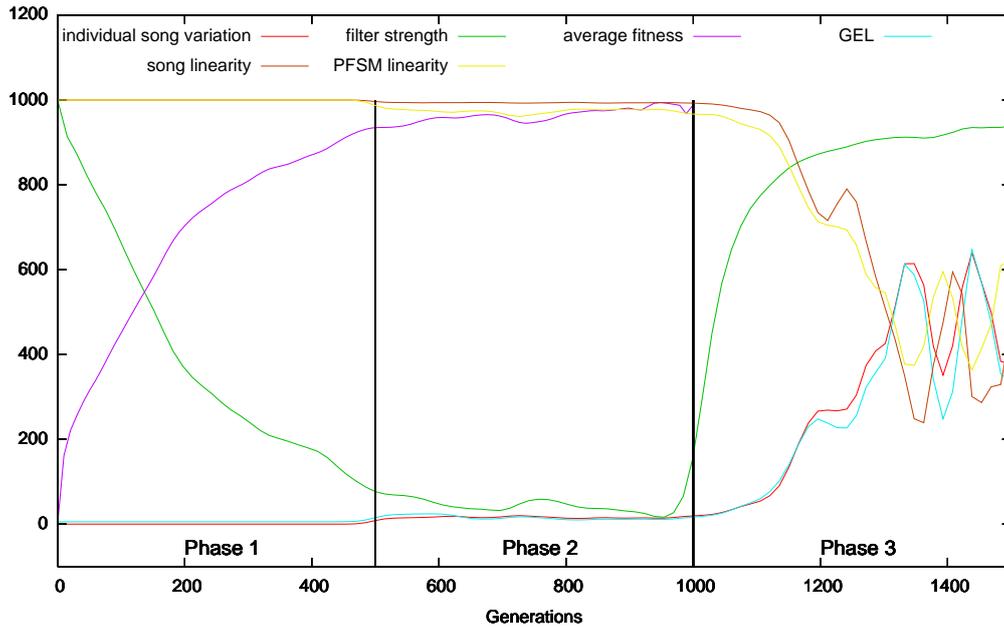


Figure 1: These results are averages taken over 10 separate runs of the simulation with a different random number generator seed for each run. It should be noted that all these measures have been normalised and linearly mapped to give a value within 0 and 1000, where the minimum value seen in the run is set to 0 and maximum values seen in the run is set to 1000. This graph therefore only shows the relative change each of the measures over the course of a run, not the absolute values of each measure. We have also smoothed the lines in the graph to better allow us to see the overall trends. More detailed results are available upon request.

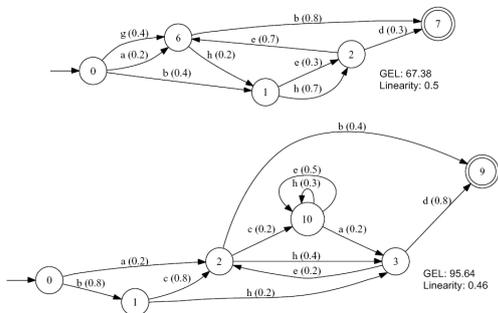


Figure 3: Two example PFSMs from the population at the end of phase 3.

strength², enable the birds to filter out the hetero-specific songs introduced in phase 2 without any fitness decrease. We see that all 5 measures stay roughly the same throughout this phase, indicating that this is a fairly stable state. When we ‘domesticate’ the population in phase 3 we see a significant change in its behaviour. Immediately we see that the filters begin to weaken, and we see that the GEL and

²Note that the strongest filter would give a value of 0, and the weakest 1

the individual variation measures also rise steadily throughout the phase indicating that the population’s PFSMs are getting larger and the birds have a more varied song repertoire. At the same time we see both the song, and underlying PFSM linearity drop, indicating that the songs a bird will sing have comparatively more varied note transitions.

This behaviour seems to be a result of the fact the strength of the population’s filters is no longer being selectively maintained, that it they have been *masked* from selection. This allows mutations to accumulate and for the filters to become steadily weaker. This allows some of the hetero-specific songs to pass through the filter when a bird is selecting its training set, which results in the bird inducing a more varied PFSM. Essentially the domesticated populations is able to learn from much more varied sources and so early auditory experience has much more of an effect on adult song behaviour.

These results are comparable to those provided in (Wiles et al., 2002), who show with a computational model how a similar masking effect, in this case a fruit-rich environment rather than domestication, may have played a role in the loss of the ability of anthropoid primates to synthesize vitamin C. Their

model, however, goes further than ours and shows that if the selection pressure were later *unmasked* this could result in selection for other abilities, e.g. colour vision, that maintained the levels of vitamin C available.

4.2 Song complexity?

Okanoya (2002) argues that the Bengalese finch has a much more “complex” song than the munia. His measure of complexity is the song linearity, defined as the total number of unique song notes divided by the number of unique note-to-note transitions. He finds that the average song linearity of the munia is around 0.8 while the Bengalese finch song has a value of around 0.4 (p. 56). We provide results for this measure over the course of our simulations in the graph above, but on average we also see a higher value, around 0.95, for the ancestral population and a lower value, around 0.6, for the domesticated population.

While this measure seems a reasonably intuitive measure of song complexity, it should be noted that this measure will classify an entirely random song as maximally complex. We do not want to equate randomness with complexity, but we find it hard to define a measure that can differentiate between the two. Any standard measure of the information content of a song will not be able to do so; a random song is maximally informative in information-theoretic terms. However we consider that two measures, the GEL of a bird’s PFSM taken together with the linearity of the PFSM provide a reasonable estimate of the complexity of a song. A PFSM with a very small GEL and a low linearity is likely to produce more random songs, as it approaches a one state PFSM with multiple transitions back to the same state. A PFSM with a large GEL, but a very high linearity (as we see in the ancestral population in the model) will produce an entirely linear song. A PFSM with a large GEL and a relatively low linearity will produce songs that we are more happy to refer to as complex, as the GEL indicates that it has many states, and so different notes will be used in different contexts, but each state also has several transitions which means that different transitions can be made from each context. Our results demonstrate that the domesticated population does have a higher GEL and a lower PFSM linearity than the wild population and so we are tentatively happy to agree that domestication has caused an increase in song complexity. However, we are still working on developing a more satisfactory measure of song complexity.

4.3 Comparison with the biological data

Comparing these results with the data available for the Bengalese finch we find that the model does seem to capture some of the phenomena involved. Okanoya has shown that a munia chick which is not exposed to conspecific song will not sing a normal song, which seems to fit with the model. He has also shown that while Bengalese chicks can readily learn munia songs, munia chicks cannot learn the more complex Bengalese songs. In the model this difference is attributable to their different filters. The difference in the values for the song linearity in the ancestral and domesticated populations also seem to match fairly well.

As it stands though, the model does not explain why the female munia prefers the Bengalese song. We would argue that a bias for complexity song may have been latent in the munia, and the fact that the munia females prefer the more complex song does not prove that this was the driving force for the change in song behaviour, although introducing such a preference into the model may help to tease these pressures apart. Okanoya (2004) demonstrates that the Nif region of the Bengalese finch’s brain is necessary for it to be able to sing the more complex song; when surgically lesioned a Bengalese finch with previously complex song will sing a simpler, more munia-like song. We would argue that the model remains neutral to this datum, as it is possible that the munia does have this pathway present in its brain but, because it only ever learns a simpler song, does not use it.

5 Discussion

Our results demonstrate that an increase in song complexity (in some sense) and increased influence from early learning can arise *without* direct selection on either trait, simply through the process of domestication, but what is the significance of this result for the study of human language? Can studying the evolution of learning and complexity in bird song inform our study of the origins of complex language in our species? We believe that understanding the mechanisms behind the emergence of the Bengalese finch’s song, and indeed the evolution of bird song in general, is valuable for evolutionary linguistics in two ways.

Firstly, it has been argued that iterated learning is a key mechanism for the origins of syntax in human language (Kirby and Hurford, 2002). It is striking that human language differs from most other communication systems both in being transmitted through iterated learning and in having complex syntactic

structure. We say “most” here but not “all”. We appear to be in a very exclusive club with songbirds as another member. Of course, there are important differences between iterated learning in humans and birds. For example, in the former a central constraint on transmitted languages is that they be *expressive*, in that strings must convey complex meanings. Bird song does not carry meaning in the same way, although a diversity of songs may play a role as a sexual display (Catchpole and Slater, 1995). Nevertheless the co-occurrence of iterated learning and signal complexity in both songbirds and humans combined with the rarity of either anywhere else in nature cannot be ignored.

Secondly, and more specifically, by uncovering the crucial role of selective *masking* in the case of the Bengalese Finch, we bring a new mechanism to the table for discussions of the origins of human syntax. It is quite possible that we should not be looking for selective advantages of a culturally transmitted syntactic language, but rather asking what selective forces may have been shielded in our recent evolutionary past. The lifting of selection pressure, and the subsequent diversification of behaviour could have been the necessary precursors of a system of iterated learning for language. What remains to be understood is exactly what more is required for any subsequent modification and synergistic reorganisation of the neural mechanisms underlying these new behaviours.

We feel that the answer to this question is best pursued through computational modelling of the vocal behaviour of both birds and humans.

Acknowledgements

We would like to thank Kazuo Okanoya and Terry Deacon for very useful discussion on this work. We would also like to thank the members of the Language Evolution and Computation Research Unit at the University of Edinburgh for helpful comments and feedback. The first author is supported by an EPSRC studentship.

References

- Henry Brighton. Compositional syntax from cultural transmission. *Artificial Life*, 8(1):25–54, 2002.
- Seth Bullock. Are artificial mutation biases unnatural? In D. Floreano, J.-D. Nicoud, and F. Mondada, editors, *Fifth European Conference on Artificial Life (ECAL99)*, pages 64–73. Springer-Verlag, 1999.
- C. K. Catchpole and P. J. B. Slater. *Bird Song: Biological themes and variations*. Cambridge University Press, 1995.
- Charles Darwin. *The Descent of Man, and Selection in Relation to Sex*. John Murray, London, 2nd edition, 1879. Reprinted in 2004 by Penguin.
- Terrence Deacon. Multilevel selection in a complex adaptive system: the problem of language origins. In Bruce Weber and David Depew, editors, *Evolution and Learning: the Baldwin Effect Reconsidered*, pages 81–106. MIT Press, Cambridge, MA, 2003.
- Allison J. Doupe and Patricia K. Kuhl. Birdsong and human speech: Common themes and mechanisms. *Annual Reviews of Neuroscience*, 22:567–631, 1999.
- Erich D. Jarvis. Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences*, 1016:749–777, 2004.
- Simon Kirby. Spontaneous evolution of linguistic structure: an iterated learning model of the emergence of regularity and irregularity. *IEEE Journal of Evolutionary Computation*, 5(2):102–110, 2001.
- Simon Kirby and James R. Hurford. The emergence of linguistic structure: An overview of the iterated learning model. In A. Cangelosi and D. Parisi, editors, *Simulating the Evolution of Language*. Springer Verlag, London, 2002.
- R. F. Lachlan and M. W. Feldman. Evolution of cultural communication systems: the coevolution of cultural signals and genes encoding learning preferences. *Journal of Evolutionary Biology*, 16: 1084–1095, 2003.
- Robert F. Lachlan and Peter J. B. Slater. The maintenance of vocal learning by gene-culture interaction: the cultural trap hypothesis. *Proceedings of the Royal Society of London. B*, 266:701–706, 1999.
- Daniel Livingstone. The evolution of dialect diversity. In Angelo Cangelosi and Domenico Parisi, editors, *Simulating the Evolution of Language*, chapter 5, pages 99–118. Springer Verlag, London, 2002.
- Kazuo Okanoya. Sexual display as a syntactic vehicle: The evolution of syntax in birdsong and human language through sexual selection. In Alison Wray, editor, *The Transition to Language*, chapter 3. Oxford University Press, Oxford, 2002.

- Kazuo Okanoya. The bengalese finch: A window on the behavioral neurobiology of birdsong syntax. *Annals of the New York Academy of Sciences*, 1016:724–735, 2004.
- Kazutoshi Sasahara and Takashi Ikegami. Song grammars as complex sexual displays. In *Artificial Life 9*, 2004.
- C. E. Shannon. A mathematical theory of communication. *Bell Systems Technical Journal*, 27:379–423 and 623–656, 1948.
- P. J. B. Slater. Fifty years of bird song research: a case study in animal behaviour. *Animal Behaviour*, 65: 633–639, 2003.
- Tracy K. Teal and Charles E. Taylor. Effects of compression on language evolution. *Artificial Life*, 6: 129–143, 2000.
- Janet Wiles, James Watson, Bradley Tonkes, and Terrence Deacon. Strange loops in learning and evolution, 2002. available from: citeseer.ist.psu.edu/wiles02strange.html.
- A. Zahavi. Mate selection – a selection for a handicap. *Journal of Theoretical Biology*, 53:205–214, 1975.