

Coevolution of Birdsong Grammar without Imitation

Kazutoshi Sasahara and Takashi Ikegami

Department of General Systems Studies, Graduate School of Arts and Sciences,
University of Tokyo,
3-8-1, Komaba, Meguro-ku, Tokyo 153-8902, Japan
{sasahara,ikeg}@sacral.c.u-tokyo.ac.jp
<http://sacral.c.u-tokyo.ac.jp>

Abstract. The mating song of the male Bengalese finch can be described by a finite-state grammar and has the feature that more complex songs are preferred by females [1]-[3]. These facts suggest that complex song grammars may have evolved via sexual selection. How, then, do the female birds gauge a song's complexity? Assuming that they can measure the complexity of a song while communicating with a male, but without making a model of the song, we studied the evolution of song grammars. In our simulation, it was demonstrated that song grammars became more complex through communication between coevolving males and females. Furthermore, when singing and listening were subject to fluctuations, peculiar features were observed in communication and evolution.

1 Introduction

From the standpoint of Chomsky's theory, the most unique function of human language is its "recursiveness", which enables potentially infinite expressions from finite elements [4]. This plays a fundamental role in grammar. Since "linguistic behavior does not fossilize", the questions of how mankind attained this function and how language has been complicated in modern times are difficult to deal with scientifically and they remain significant open questions.

In addition to mankind, songbirds and whales have a similar capability. They can combine finite sound elements recursively to have vocal communication [5]. Therefore, their song can be regarded as analogous to human language. The study of such language-like behavior is one way to understand the evolution of language.

We focus on the Bengalese finch (*Lonchura striata var. domestica*), in particular the song of the male as it courts a female. This song consists of a combination of chunks, each of which is a sequence of sound elements. It can be described by a finite-state grammar, unlike alarm calls and threats. In addition, it was shown in Okanoya's experiments that more complex courtship songs were preferred over monotonous ones, and that they promoted the reproductive behavior of females [1][2]. These facts suggest that males with complex songs have been chosen by females and that song grammars have evolved as a result of sexual selection [3].

The handicap principle [6], which says that the complex song of a male may be a genuine indicator of his superiority, supports this scenario.

How do female birds perceive the complexity of a song? One possibility proposed by Okanoya is that she can judge it through communication, without making a model of the male song. Based on the observation that the female Zebra finch chirps (interjects) in synchrony with the male, we supposed that the female Bengalese finch also interjects in this way, measuring how many interjections succeed according to her preference to evaluate the song quality. What is significant in this hypothesis is that without a model of the song in the female brain, she can still judge its quality according to her preference and then complex song grammars may evolve. To explore this, we attempted to model the evolution of song grammar through communication. Furthermore, we introduced the effects of mis-singing and mis-listening and studied how such fluctuations influenced communication and evolution. In this paper, we verify our hypothesis, examine the behaviors observed, and argue that song grammars become more complex as a result of coevolution.

2 Modeling

Here we model the coevolution of males and females as a communication game. The song grammar of a male bird is expressed as a sequential machine[8], $M_i = (Q, \Sigma, \Delta, \delta, \lambda, q_0)$, which is a finite automaton(FA) with output function, where Q is a finite set of states, q_0 is an initial state, Σ is a finite set of input symbols, Δ is a finite set of output symbols, δ is a state transition function, $Q \times \Sigma \rightarrow Q$, λ is an output function, $Q \times \Sigma \rightarrow \Delta$. In this model, $\Delta = \{blank, A, B, C, D, E, F, G, H\}$, each letter represents a song chunk and “blank” represents a silent interval between chunks. The whole output sequence expresses a courtship song. Furthermore, two types of dummy male are introduced as probes for investigating evolutionary features of the system. One is a “random bird”, which has a random number generator instead of a FA and sings a random song. The other is a “monotonous bird”, which has a clock generator that outputs chunks and blanks alternately. He sings a simple periodic song.

Female birds, on the other hand, have another type of FA, $M'_i = (Q, \Sigma, \delta, q_0, F)$, which represents her preference, where Q , Σ , δ , and q_0 are the same as above, F is a set of accepting states, which is a subset of Q . She changes her internal state by listening to the song and interjects if she is in an accepting state. If she is able to interject when the male bird sings *blank*, we say that the interjection is successful. Examples of male and female FAs are shown in Fig.1.

Moreover, the song fluctuation effects, mis-singing and mis-listening, are taken into account. Some chunks are easily mistaken (in both singing and hearing) while others are not. Specifically, in the pair (A,B), “A” tends to be mistaken for “B” once every three times and vice versa. Similarly, the mistake rate is 1/5 for (C,D), 1/10 for (E,F) and 1/20 for (G,H). Hence, the male does not always sing his grammar exactly and the female does not always accept his song exactly, even if it is perfect.

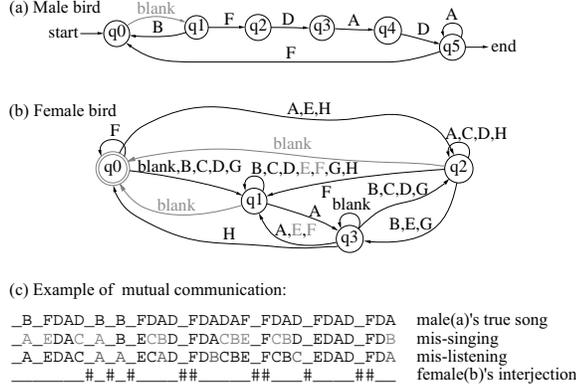


Fig. 1. Examples of finite automata (FA) and communication. (a) Shows a male FA and (b) shows a female. The double circle represents an accepting state. (a) and (b) are a well-suited pair that achieved good communication at $t = 5,000$. (c) Illustrates a communication process including song generation, interjection and fluctuations in singing and listening. “_” denotes “blank” and “#” denotes interjection.

Given n types of males and females, communication is performed as follows. Every time step, each normal male with a FA sings for length L_i^{song} chunks, according to his grammar, M_i . Dummy birds sing according to their own mechanisms. Note that they make mistakes in singing at the rates given above. Subsequently, each female interjects to the songs in accordance with her preference, M'_i . Females also make mistakes in hearing. An example of communication is illustrated in Fig.1.

In communication between the populations of males of the i -th type and females of the j -th type, their scores are calculated as follows.

$$a_{ij} = \frac{N_j^{success}}{N_i^{blank}} - C \cdot N_i^{node} \quad (1)$$

$$b_{ji} = \frac{N_j^{success}}{N_j^{interj}} - C \cdot N_j^{node} \quad (2)$$

where a_{ij} is the score of the i -th male population, b_{ji} is the score of the j -th female population, $N^{success}$ is the number of successful interjections, N^{blank} is the total number of blanks, N^{interj} is the total number of interjections and N^{node} is the number of nodes in M_i , N^{node} is the number of nodes in M'_i . The first terms denote the success rate of interjection and the second terms denote the cost (in nodes) they pay for their ability to sing complex songs. Note that the dummy birds have no node-cost because they have no FA.

In following cases, however, it is considered that the communication ends in failure, and both male and female are given zero score: (i) Every element of the

song is *blank* (inappropriate song), (ii) every reaction of the female is interjection (inappropriate interjection), (iii) the female interjects exactly when the male sings *blank* and never otherwise (no novelty). We consider (iii) so that we can take into account the importance of novelty of songs for females' preferences [7].

Let x_i be the relative population of the i -th male bird, and y_i be that of the i -th female bird. Supposing that they produce offspring relative to their scores as defined previously, the relative populations are calculated by the replicator equations [9], with the discrete-time Runge-Kutta method:

$$\frac{dx_i}{dt} = x_i \left(\sum_{j=1}^n a_{ij} y_j - \sum_{k=1}^n \sum_{l=1}^n x_k a_{kl} y_l \right) \quad (3)$$

$$\frac{dy_i}{dt} = y_i \left(\sum_{j=1}^n b_{ij} x_j - \sum_{k=1}^n \sum_{l=1}^n y_k b_{kl} x_l \right) \quad (4)$$

where, $0 \leq a_{ij} \leq 1$, $0 \leq b_{ij} \leq 1$, and $\sum_i x_i = \sum_i y_i = 1$.

We define the linearity of a FA, LI , as $LI \equiv N^{node}/N^{arrow}$, where N^{arrow} is the number of arrows leaving from a node. If $N^{node} = N$, this value ranges between $1/N \leq LI \leq 1$ as N^{arrow} varies from N^2 to N . More complex FAs have lower values of LI .

The following mutations occur with the same probability R^{mutat} to males and females at each time step:

- (a) **Arrow Mutation:** Change the transitions of the FA with the number of nodes remaining fixed.
- (b) **Node Mutation:** Change the number of nodes (± 1) and then add or remove arrows as required.
- (c) **Length Mutation:** Change $L^{song}(\pm 1)$.
- (d) **Selection1:** Remove the males in proportion to LI or remove a randomly selected females.

Note that sexual selection doesn't involve any relation between complex sexual display and male survival superiority. In selection1, however, we hypothesize that having a complex grammar is correlated to male survival ability. Therefore, males with higher values of LI are regarded as less able to survive. In addition, the following selection is performed:

- (d) **Selection2:** Remove the population with a relative population lower than the threshold, V^{thresh} .

The dummy males only mutate according to (c) and (d). The females are prone to every mutation except (c). After the selections, a normal bird type is randomly chosen and first mutated according to (a) or (b), and then mutated according to (c), before being reintroduced into the system as a new bird type with relative population $1/n$. In this way, the number of populations is kept constant.

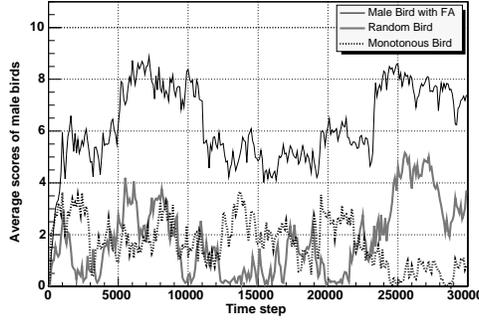


Fig. 2. Average scores of males. The solid line shows the average score of normal males, the thick line shows that of the random bird, and the dotted line shows the monotonous bird.

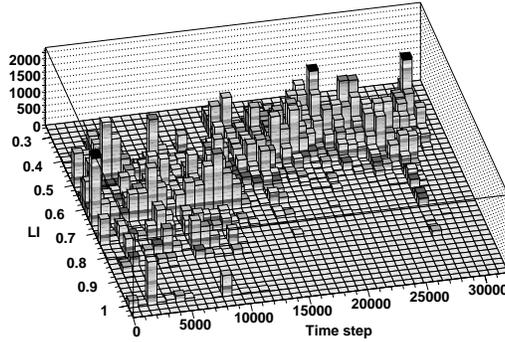


Fig. 3. Time evolution of linearity(LI). We observe that evolution stagnates around $t = 10,000$ and that a rebound phenomenon occurs between $t = 23,000$ and $27,000$.

3 Simulation Results

Our simulations consisted of ten types ($n = 10$) each of both male and female populations, where one of the male populations was comprised of random birds, and another of monotonous birds. The other eight were normal birds with FAs. The initial relative populations were $x_i = y_i = 0.1$. Every FA was constructed randomly with the initial conditions $N^{node} = 2$ and $L^{song} = 5$. Other significant parameters were $C = 0.001$, $R^{mutat} = 0.06$, $V^{thresh} = 0.001$, $\max L^{song}$, $\max N^{node}$ and N^{mode} were 50.

The average scores of the males as a function of time appear in Fig.2. We see from Fig.2 that the average score of the normal males increased rapidly until about $t = 2,000$ and then their score varied between around 4 and 8. Very similar evolution was observed for the females. On the other hand, the dummy birds couldn't maintain high scores. The normal males' scores are much larger than those of dummy birds. This clearly indicates that the females could interject properly to the output sequences from the song grammars, and could react to

behavior between periodic and random. Note that the monotonous birds did obtain non-zero score. Without fluctuations, the monotonous males would suffer complete interjection and could obtain low scores only when they were matched with monotonous females. Mis-singing and mis-listening, however, enabled them to acquire higher (but still low) scores. On the other hand, some proper complex songs might yield lower scores because of fluctuations. Song fluctuation has an effect of obscuring the complexity of messages which expresses "honest signal of male superiority". Therefore, the handicap principle does not work perfectly in a noisy environment.

The time evolution of linearity(LI) appears in Fig.3. From this, it can be seen that the males' grammars were generally having smaller values of LI , i.e. they were becoming more complex over time. This shows clearly that the song grammars evolved to become more complex as a result of communication. Furthermore, the "rebound phenomenon" - that complexity of grammars returned to being relatively simple after being somewhat complex - was observed. In Fig.3, it can be seen between $t = 23,000$ and $27,000$ and it is consistent with the period in which males' scores were rising again in Fig.2. This rebound phenomenon is due to males' grammars becoming so complex that females couldn't successfully interject, so that they began to choose slightly less complex grammars. Thus, we conclude that song grammars cannot become arbitrarily complex and instead they attain complexities within the range that females are able to interject appropriately. In a sense, communication itself seems to be the trigger that causes the complexity of grammars to increase, at the same time determining its maximum. Subsequently, we may observe "stagnated evolution" - that simple males were selected and dominated the system around $t = 5,000$. In this period, communications seemed to be successful, as shown in Fig.2. However, since the females became too complex (by increasing their number of nodes) at this time, after that the simple males suffered perfect interjection and disappeared from the system. Without selection1 (i.e. without considering the relation between complex grammar and survival superiority), we did not observe any clear contrast between the rebound phenomenon and the stagnated evolution, even though the song grammars gradually became complex and saturated at the higher LI comparing with the case of selection1. This suggests that with only sexual selection the complexity cannot develop enough. We hypothesize that the higher complex grammars can only attain at the expense of the lower complex grammars.

Fig.4 shows the change of the average values characterizing the communication: the song length, the number of kinds of chunks and the number of nodes in males and females. It was found that L^{song} had almost reached a maximum after $t = 6,000$. This indicates that females were apt to succeed in interjection with the longest songs, even taking into account the risks of mis-singing, mis-listening and mis-interjection. Moreover, it turned out that the number of nodes of female FAs decreased at around $t = 10,000$ which was the stagnation period, and around $t = 23,000$ when the rebound phenomenon occurred. Females with FAs of only a few nodes are not sensitive to the order of chunks and expresses preferences to the chunks themselves. On the other hand, if the number of nodes

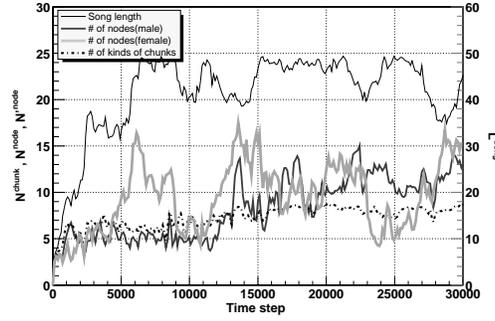


Fig. 4. Dynamics of communication features. Each line shows a value averaged over all birds. The normal line shows the song length (L^{song}), the medium line shows the number of nodes in males $FA(N^{node})$, the thickest line is the number of nodes in females $FA(N^{node})$, and dotted line shows the number of kinds of chunks (N^{chunk}), used in songs.

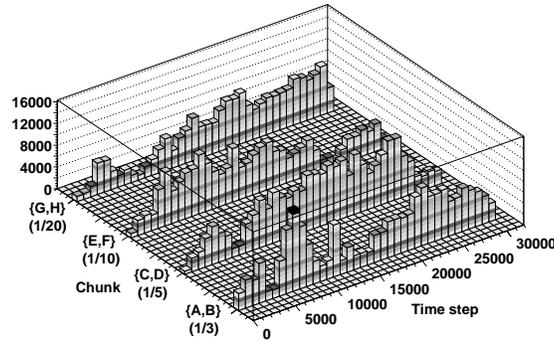


Fig. 5. Trend of songs favored by females. Chunks used in songs oscillate over time.

increases, the females become more complex and sensitive to phrasing. Such a correlation between the number of nodes in females and the linearity of the males they prefer is a reasonable trait.

Finally, Fig.5 shows the number alternations of chunks used in the songs over time, which includes the pairs $\{A,B\}$, $\{C,D\}$, $\{E,F\}$, $\{G,H\}$. Even though these pairs have different mistake rates, we didn't find one to be evolutionarily superior. Since the females did not care about the mistake rates, the popularity of different chunks was free to oscillate over time. Indeed, novelty in songs (due to different chunks changing their popularity) can be regarded as a feature of songs favored by females. At $t=5,000$, $\{E, F\}$ was the most popular chunk and we could find the fluctuation-absorbing structure in a female FA in Fig.1. That is, since E and F were on the same arrow of her FA, the interjection could be successful even if she mistook E for F. Such structures were commonly found in simulations with song fluctuation.

4 Conclusion

In this paper, we did not require the driving force behind grammar evolution to be a diverse outside environment, but only to be communication itself [10]-[14]. We focused on whether song grammars could evolve to become complex through relatively brief, song-interjection communication. Consequently, we could clearly demonstrate that song grammars could evolve to become complex as a result of females' interjection. This supports Okanoya's hypothesis. In addition, it turned out that the song grammars could not become arbitrarily complex, and instead they evolved towards a boundary where interjection is successful, but not perfect. For example, the rebound phenomenon, which showed that overly complex grammars would revert to simpler ones, was observed.(see Fig.2 and Fig.3) That is, the complexity evolved by sexual selection must be understandable by the cognitive system of partner.

Subsequently, we investigated how song fluctuation in singing and listening affected communication and evolution. In such an environment, it was possible that the simple males might be selected and the complex males might not be. The song grammars, however, generally have a tendency to become complex, although not only excellent grammars are inherited in the presence of fluctuation. Moreover, particular fluctuation-absorbing structures were found. In this model, such fluctuation functioned as a mechanism which influenced evolution in a form distinct from mutation.

Acknowledgements: This work was partially supported by Grant-in aid(No. 09640454) from the Ministry of Education, Science, Sports and Culture. KS would like to thank Okanoya for his helpful suggestions.

References

1. Honda, E., and Okanoya, K.(1999), *Zoological Science*, **16**, pp.319-326.
2. Hosino, T., and Okanoya, K.(2000), *NeuroReport*, **11**, pp.2091-2095.
3. Okanoya, K.(2002), *The Transition to Language*, pp.46-63, Oxford University Press.
4. Hauser, M.D., Chomsky, N. and Fitch, W.T.(2002), *Science*, **298**, pp.1569-1579.
5. Aitchison, J.(2000), *The Seeds of Speech: Language Origin and Evolution*, Cambridge University Press
6. Zahavi, A. and Zahavi, A.(1997), *The Handicap Principle*, Princeton University Press.
7. Werner, G.M. and Todd, P.M.(1997), *Fourth European Conference on Artificial Life*, pp.434-443.
8. Hopcroft, J.E. and Ullman, J.D.(1979), *Introduction to Automata Theory, Languages and Computation*, Addison Wesley.
9. Hofbauer, J., Sigmund, K.(1998), *Evolutionary Games and Population Dynamics*, Cambridge University Press.
10. Suzuki, J. and Kaneko, K.(1994), *Physica D*, **75**, pp.328-342.
11. Hashimoto, T. and Ikegami, T.(1996), *Biosystems* **38**, pp.1-14.
12. Steels, L.(1999), *The Talking Heads Experiment. Volume 1. Words and Meanings*, Antwerpen.
13. Nowak, M., Komarova, N.L. and Niyogi, P.(2001), *Science*, **291**, p.114-118.
14. Nowak, M., Komarova, N.L. and Niyogi, P.(2002), *Nature*, **417**, pp.611-617.