# Evolution of Birdsong Syntax by Interjection Communication

**Kazutoshi Sasahara**
Laboratory for Biolinguistics
RIKEN Brain Science Institute (BSI)
2-1 Hirosawa, Wako-shi
Saitama 351-0198
Japan
sasahara@brain.riken.jp

**Takashi Ikegami**
Department of General Systems
    Studies
Graduate School of Arts and Sciences
The University of Tokyo
3-8-1, Komaba, Meguro-ku
Tokyo 153-8902
Japan
ikeg@sacral.c.u-tokyo.ac.jp

**Abstract** Animals use diverse forms of communication, from sound signals to body postures. Recent ethological studies have reported a unique syntactic communication of a songbird, the Bengalese finch (*Lonchura striata* var. *domestica*). Male Bengalese finches sing complex courtship songs, which can be reconstructed by finite automata, and female Bengalese finches prefer complex songs, as opposed to monotonous or random ones. These facts suggest that the song syntaxes of male birds may have evolved as a result of sexual selection by female birds. Inspired by this hypothesis, we developed a communication model that is a system coupling different types of automaton, one for song production by males and another for song evaluation by females. We applied this model to study the evolution of syntactic animal communication in terms of the self-organization of coevolving automata. Three types of courting strategies as well as a relationship between the song syntax and female preference emerged. We argue that despite the simple communication involved, the complexity and diversity of song syntaxes can evolve via diverse female preferences.

**Keywords**
Coevolution, finite automaton, interjection, novelty, sexual selection, song syntax

## 1 Introduction

### 1.1 Evolution of Animal Communication

Animals communicate through a wide variety of media and methods, including body postures and movements, and physical signals such as sounds and chemicals. Honeybees, for example, indicate the location of nectar-bearing flowers to colony mates through a complex series of dances; vervet monkeys emit different alarm calls depending on the perceived predators [5, 17]. The evolution of animal communication is one of the central issues in ecology and ethology. Why and how did such diverse forms of animal communication evolve, and how do they differ from linguistic communication in humans? To address these questions, animal communication is modeled and simulated in artificial life (ALife) research.

Previous ALife research on the evolution of artificial animals has demonstrated that communication could have evolved for social cooperation [10] and for finding mates [22]. Other studies have addressed different aspects of animal communication, such as signal evolution, including complex signals involving chaotic dynamics, which are difficult to imitate [19], and the diverse signals resulting from sexual selection [23], the evolution of categories shared by artificial agents [1], and the emergence of meaningful symbols in artificial agents [18].

Simulations of animal communication are required to understand the evolution of human communication; however, previous simulation models did not use ethological evidence to analyze the

syntactical aspects of animal communication. In this article, we model the coevolution of syntactical animal communication on the basis of recent studies of songbirds. We begin by briefly reviewing a case study on the Bengalese finch (*Lonchura striata* var. *domestica*) and describe the objectives of our study.

## 1.2 Song Syntax as a Complex Sexual Display

There are thousands of songbird species, each producing a repertoire of songs and other calls to communicate with conspecifics. Recent studies of the Bengalese finch have reported unique features of its courtship songs with respect to mechanism, function, development, and evolution:

- The song of the male Bengalese finch is neither monotonous nor random; it consists of a combination of chunks, each of which is a fixed sequence of a few song notes. Unlike alarm calls and threats, the song can be reconstructed by a finite automaton, which we call song syntax (see Figure 1) [6, 12, 13]. In contrast, most other songbirds sing stereotyped songs or nearly random songs without song syntax [2].

- Song syntax is managed by song circuitry in the brain. The circuitry hierarchy (viz., the song control nuclei NIf, HVC, and RA) directly corresponds to the song hierarchy. Namely, the NIf controls the transition of song syntax, the HVC organizes the chunks, and the RA is responsible for song notes [8]. Furthermore, there are anatomical differences in the song circuitry of males and females; for example, females lack the HVC [21].

- Females prefer more complex songs over monotonous ones, displaying more reproductive behaviors, such as increasing the frequency of nest-making and mating poses, in the presence of the former [12, 13].

- The Bengalese finch is a domesticated strain of the wild white-backed munia (*Lonchura striata*). After 250 years of domestication, the songs of the male Bengalese finch have become much more complex than those of its ancestor (see Figure 2) [12, 13].

These findings suggest that the complex song syntax of the male Bengalese finch has evolved as a result of female preference for complexity. Although this hypothesis is plausible, some questions remain: How do females detect the song complexity? What structural features of song syntax evolve
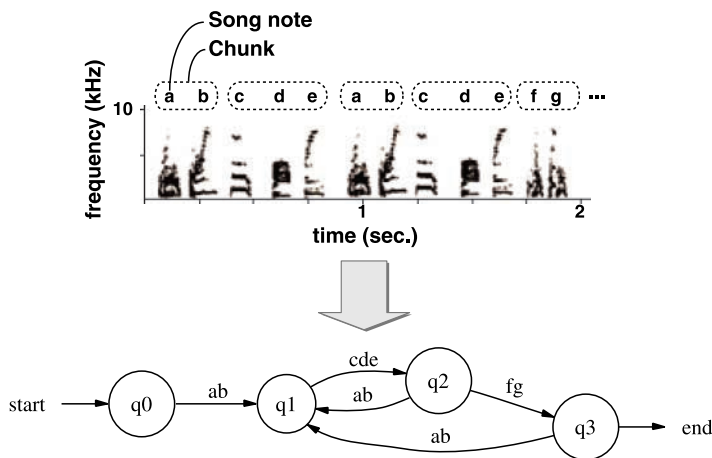


Figure 1. An example of courtship song and song syntax of Bengalese finch. (Data supplied by K. Okanoya; modified with permission.)
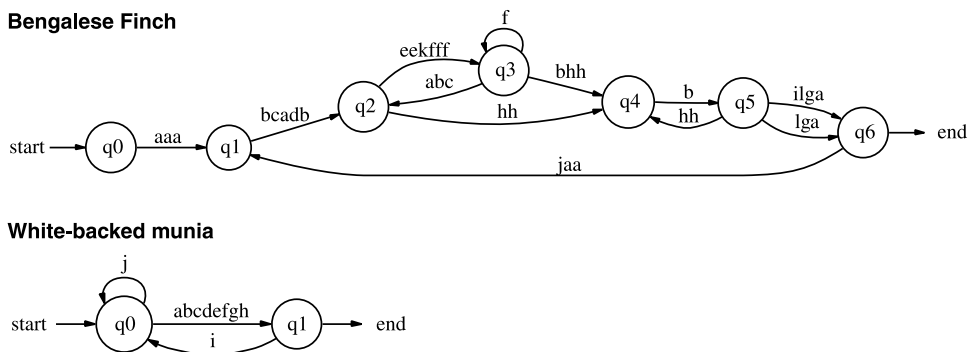
**Bengalese Finch**



**White-backed munia**



Figure 2. Examples of the song syntax of the Bengalese finch (top) and white-backed munia, which is the ancestral species (bottom). The song syntax of the Bengalese finch is much more complex than that of its ancestor. (Data supplied by K. Okanoya; modified with permission.)

via sexual selection? To address these questions, we simulated the coevolution of male song syntax and female preference using a simple ALife model, based on available experimental data.

## 2 Model

### 2.1 Key Concepts

To model song communication, it is necessary to consider two concepts.

First, how do female birds perceive *song complexity*? Since females tend to prefer complex songs, if they can have an absolute measure of song complexity (for example, an ability to measure the song's randomness or entropy), the evolution of song syntaxes might just be a maximization process for this measure. However, we find diverse song syntaxes in male birds, from complex ones with many stochastic transitions to simple ones with linear transitions. Thus, that is not the case for song communication. An alternative possibility is an interaction-based indirect measure of song complexity. When listening to monotonous or random songs, females generally show a lower frequency of reproductive behaviors, while songs generated by an artificial finite automaton (FA) enhance the reproductive behavior of females [13]. Interestingly, however, a few females still prefer monotonous or random songs.

These facts suggest that females have special preferences for the arrangement of song chunks depending on an evaluation of communicational compatibility with males [15, 16]. As a possible form of this interaction, we introduce *interjection communication*, in which a female interjects (by wagging her tail or chirping softly, for example) in synchrony with a male's song, measuring how many interjections succeed with her own preference before she evaluates her satisfaction with the song.[1]

The second concept concerns *song novelty*. Since courtship songs are sexual displays to attract females, novelty plays an important role in song communication [25]. Several analytic and simulation studies support this assertion [9, 11, 23]. Thus, we introduce song novelty into the interjection communication in terms of the unpredictability of song chunks. The song novelty condition is defined as follows:

- A female bird must make more than $N_{novel}$ mistakes while interjecting.

---

[1] In fact, a female zebra finch (*Taeniopygia guttata castanotis*) has been observed to interject by chirping in blanks in a male zebra finch's song [20]. On the other hand, a female Bengalese finch never chirped, but showed rhythmic body movement. We assume that this body movement can function as an interjection for a female Bengalese finch.

This condition is chosen because a song that a female can interject perfectly (i.e., one that she can predict completely) cannot be novel. Unless this novelty condition is satisfied, the male singing is not suitable for mating. The song novelty is not based on the song string itself, but on the interjection communication between males and females.

In our model, song complexity and novelty are relative notions, which are based on the difficulty of successful interjections depending on diverse females' preferences. The modeling is described in detail in Section 2.3.

## 2.2  Male and Female Birds

We model the communication of artificial birds after considering both interjection communication and song novelty. Males and females are modeled with different types of FA [7].

The song syntax of a male, $G$, is represented by a sequential machine defined by the following 6-tuple:

$$G = (Q, \Sigma, \Delta, \delta, \lambda, q_0), \tag{1}$$

where $Q$ is a finite set of states, $q_0$ is an initial state, $\Sigma$ is a finite set of input symbols, $\Delta$ is a finite set of output symbols, $\delta$ is a state transition function, $Q \times \Sigma \rightarrow Q$, $\lambda$ is an output function, and $Q \times \Sigma \rightarrow \Delta$. For the output symbols $\Delta = \{blank, A, B, \ldots, J\}$, each letter represents a song chunk and *blank* denotes a silent interval between the chunks. The input symbols $\Sigma$ are identical to the output symbols $\Delta$, except for the start symbol, which is a primary input to start generating a sequence of chunks. A male sings a song (i.e., a chunk output) with a song syntax $G$. A combination of chunks between blanks constitutes a phrase, and the entire output sequence expresses a song. Note that for simplicity, the input symbols are omitted in all the illustrations of a song syntax.

The preference of a female, $P$, is expressed by a different type of FA as follows:

$$P = (Q, \Sigma, \delta, q_0, F), \tag{2}$$

where $Q$, $\Sigma$, $\delta$, and $q_0$ are the same as above, and $F$ is a set of accepting states, which is a subset of $Q$. This FA determines a female's reaction to the songs. A female preference $P$ has two internal states for the input sequences: accepting states for "interject" and non-accepting states for "no reaction." A female changes her internal state by listening to a chunk in a song (receiving the output from a male as the input), and interjects when her internal state becomes an accepting state. When a female hears preferred patterns of chunks and interjects to breaks (i.e., *blanks*) in a song, it is deemed a successful interjection. Note that an output sequence consisting of only blanks is not a song, and is therefore ignored. Examples of male and female FAs are illustrated in Figure 3.
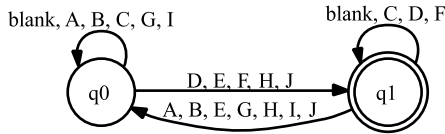
## 2.3  Interjection Communication

We model an ecosystem by using the above-mentioned different types of FA. At each time step, a male randomly chooses a female and sings a song of length $L_{song}$ with his song syntax $G$. Each male sings for a maximum of $L_{song}^{total}$ and attracts several females during this period, where $L_{song}^{total}$ denotes the total song length. For example, if a male has $L_{song} = 20$ and $L_{song}^{total} = 50$, he sings 20 chunks for two females and the remaining 10 chunks for another female. In this way, he sings to three females during one time step.

Females interject all the songs they hear according to their song preference $P$. An example of interjection communication is shown in Figure 3. After an interjection communication, a female

**Song syntax (G)**
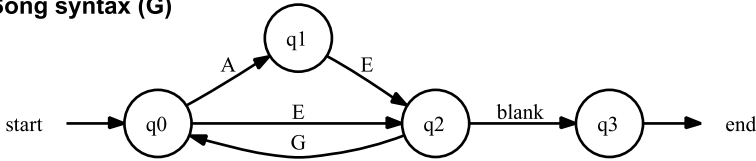
**Preference (P)**

**Interjection Communications**

| **Male A:** | `AEGE_E_E_AEGE_EGAEGE_EGAE_E_AE_E_` . . . |
|---|---|
| **Female A:** | `_#_#`**#**`___#`**#**`_#_#`**#**`____#_#`**#**`____#`**#**`____#`**#**`_`**#**`_` . . . |
| **Female B:** | `_##___#_#____##____##____#__#_#_##_#_` . . . |

Figure 3. Example of song syntax (G), preference (P), and interjection communications. Male A sings with song syntax G. The string of letters represents a song of male A, where '_' denotes a silent interval. Female A listens to the song and interjects according to her song preference P, where the double circle denotes an accepting state of P. In this communication, some interjections are successfully performed (shown by boldface '**#**'), while the others are failures (shown by italic '*#*'). Female B with different P does not succeed in interjection.

evaluates communicational compatibility with a male by assigning a communication score, which is the sum of three weighted terms:

$$S = \frac{1}{w_1 + w_2 + w_3} \left( w_1 \cdot QN + w_2 \cdot QL + w_3 \cdot DC \right), \tag{3}$$

where

$$QN = \text{quantity of interjection} = \frac{1}{N_{\text{interj}}^{\text{th}}} \quad \min(N_{\text{interj}}^{\text{succ}}, N_{\text{interj}}^{\text{th}}),$$

$$QL = \text{quality of interjection} = \frac{N_{\text{interj}}^{\text{succ}}}{N_{\text{interj}}^{\text{all}}},$$

$$DC = \text{density of chunks} = \frac{N_{\text{chunk}}}{L_{\text{song}}},$$

and where $w_i$ denotes the weight to each score term. The communication score considers both the quantity and quality of interjection, and the density of chunks in a song. $QN$ is proportional to the number of successful interjections, $N_{\text{interj}}^{\text{succ}}$, below the threshold $N_{\text{interj}}^{\text{th}}$. If $N_{\text{interj}}^{\text{succ}} \geq N_{\text{interj}}^{\text{th}}$, the female's

evaluation is saturated and $QN$ becomes one. $QL$ denotes the success rate of interjection: the ratio of the total number $N_{\text{interj}}^{\text{all}}$ of interjections to the number $N_{\text{interj}}^{\text{succ}}$ of successful interjections. $DC$ represents the fraction of nonempty chunks in a song. The communication score varies in the range $0 \leq S \leq 1$. Unless the song novelty condition is satisfied, a singing male is not eligible for mating, and the communication score becomes zero.

The important point in this interjection communication is the relative measure of song complexity and novelty by females. For example, as shown in Figure 3, although a song by male A is good and is preferred by female A, this does not mean that other females will also prefer the same song. Moreover, female A will not always prefer the song of male A if his song syntax has branches, which enable stochastic outputs.

## 2.4  Evolution

The coevolution of males and females is driven by interjection communication. Each female selects a male according to the communication scores, with the highest score representing her choice of mate. In this model, monogamy is assumed, so that each male mates with only one female, although he sings to many. By assuming that females produce offspring in proportion to their communication score, the number of offspring is calculated as follows:

$$\#\text{offspring} = C_{\text{offs}} \cdot S. \tag{4}$$

An offspring, whose gender is randomly assigned, is added into the ecosystem as a new bird. Since chicks learn songs from their fathers, or may have song preferences similar to that of their mothers due to their upbringing, their characters become similar to those of their parents. Therefore, in our model, the chicks inherit FAs (i.e., song syntaxes and song preferences) similar to those of their parents; these are varied using one of the following operations. We use the same occurrence rate for each mutation.

(a) **Arrow mutation**: With $N_{\text{node}}$ remaining fixed, one of the following—changing, adding, or deleting (only in males)—happens to a randomly selected arrow. An arrow is connected to another node or is deleted; otherwise the arrow has a new connection with probability $1/N_{\text{chunk}}^{\text{total}}$.

(b) **Node mutation**: A new node is added as the last node ($q_{\text{last}}$), and it has a transition from the next node ($q_{\text{last}-1}$); otherwise, a randomly selected node is removed, and all the connections to that node are also removed.

(c) **Random mutation**: A new FA is created randomly. $N_{\text{node}}$ is set to be the same as that of the parent. Appropriate connections are made with probability $1/N_{\text{chunk}}^{\text{total}}$.

The list (a)–(c) expresses the possible inaccuracy of a chick inheriting parental characteristics in that order; that is, (a) represents the highest accuracy of inheritance, and (c) represents complete failure to inherit any parental characteristics. In addition, the following mutation is performed for male chicks:

(d) **Song length mutation**: Change $L_{\text{song}}$ ($\pm 5$) and $L_{\text{song}}^{\text{total}}$ ($\pm 2$).

There are two biological constraints: lifetime and ecological capacity. Every bird has a lifetime $T_{\text{life}}$, after which it is removed from the ecosystem. To limit the maximum size of the ecosystem, some birds are selected at random and removed to reflect the ecological capacity. The number of birds removed is proportional to the total number of birds:

$$\#\text{birds removed} = C_{\text{eco}} \cdot (N^{\text{male}} + N^{\text{female}}). \tag{5}$$
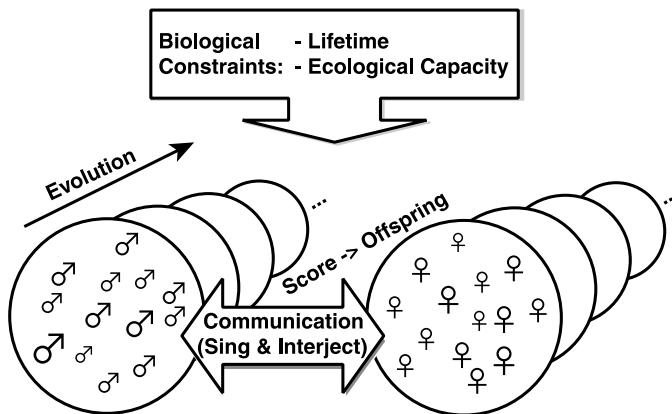
Figure 4. Schematic representation of the evolution of artificial birds. Under biological constraints, the male and female birds (different types of FA) communicate, replicate, and coevolve.

In summary, the coevolution of artificial birds progresses via interjection communications, genetic mutation, and biological constraints, over time. A schematic of our simulation is shown in Figure 4.

## 3   Simulations

We show two representative types of simulation results: one single experiment, and systematic experiments with parameter control. The simulation condition is listed in Table 1.

### 3.1   A Case Study

This section describes typical results of a single simulation run. At the initial state, every male bird had a random FA with $N_{\mathrm{node}} = 2$, at which a transition arrow of the FA is created with a probability of $1/N_{\mathrm{chunk}}^{\mathrm{total}}$ ($= 1/11$), where the song chunks consist of {*blank*, A, B, . . . , J}. The song length and total song length were set to $L_{\mathrm{song}} = 10$ and $L_{\mathrm{song}}^{\mathrm{total}} = 50$, respectively; hence, males can initially sing to five females. Meanwhile, every female bird initially had a random FA with $N_{\mathrm{node}} = 2$. The initial population size of males and females is 100. In this simulation, we adopted the neutral score weights $w_1 = w_2 = w_3 = 2$, and the song novelty condition $N_{\mathrm{novel}} = 1$. The other parameters were set as listed in Table 1.

Table 1. Simulation condition.

| | |
|---|---|
| Initial birds | $N^{\mathrm{male}} = N^{\mathrm{female}} = 100$, $N_{\mathrm{node}}^{\mathrm{male}} = N_{\mathrm{node}}^{\mathrm{female}} = 2$, $L_{\mathrm{song}} = 10$, $L_{\mathrm{song}}^{\mathrm{total}} = 50$ (male only) |
| Communication | $w_1 = w_2 = w_3 = 2$ (Sections 3.1.2, 3.2.3), $w_i$: varying (Section 3.2.4), $N_{\mathrm{chunk}}^{\mathrm{total}} = 11$, $N_{\mathrm{interj}}^{\mathrm{th}} = 100$ (see Equation 3), $\max(L_{\mathrm{song}}^{\mathrm{total}}) = 500$, $N_{\mathrm{novelty}} = 1$ (except Section 3.2.2), $N_{\mathrm{novelty}}$: varying (Section 3.2.2) |
| Mating | $C_{\mathrm{offs}} = 4/3$ (see Equation 4) |
| Lifetime | $L_{\mathrm{life}}$: randomly assigned between 1 and 10 |
| Ecological capacity | $C_{\mathrm{eco}} = 0.1$ (see Equation 5) |

### 3.1.1 Stepwise Evolution of Population

The population dynamics of the artificial birds are shown in Figure 5. We observe a *stepwise* evolution, in which the populations of males and females increased rapidly at $t = 4,000$; subsequently, they increased gradually toward a maximum, as determined by the ecological capacity. Whether such a stepwise change occurs depends on both the parameters $N_{interj}^{th}$ (Equation 3) and $C_{offs}$ (Equation 4) that affect the number of offspring. Stepwise evolution is observed for a wide range of these parameters, provided they are not too big or small (i.e., the reproduction rate is not too high or low).

### 3.1.2 Emergence of Courtship Strategies

The evolution of song communication is depicted in Figure 6, which shows the average values at each time step: the communication score ($S$), the song length ($L_{song}$), the total song length ($L_{song}^{total}$), and the success rate of communication ($R_{commu}^{succ}$), which is the ratio of communications with nonzero scores to total communications. Although the total song length was much longer, before $t = 4,000$ the males sang relatively short songs. A comparison of $L_{song}$ with $L_{song}^{total}$, shown in Figure 6, reveals that males were able to successfully find mating partners by singing to many females, and females heard many songs before $t = 4,000$.

Immediately thereafter, both the average communication score and the success rate of communication increased rapidly, indicating that the males that could sing longer songs appeared and the communication became much better. This caused stepwise evolution, as described in Figure 5. At this point, longer songs were increasingly preferred, and the males began to sing longer songs to very few females during the time $4,000 \leq t \leq 15,000$ (see Figure 6).

After $t = 15,000$ in Figure 6, this trend stopped, and the males began singing songs of moderate length ($L_{song} = 200$) to a few females. During this period, the communication score increased even though the success rate of communication began decreasing. At this point, three types of courtship strategies can be associated with the song length: (i) short songs to many females ($0 \leq t \leq 4,000$), (ii) longer songs to one, or very few, females ($4,000 \leq t \leq 15,000$); and (iii) songs of moderate length to a few females ($15,000 \leq t \leq 20,000$). These observations indicate that the emergence of longer songs affected interjection communication and led to the transition from courtship strategy (i) to (ii), and that males eventually evolved courtship strategy (iii), which avoided the risk of failure inherent in singing for a length $L_{song}^{total}$ to only one female.
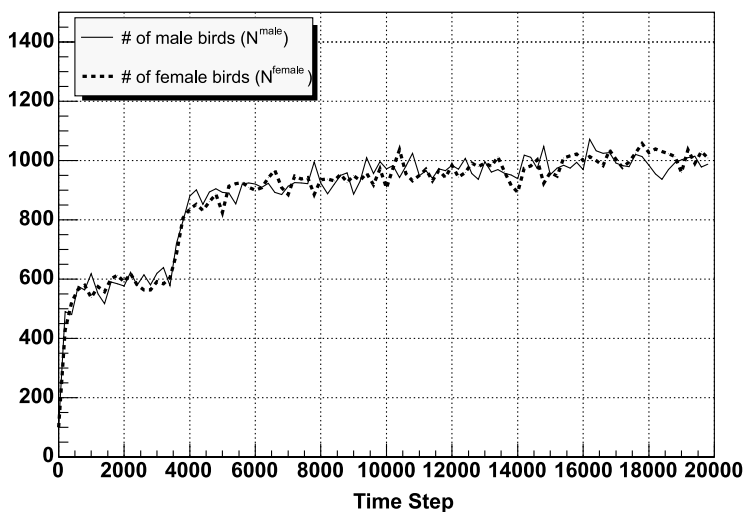


Figure 5. Population dynamics: Stepwise evolution is observed. After around $t = 4000$, the population size rapidly increases. This trait is closely related to the courtship strategies of male birds.
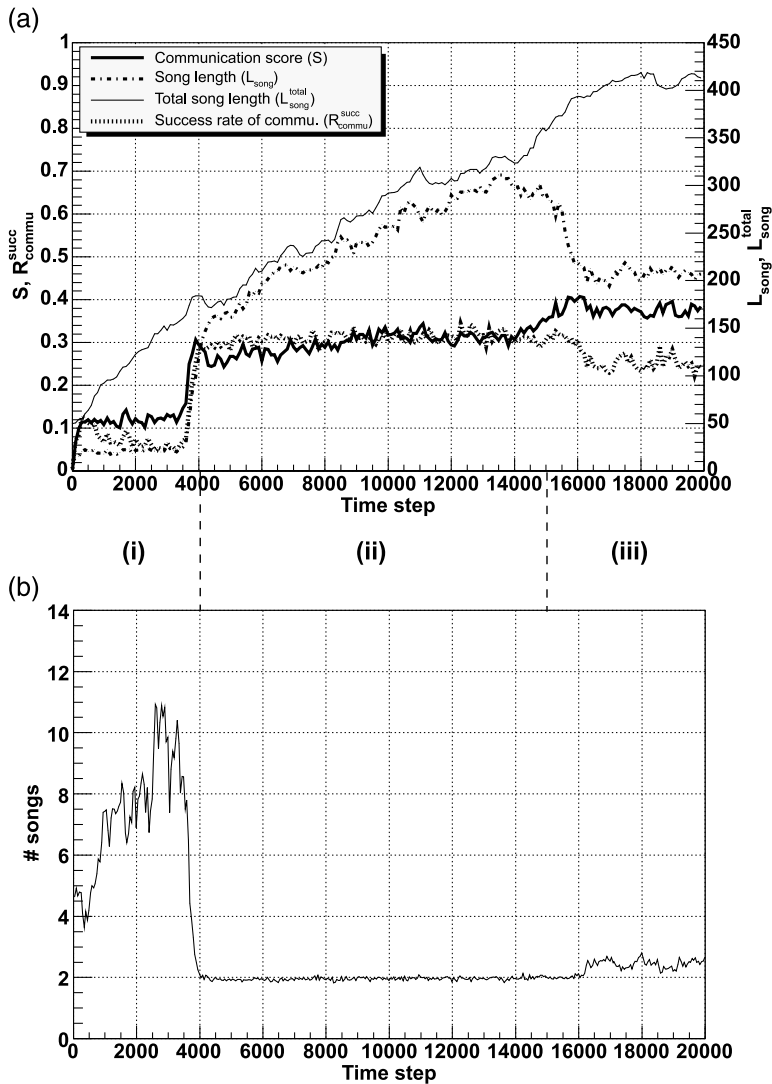
Figure 6. Evolution of interjection communication. (a) Each line shows the average values of communication score ($S$), song length ($L_{song}$), total song length ($L_{song}^{total}$), and success rate of communication ($R_{commu}^{succ}$). Immediately after $t = 4,000$, male birds with longer songs appear. Subsequently, the communication scores and success rate of communication increase rapidly, and males begin to sing longer songs to very few females at $4,000 \leq t \leq 15,000$. After $t = 15,000$, males with songs of moderate length succeed in communicating with females and become dominant. (b) The number of songs that were heard by females and that were changed according to the transition from courtship strategy (i) to (iii).

### 3.1.3  Asymmetry between Song Syntaxes and Preferences

In automaton theory, it is often held that the complexity of an FA increases in proportion to its number of states (that is, the number of nodes) [3]. Thus, the evolution of the number of nodes reflects the change in the complexity of interjection communication.

After males with longer songs emerged at $t = 4,000$, the number of nodes of the males ($N_{node}^{male}$) increased drastically, as shown in Figure 7, and the males began singing much longer, as shown in Figure 6. This indicates that the males that have longer songs and song syntaxes with many nodes made effective use of these syntaxes to produce novel songs. Examples of the interjection communication are shown in Figure 8.
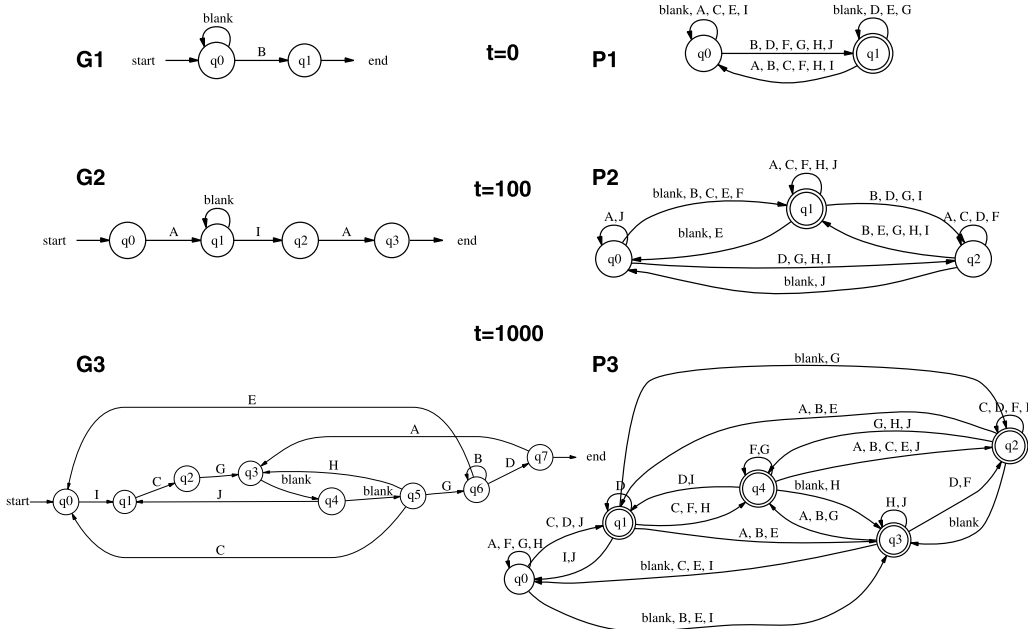
Figure 7. Evolution of finite automata in males and females. The average numbers of nodes in males and females ($N_{node}^{male}$ and $N_{node}^{female}$) increase and then oscillate over time. After $t = 4,000$, the number of nodes of males becomes greater than that of females.

```
t = 0
G1(N_node=2,LI=0.67,L_song=10,L_song^total=50) vs.P1(N_node=2)
_B__BB_B__
_###_##___
score = 0.34


t = 100
G2(N_node=4,LI=0.8,L_song=18, L_song^total =48) vs.P2(N_node=3)
A__IAAIAA__IAA___I
_#____###_#____#__
score = 0.38


t = 1000
G3(N_node=8,LI=0.46,L_song=13,L_song^total =85) vs.P3(N_node=5)
ICG_JCG_JCG__
#__##__##__#_
score = 0.41


t = 1000
G4(N_node=11,LI=0.55,L_song=23,L_song^total =92) vs.P4(N_node=5)
FIHFGEB__DDB__DFIHFDEBB
_____#____#____#___
score = 0


t = 3900
G5(N_node=9,LI=0.69,L_song=99,L_song^total =183) vs.P5(N_node=13)
__GIG__GI__GCDGGIG__GI__GCD__GC_I_I_I__GI__GC_I__GIGGIGGI__GCD__GCDG____I_I_I_____GI__GC_I__GIG_
##__###__#____#___#_____#__#_#_#_#_##___#_#_##____#____##__#_#__#_#####_#_#_#####____#_#_##___#
score = 0.56


t=10000
G6(N_node=17,LI=0.49,L_song=296, L_song^total =298) vs.P6(N_node=11)
___F_F_AA_FFA_C_CDAA_FFA_C_I__C_CDJ__AAA_FFAA_F_A_C_CDJ_FAAAA__FA___F_C_I_AA_CJFAAAA_CJF_C_CDACFA_CJFA___AA__...
###_###____##_##_#____##_#_##_##_##_#__###_#_____#___#_###__####_##_#_#_##_#__##_____#__##_#_____####...
score = 0.68


t=15000
G7(N_node=17,LI=0.38,L_song=306,L_song^total =358) vs.P7(N_node=8)
__A__GIBG__A__G__A__GIFD_BGI_BF__G__A_F__GIBA__A__GA__A__A__A_F__F__GA__A_F__A__GIBGI__FDAJJFDG_FDG__BA__...
_#__#_#__#_#_#__#__##_##__##___##__##_##_##__##__##_##_##_##_##_##_##__##_##_##_##_____#_#_#____#__#_#_#_#...
score = 0.78


t=18000
G8(N_node=4,LI=0.67,L_song=205,L_song^total =422) vs.P8(N_node=6)
A__AA__A__A__A__AA__AA__A__AA__A__A__A__AA__AA__AA__A__A__A__A__AA__AA__A__A__A__AA__A__A__A__AA__A...
###_####__#####__#####_###__#####_####__#####_###_###_###__######__#####_###_####__######__#####__#####_...
score = 0.67
```

Figure 8. Examples of the interjection communication illustrated in Figure 6. The pairs without G4 versus P4 lead to suitable communications. The songs become longer and more complex, especially after $t = 3,900$. However, even in the last stage of evolution, simple songs still exist at $t = 18,000$.

Figure 9. Examples of song syntaxes (*G*) and preferences before the emergence of longer songs (*P*): *G1–G3* represent song syntaxes, and *P1–P3* represent female preferences. *G1/P1*, *G2/P2*, and *G3/P3* are well-suited pairs that yielded a good communication score in the early stages of evolution (also see Figure 8). *G3* is a more complex syntax than either *G1* or *G2*; it has many internal branches. The double circle represents the accepting state at which the females interject.

Simultaneously, Figure 7 shows that the number of nodes of the males was inhibited within an upper bound $N_{\text{node}}^{\text{male}} = 25$. If a male with $L_{\text{song}} = 50$ has many nodes in his FA, such as $N_{\text{node}}^{\text{male}} = 200$, he can use only a part of all the internal states in the FA (a maximum of 25%). This is why the number of nodes of the males does not increase monotonically. Therefore, there is an appropriate song length in which each male can organize chunks and blanks in complex ways.

Conversely, the number of nodes of females did not change as drastically. In the latter stages of evolution, the FAs of females were much simpler than those of males, with respect to the number of nodes.

### 3.1.4 Complexity of Song Syntaxes

In addition to the number of nodes of an FA, we must consider a structural measure of song syntax for analysis. To achieve this, we introduce the *linearity* (*LI*) of song syntax:

$$LI \equiv N_{\text{node}}^{\text{male}}/N_{\text{arrow}}, \tag{6}$$

where $N_{\text{arrow}}$ is the number of arrows. If $N_{\text{node}}^{\text{male}} = N$, then *LI* ranges between $1/N \leq LI \leq 1$ as $N_{\text{arrow}}$ varies from $N^2$ to $N$. Therefore, more complex song syntaxes have lower values of *LI*. The complexity of a song syntax might be characterized by considering both its linearity and the number of nodes.[2]

Figure 9 shows examples of the song syntaxes and preferences. As evolution proceeds, the song syntaxes become more complex (also see Figure 10), and males produce more complex songs, as

---

2  It is insufficient to characterize the complexity of song syntax using only *LI*, because if there are, for example, song syntaxes with ($N_{\text{node}}^{\text{male}}$, $N_{\text{arrow}}$) = (2, 3) and (6, 9), respectively, both will have the same *LI* ($\approx$ 0.67), whereas the latter syntax is in fact more complex because it has many transition states. Therefore, both *LI* and $N_{\text{node}}^{\text{male}}$ must be considered.
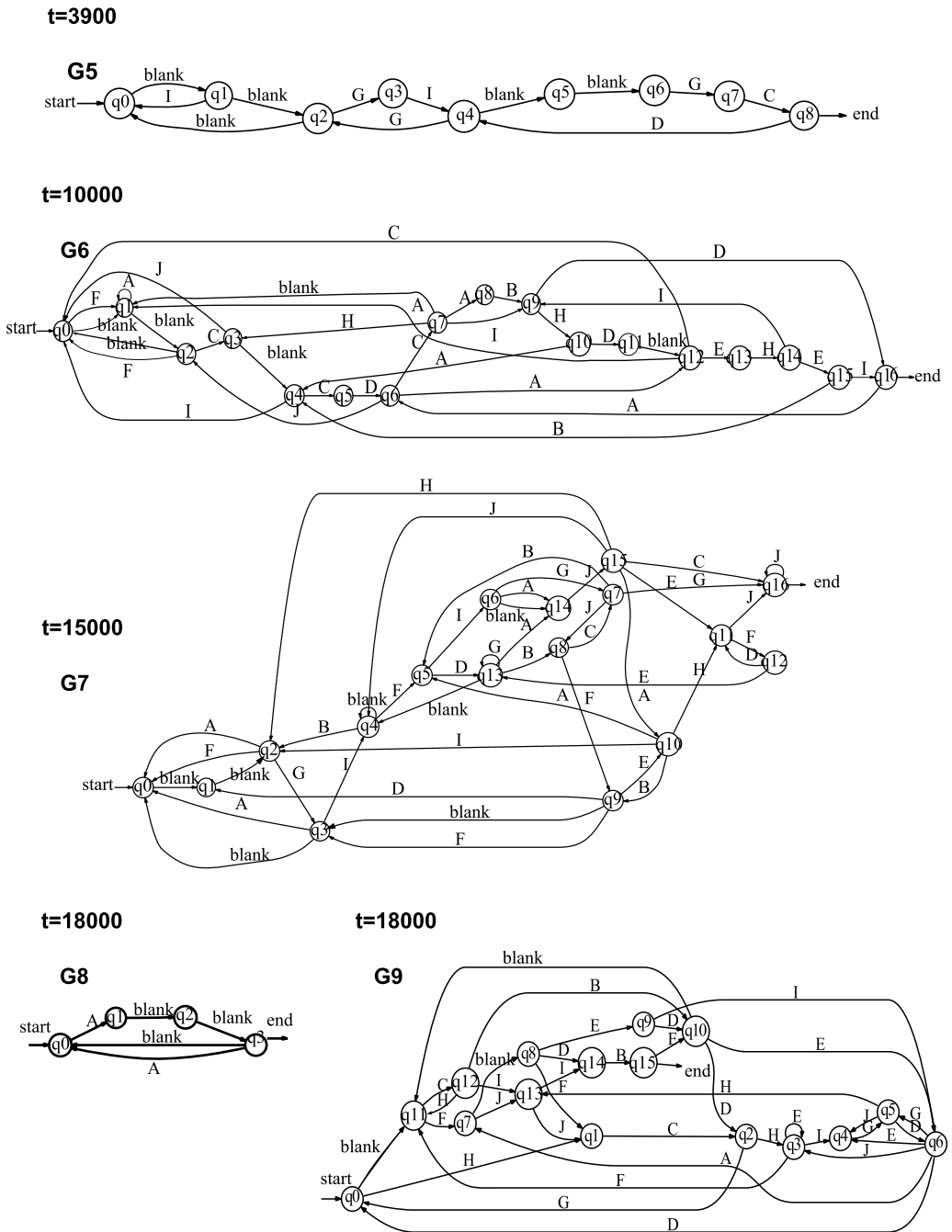
Figure 10. Examples of song syntaxes after the emergence of longer songs: These song syntaxes appeared in the ecosystem after $t = 4,000$ and led to good communication with females, as shown in Figure 8.

shown in Figure 8. Complex song syntaxes including branches can be organized into nondeterministic chunks, thereby avoiding the perfect interjection, which is a prerequisite for the song novelty condition.

Figure 11 shows the average $LI$ of song syntaxes and rate of novel songs. $LI$ decreased gradually from 0.8 to 0.4, and never became less than 0.3, at which point the song syntaxes had three arrows
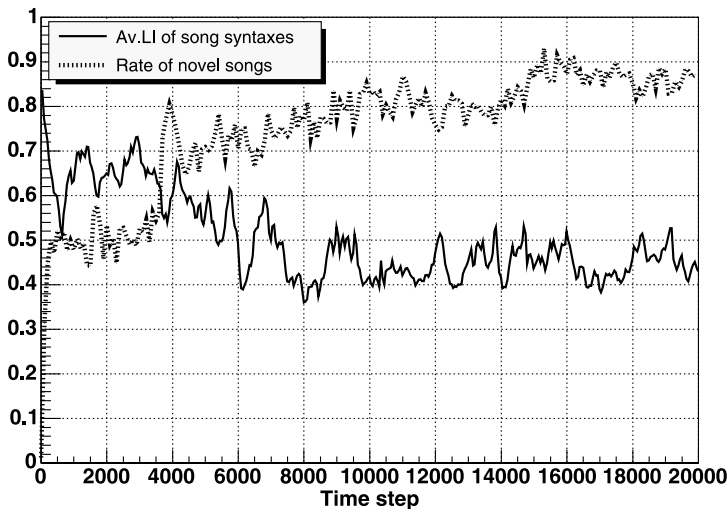
Figure 11. Evolution of the linearity (*LI*) of song syntaxes and song novelty. Song syntaxes gradually became more complex (average *LI* decreased), and the rate of novel songs evolved stepwise. After the emergence of males with longer songs at *t* = 4,000, almost all songs are judged novel.

per node. These syntaxes were not overly stochastic, and females could properly interject, as shown in Figure 8. This result illustrates that an upper bound of song complexity (that is, a lower bound of *LI*) evolved via the interjection communication.

Females rejected almost half the songs due to lack of novelty before *t* = 4,000. From this point onward, the rate of novel songs increased stepwise, and each step correlated with a transition in courtship strategy, that is, between strategies (i) to (iii) that were mentioned in Section 3.1.2. Therefore, the courtship transitions correlated significantly with the increase in song novelty, which was judged by the females.

Figure 12 shows the distribution of the song syntaxes in the $LI$-$N_{node}^{male}$ space. The song syntax distribution was linked to the transition in male courtship strategies: (a) the distribution of the song syntaxes shifted towards more complex syntaxes until *t* = 4,000, at which point males sang simple, short songs; (b) the distribution continued to shift to still more complex syntaxes and then clustered at *t* = 15,000, at which point most of the song syntaxes had similar complexity, and the males sang longer songs to approximately one female; (c) the song syntaxes became distributed over the $LI$-$N_{node}^{male}$ space, and the males sang songs of moderate length to a few females. Although the average $LI$ values at *t* = 15,000 and *t* = 20,000 do not show a significant difference (as shown in Figure 11), the song syntaxes are much more diverse at *t* = 20,000 in the $LI$-$N_{node}^{male}$ space; interestingly, at this point, simple song syntaxes still existed (for example *G8* in Figure 8). Therefore, this evolution is not a simple maximization process in song complexity, and the diversity of song syntaxes is still maintained.

## 3.2 Systematic Studies

This subsection describes systematic experiments with changing parameters, focusing on different aspects of the coevolution of artificial birds.

### 3.2.1 Relationship between Song Syntax and Preference

Simulations were performed by fixing the number of nodes of females ($N_{node}^{female}$) from 2 to 50. The other parameters were identical to the previous simulation (see Table 1). The results are shown in Figure 13. Each pair of bars denotes the average communication score and *LI*, and the line graph shows the number of nodes of males ($N_{node}^{male}$) with ten simulations.
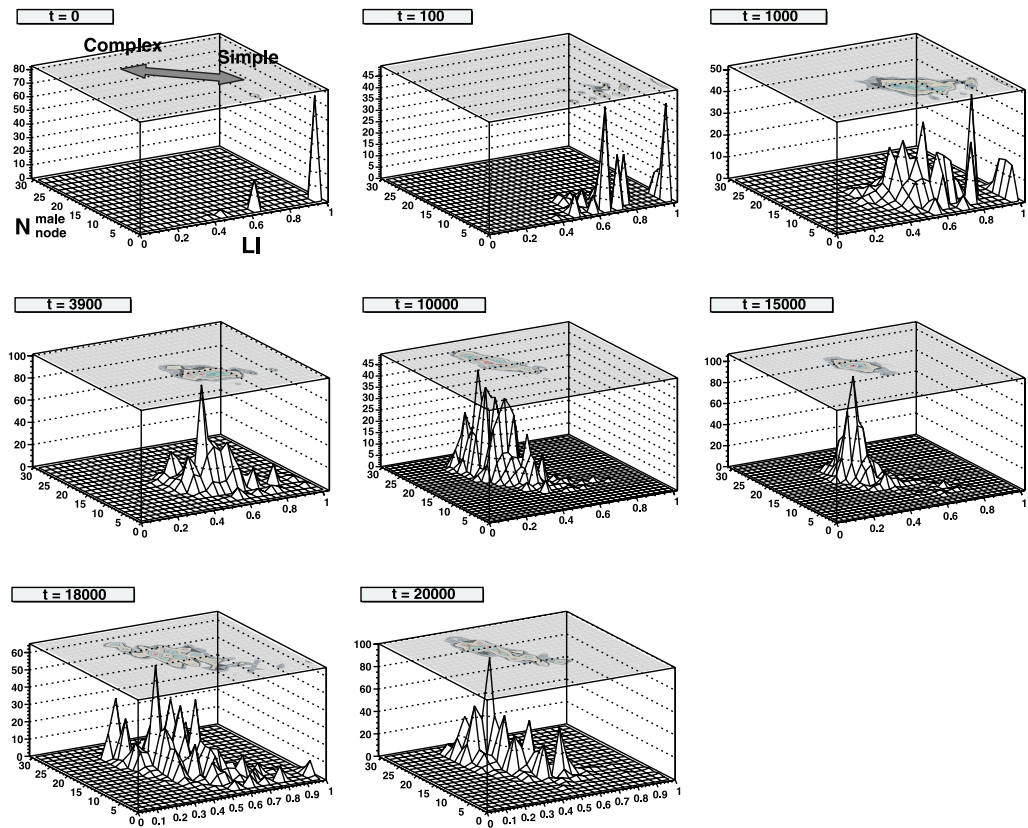
Figure 12. The distribution of the song syntaxes in the $LI$-$N_{node}^{male}$ space: Song syntaxes with greater $N_{node}^{male}$ and smaller $LI$ are considered more complex. The volume of the histogram is proportional to the number of male birds. A projection of the histogram is shown on the top surface of each cube. The distribution of the song syntaxes shifted towards more complex syntaxes with time and clustered at $t = 15{,}000$. Subsequently, in the evolutionary process, the diversity of song syntaxes recovered.
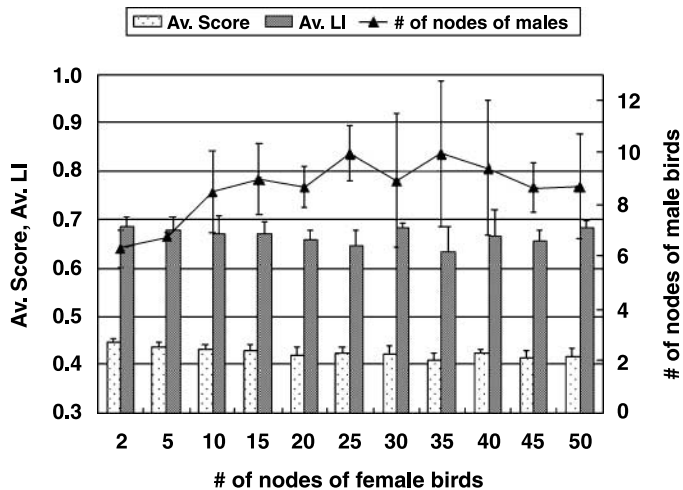


Figure 13. The relationship between song syntax and female preference. Each bar graph shows the average communication score and $LI$, and the line shows the average number of nodes of males as a function of the number of nodes of females. The error bars represent the standard deviation. $LI$ and communication score are not greatly affected by the number of nodes of females.

As shown in Figure 13, the number of nodes of males and that of females were correlated. However, both *LI* of song syntaxes and communication score were insensitive to the number of nodes of females. According to these observations, the structural measure of song syntaxes, *LI*, can evolve independently of the number of nodes of females; simple females with only two nodes could induce the same *LI* and communication score as complex females with many more nodes.

Regarding interjection communication, females only need to recognize previous phrases to succeed in the interjection, even in the case of complex songs. For example, in Figure 8, the female with *P7* interjected well in the case of the known phrases "A__" and "F__" produced by the complex male with *G7* in most, if not all, of his songs.

### 3.2.2   Tradeoff between Song Familiarity and Novelty

This section discusses the influence of song novelty on the evolution of song syntax complexity. Simulations were performed with the same parameters as in previous simulations, while the song novelty condition $N_{novel}$ was changed.

The results, summarized in Figure 14, show the average values based on 10 runs, the numbers of nodes in males and females ($N_{node}^{male}$ and $N_{node}^{female}$), and *LI* of song syntaxes. There were no successful simulations when $N_{novel} \geq 6$, which is an extremely difficult condition for successful communication; this is because the initial males could sing only for a length $L_{song} = 10$. As shown in Figure 14, there are significant differences among the FAs, depending on the song novelty. When $N_{novel} = 0$, the evolution of complex FAs was inhibited in both males and females, that is, all the FAs had a small number of nodes. Songs for which $N_{novel} = 0$ are considered familiar songs (that is, non-novel songs), because females can predict them almost perfectly. In contrast, when $N_{novel} \geq 1$, all the respective values are similar.

A tradeoff between song familiarity and novelty determines whether the song novelty condition is zero; that is, simple song communication can evolve when $N_{novel} = 0$. However, in order to evolve complex song communication, $N_{novel}$ must be at least one.
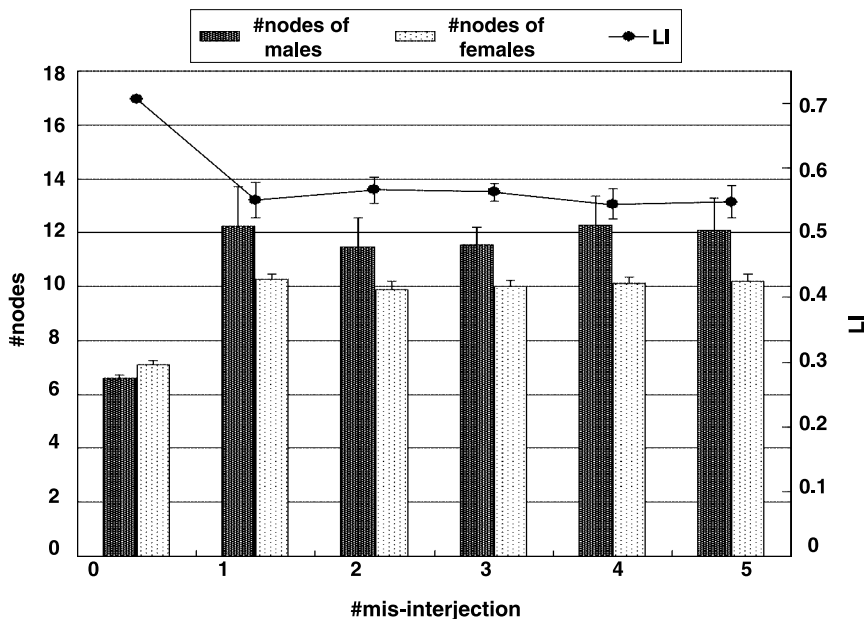


Figure 14. Dependence of song novelty condition. The number of nodes and linearity of song syntaxes ($N_{node}^{male}$ and *LI*) are similar for each $N_{novel}$ greater than one. In contrast, $N_{node}^{male}$ is considerably lower and *LI* is much higher at $N_{novel} = 0$. The error bars represent standard deviations.

### 3.2.3   Random Interjection and Random Mating

To compare the interjection communication with other factors, we introduce *random mating* and *random interjection*. In random interjection, each female interjects randomly in every song of a preferred male, and then selects a mate with the highest communication score. In random mating, each female interjects in the songs of many males, but then selects a mate randomly, regardless of the communication score.

Simulations were performed using interjection communication, random interjection, and random mating. The result, summarized in Figure 15, shows the average values of the communication score and *LI* of the song syntaxes, which were calculated from 10 runs. A comparison of *LI* among the three factors reveals that all the factors produced comparable *LI* of song syntax. Apparently, the two introduced methods could enhance the complexity of song syntaxes through a random drift process. However, the song syntaxes evolved via the interjection communication have significantly lower linearity than that of the other factors, because its dynamic range is very narrow. The major difference among the three factors was the communication score, as shown in Figure 15. The interjection communication led to a considerably more successful communication than that through other random means. Therefore, the complexity of a song syntax is not measured by *LI* only; *LI* should be considered together with the communication score. These results clearly show that interjection communication can sufficiently enhance the complexity of song syntaxes to self-organize coherent communication with females.

### 3.2.4   Weights of Communication Scores

Thus far, we have discussed the evolution based on communication scores with the neutral weights ($w_1 = w_2 = w_3 = 2$). In this section, we discuss the evolutions of song communication when the weights of the communication scores are varied. Under the condition $w_1 + w_2 + w_3 = 6$, $w_1$ and $w_2$ are changed. The song novelty condition is fixed as $N_{\mathrm{novel}} = 1$.

Figure 16 shows the results based on ten runs. The phase diagrams of (a) song length, (b) number of nodes of males ($N_{\mathrm{node}}^{\mathrm{male}}$), and (c) *LI* of song syntaxes are shown. The white part in the figures represents either no parameter settings (e.g., $w_1 = w_2 = 4$ is prohibited because the sum of $w_i$ is fixed at 6) or absence of successful simulations, where all the birds became extinct due to bad communication during the initial stages.

As shown in Figure 16a, songs tend to become longer as $w_2$ decreases, during which time $QL$ contributes less and females begin to prefer longer songs. The stepwise evolution of the song length is shown in the gray region in Figure 16a. When $w_1 = w_2 = w_3 = 2$, 83 out of 100 simulation runs
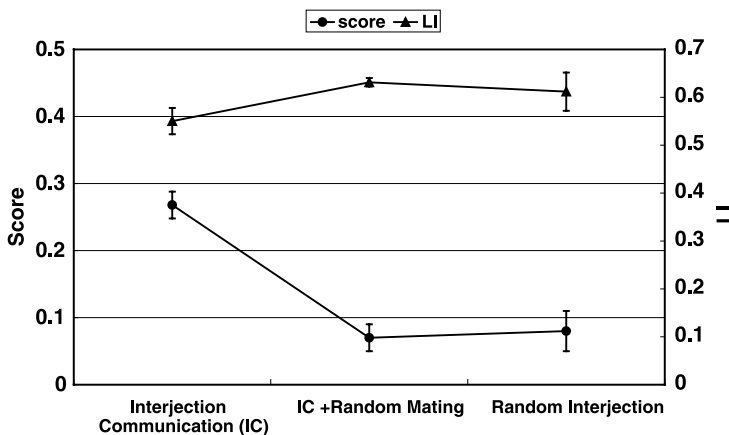


Figure 15. A comparison of the different communication methods. Interjection communication (IC) is advantageous for evolving complex and coherent song communication as compared with random interjection communication and random mating.
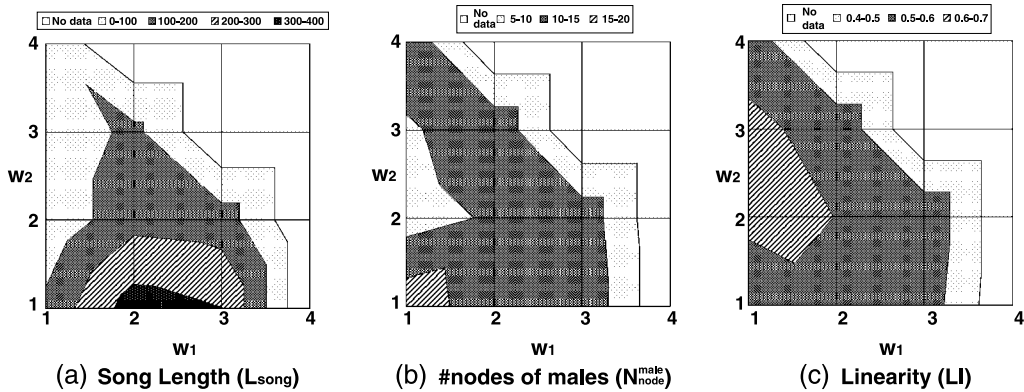
Figure 16. Phase diagrams of the features of male birds, including (a) song length, (b) number of nodes, and (c) linearity of song syntaxes. As $w_2$ decreases, longer songs tend to evolve. The most complex song syntaxes evolve for intermediate values of $w_1$ and $w_2$.

resulted in the three-step evolution of song length; all the other conditions led to a two-step evolution without the second stage of evolution (that is, (ii) in Figure 6) [16].

Since $w_1$ and $w_2$ increased simultaneously, both $N_{node}^{male}$ and $LI$ increased (i.e., the complexity of song syntax decreased). The most complex song syntaxes evolved with intermediate values of $w_1$ and $w_2$; the simplest song syntaxes evolved for $w_1 = 1$ and $w_2 \approx 2$.

## 4  Discussion and Conclusion

We demonstrated the coevolution of male song syntaxes and female preferences in songbirds with a coupling system having different types of finite automata. Although this model was a very simple one, it helped elucidate important features of syntactic animal communication.

Our simulations showed that the male song syntax became more complex (i.e., less linear) when females used interjection communication, as opposed to random communication or random mating. The linearity of the song syntax was relatively independent of the complexity of female preference (i.e., even females with two nodes in their preference automata evolved less linear song syntaxes). As previously mentioned, the Bengalese finch is a domesticated strain of the white-backed munia. Domestication might have freed the birds from predatory pressures, thereby helping to facilitate the evolution of a complex song syntax [14, 24]. However, our simulations revealed another possibility. The ability of females to recognize novelty, which is defined by incomplete interjection, might control the evolution of complexity and diversity in song syntaxes.

Werner and Todd [23] demonstrated similar simulations by focusing on the evolution of diverse signals driven by sexual selection. In their model, a "song" of a male was encoded as a genotype, not a song-generating "machine"; the sizes of machines for males and females were fixed. In addition, their study lacked a discussion of complexity evolved via sexual selection. On the other hand, our model not only showed diverse signals, but also presented the diverse structures of these signals (i.e., song and preference automata); further, the structures themselves could evolve. As a result, we obtained new results: The song preferences of females could control (not only enhance, but also constrain) the complexity and diversity of song syntaxes, and the dynamic transitions of male courtship strategies originated from interjection communication.

Our syntax evolution model differs from that of Hashimoto and Ikegami [4], which used a common formal syntax for both speech and recognition for each agent. This model showed that an ensemble of agents evolve a complex syntax, which is not identical to universal grammar that can potentially generate and recognize any sentence; instead, it is a dominant syntax that is shared by many agents (called net grammar). The model assumed *symmetric* communication, where each agent

plays the role of both speaker and listener. In contrast, we simulated asymmetric communication; namely, males sing and females listen. Different song syntaxes evolved, but they were not shared by all agents. In addition, as reflected by our simulations, the division of labor might bifurcate syntax similarly to the way in which human language is shaped by society.

In this model, we neglected the spatial structure. Due to the lack of spatial constraint (it was possible for every male bird to communicate with every female bird), our ecosystem became saturated with similar males and females, and the rate of evolution slowed after $t = 20,000$. To consider open-ended evolution of syntactic animal communication, we must consider spatial factors, such as the spatiotemporal patterns of coevolving male and female birds.

## Acknowledgments

## References

1. Cangelosi, A., & Parisi, D. (1998). The emergence of a "language" in an evolving population of neural networks. *Connection Science, 10*(2), 83–97.

2. Catchpole, C. K., & Slater, P. J. B. (1995). *Bird song: Biological themes and variations.* Cambridge, UK: Cambridge University Press.

3. Crutchfield, J. P. (1994). The calculi of emergence: Computation, dynamics and induction. *Physica D, 75,* 11–54.

4. Hashimoto, T., & Ikegami, T. (1996). Emergence of net-grammar in communicating agents. *Biosystems, 38,* 1–14.

5. Hauser, M. D. (1997). *The evolution of communication.* Cambridge, MA: MIT Press.

6. Honda, E., & Okanoya, K. (1999). Acoustical and syntactical comparisons between songs of the white-backed munia (*Lonchura striata*) and its domesticated strain, the Bengalese finch (*Lonchura striata* var. *domestica*). *Zoological Science, 16,* 319–326.

7. Hopcroft, J. E., & Ullman, J. D. (2006). *Introduction to automata theory, languages and computation,* (3rd ed.). Reading, MA: Addison-Wesley.

8. Hosino, T., & Okanoya, K. (2000). Lesion of a higher-order song control nucleus disrupts phrase-level complexity in Bengalese finches. *NeuroReport, 11,* 2091–2095.

9. Iwasa, Y., Pomiankowski, A., & Nee, S. (1991). The evolution of costly mate preferences. II. The 'handicap' principle. *Evolution, 45,* 1431–1442.

10. MacLennan, B. (1992). Synthetic ethology: An approach to the study of communication In C. G. Langton, C. D. Taylor, D. Farmer, & S. Rasmussen (Eds.), *Artificial life II* (pp. 631–658). Reading, MA: Addison-Wesley.

11. Miller, G. F., & Todd, P. M. (1993). Evolutionary wanderlust: Sexual selection with directional mate preference. In M. Jean-Arcady, L. R. Herbert, & W. W. Stewart (Eds.), *From Animals to Animats 2: Proceedings of the Second International Conference on Simulation of Adaptive Behavior* (pp. 21–30). Cambridge, MA: MIT Press.

12. Okanoya, K. (2002). Sexual display as a syntactical vehicle. In A. Wray (Ed.), *The transition to language* (pp. 46–63). New York: Oxford University Press.

13. Okanoya, K. (2004). Song syntax in Bengalese finches: Proximate and ultimate analyses. *Advances in the Study of Behavior, 34,* 297–346.

14. Ritchie, G., & Kirby, S. (2005). Selection, domestication, and the emergence of learned communication systems. In *Proceedings of the Second International Symposium on the Emergence and Evolution of Linguistic Communication.*

15. Sasahara, K., & Ikegami, T. (2003). Coevolution of birdsong grammar without imitation. In B. Wolfgang, C. Thomas, D. Peter, T. K. Jan, & Z. Jens (Eds.), *Advances in artificial life* (pp. 482–490). Berlin: Springer.

16. Sasahara, K., & Ikegami, T. (2004). Song grammars as complex sexual displays. In J. Pollack, M. Bedau, P. Husbands, T. Ikegami, & R. A. Watson (Eds.), *Artificial Life IX* (pp. 194–199). Cambridge, MA: MIT Press.

17. Smith, J. M., & Harper, D. (2003). *Animal signals*. New York: Oxford University Press.

18. Steels, L. (1997). Self-organizing vocabularies. In C. G. Langton & K. Shimohara (Eds.), *Artificial Life V* (pp. 179–184). Cambridge, MA: MIT Press.

19. Suzuki, J., & Kaneko, K. (1994). Imitation game. *Physica D*, *75*, 328–342.

20. Tamura, J. (2004). *Audio-visual dependence in zebra finches* (in Japanese). Master's thesis, Chiba University.

21. Tobari, Y., Nakamura, K. Z., & Okanoya, K. (2005). Sex differences in the telencephalic song control circuitry in Bengalese finches (*Lonchura striata* var. *domestica*). *Zoological Science*, *22*, 1089–1094.

22. Werner, G. M., & Dyer, M. (1992). Evolution of communication in artificial organisms. In C .G. Langton, C. D. Taylor, D. Farmer, & S. Rasmussen (Eds.), *Artificial Life II* (pp. 659–687). Reading, MA: Addison-Wesley.

23. Werner, G. M., & Todd, P. M. (1997). Too many love songs: Sexual selection and the evolution of communication. In P. Husbands & I. Harvey (Eds.), *Fourth European Conference on Artificial Life* (pp. 434–443). Cambridge, MA: MIT Press.

24. Wiles, J., Watson, J., Tonkes, B., & Deacon, T. B. (2005). Transient phenomena in learning and evolution: Genetic assimilation and genetic redistribution. *Artificial Life*, *11*, 177–188.

25. Zahavi, A., & Zahavi, A. (1997). *The handicap principle*. London: Oxford University Press.