“Piep Piep Piep - Ich Hab’ Dich Lieb”:
Rhythm as Indicator of Mate Quality

Eva van den Broek
June 2003
"Everyone wants to understand art. Why not try to understand the song of a bird?"

Pablo Picasso

The title means "Peep peep peep, I love you" derived from the song "Guilo hat euch lieb", by Guildo Horn And The Orthopaedic Stockings, Germany's entry in the Eurovision Song Contest 1998:
http://willow.dyndns.org/rachel/doh/jukebox/guilo.htm
Acknowledgements

First and foremost, I would like to thank my supervisors: Peter for having me as a student and introducing me to the scientific world, and most of all, for his everlasting patience in answering all my questions by rephrasing them. Frans encouraged me to go abroad in the first place and Carel was of great help to me with a bird's eye view.

Furthermore, I’d like to thank all members of the ABC-group in Berlin for providing me with feedback, cakes and diverting discussions for almost a year.

Special thanks to Eefje Leydesdorff, Johan Bolhuis, John Hutchinson, Miguel Rodriguez-Gironés, Joost Belthuis, Leendert van Maanen, Stefan van der Stigchel and Richard McElreath for their critical comments on the project or the manuscript;

To Anja Dieckmann for coming up with the title during our daily Gummibärchen-tasting;

To my parents for being proud of me, regardless what I told them I was busying myself with in Berlin, and to my brother Daan for drawing schematic pictures of my cryptic thoughts;

And finally, I need to thank Freek for bringing the relativity of a thesis into mind.

Parts of this study will be published as ‘Piep Piep Piep - Ich Hab’ Dich Lieb: Rhythm as an Indicator of Mate Quality’, Proceedings of the 7th European Conference on Artificial Life.
Acknowledgements ............................................................................................................. 3

1 Introduction ..................................................................................................................... 6

1.1 Research question ........................................................................................................ 8
1.2 Description of method ................................................................................................... 8
1.3 Position within Cognitive Artificial Intelligence ......................................................... 9
1.4 Status quaestionis ......................................................................................................... 9
1.5 Overview of the structure .......................................................................................... 10

2 Evolutionary simulation .................................................................................................. 12

2.1 Validity of models ....................................................................................................... 12
2.2 What can be concluded from models? ....................................................................... 14
2.3 Evolutionary models .................................................................................................... 16
2.4 Modeling signal design .............................................................................................. 18
2.5 The view from the receiver: strategic design of signals ............................................ 19
2.6 The view from the signaler: efficacy considerations on signal design..................... 23
2.7 Modeling rhythm ......................................................................................................... 25

3 Birdsong ............................................................................................................................. 27

3.1 A bird’s life ................................................................................................................... 27
3.2 Constraints on song .................................................................................................... 29
3.3 Neural temporal processing ....................................................................................... 34
3.4 Mistakes and female reaction .................................................................................... 37
3.5 Rhythm in birdsong ..................................................................................................... 39

4 The model ......................................................................................................................... 41

4.1 Description of the model ............................................................................................ 41
4.2 Parameters and their influence .................................................................................. 44
4.3 Singing and judging .................................................................................................... 47
4.4 Expectations ................................................................................................................ 49
5 Results ........................................................................................................................................51
  5.1 Outline ....................................................................................................................................51
  5.2 Template discriminability ........................................................................................................52
  5.3 Influence on female choice .......................................................................................................54
  5.4 Finding local optima ................................................................................................................56
  5.5 Stability of the solutions ..........................................................................................................58
  5.6 Evolving rhythmicity ................................................................................................................58

6 Conclusion and discussion ........................................................................................................60
  6.1 Summary ..................................................................................................................................60
  6.2 Discussion ................................................................................................................................61
  6.3 References ................................................................................................................................65

Appendix A ......................................................................................................................................70
1 Introduction

Rhythmically repeated behaviors are common in nature: locomotion, breathing, chewing, and the like all rely on regular repetition for their effectiveness. In addition to these important life functions, animal signals and displays often take a rhythmic form, from the regular flashing of fireflies or stridulations of crickets to the alarm calls of whales or the songs of birds and humans. Why would so many signal forms have such a regular appearance? The fact that also a human propensity for rhythmic utterances exists makes this question even more puzzling, but as Darwin noticed in ‘The Descent of Man’, (1878) we should not confuse our taste of beauty with questions about function:

"It is a curious fact that in the same class of animals, sounds so different as the drumming of the snipe's tail, the tapping of the woodpecker's beak, the harsh trumpet-like cry of certain water-fowl, the cooing of the turtle-dove, and the song of the nightingale, should all be pleasing to the females of the several species. But we must not judge of the tastes of distinct species by a uniform standard; nor must we judge by the standard of man's taste. Even with man, we should remember what discordant noises, the beating of tom-toms and the shrill notes of reeds, please the ears of savages." (chapter 13, section 26)

So the beauty of rhythm as perceived by humans, whether savages or not, cannot serve as an answer to the question why regularity prevails in animal signals. On the contrary, given that the function of a signal depends on grabbing the attention of others animals, surely the sender would do better displaying a surprisingly irregular pattern that is in utmost contrast with his usual rhythmic motor patterns. Still, rhythmic displays are widespread both in the visual as well as in the auditory domain. If we want to know why this very feature appears in a particular signaling system, we need to determine if displaying rhythmicity serves an adaptive function. What evolutionary forces could have shaped the form of signals in this direction?

In the citation above, Darwin suggests a possible mechanism to explain the diversity found in the design of signals: the influence of female choice on the characteristics of the signal. He noticed earlier in his work that people could effectively change the appearance of animals by selectively breeding them according to their preferences, and consecutively he wondered if female canaries would be able to practice the same sort of selective breeding, just by consistently picking males with a certain trait as partners (a 'trait' being either a visible property of the animal itself or of its display). The suggestion that female choice would exert such a great influence on evolution (and even human mate choice) was rejected furiously in the Victorian Age Darwin was living in, and only decades later mathematicians and biologist started to elaborate on the idea. In the seventies, different mechanisms
for female choice were proposed, at times referring to cheating males, then again to omniscient females. By now, mathematical and empirical evidence is gathered for partially honest signals from which females can assess male quality with certain reliability. As can be easily noticed from today’s fashion waves, preferences come and go without reason; a certain tie does not reflect male attractiveness. On the other hand, not everyone can afford an expensive golden chain. Therefore some traits could reveal something of interest to a female about a male. This thesis explores the idea of rhythm as an indicator of quality that might be used by females for partner choice.

Evidence from different disciplines points towards this line of research. To start with, rhythm seems to be used as a signal of an individual’s underlying traits in humans. Karl Grammer and his colleagues have been studying the traits that human rhythmic behavior may be able to indicate. In their studies, people were brought to the lab and asked to “dance to their own rhythm”. The participants’ movements were filmed and analyzed using neural networks which came up with surprisingly accurate estimates of the dancers’ personality traits according to the standard “Big Five” dimensions.1 At least for humans, rhythmic motion may indicate factors that are important in mate choice, showing its usefulness in courtship displays.

Further hints for the mate quality signaling function of rhythm can be found in the static visual domain. Regularly repeated forms that can easily be compared with each other can be used as quality indicators: Stripes and other regular patterns on fish, insects, snakes, and other animals might make potential errors stemming from developmental noise more salient. As Geoffrey Miller (2000) put it, “From the viewpoint of signaling theory, repetitions across space (bilateral symmetry, radical symmetry, stripes) and across time (rhythm, repetition) are efficient ways to indicate developmental stability, a major component of fitness” (p. 67). Moreover, in one cricket species the female is known to nibble the equally spaced files from the hind legs of the male during mating, thereby preventing the non-virgin male from producing a regular stridulation pattern ever after his first copulation (Sakaluk and Ivy, 1999).

Following the same line of thought, developmental instability on the neural level, as opposed to bodily irregularities, could lead to disruptions in motor control and generation of behavioral patterns. Instead of revealing virginity or developmental noise, rhythmic displays in the temporal domain therefore could heighten the salience of neural noise or disorders.

---
1.1 Research question

In this thesis, I will explore the function of rhythmic behavior as a proximate cue of underlying mate quality or fitness, through the constructing of both an evolutionary and an optimization model. Through the construction and analysis of both models, I will investigate the constraints and prerequisites that might have influenced the design of one of the best studied signaling systems in biology: bird song.

1.2 Description of method

The choice for an evolutionary approach is based on the nature of the problem. In communication, or more generally, signaling, the design of the signal and sender/receiver are mutually influencing each other over evolutionary time. As we are looking for the function of one aspect of the design, we need to describe the possible influences from the psychological/physiological apparatus of the signaling individuals before the aspect under investigation was manifest, and then compare the situation with one in which the signal has actually taken this form. Would there be any advantages for a species in which rhythmic signaling is present? Will these features arise over evolutionary time, that is, would it be advantageous on the individual level? These kinds of questions can be addressed by means of an evolutionary model. We will define ‘rhythmic behavior’ as regularly repeated behavior and focus on the temporal domain for testing the more broadly applicable ideas about signal regularity. Rhythm can be measured in the model by counting of how many repetitive phrases the song consists.

From all auditory signaling systems we chose birdsong as our domain for two main reasons. First, birdsong is used as a courtship signal (among other functions), indicating its possible quality-revealing function. In many bird species, sexual selection has resulted in complex, elaborated songs with rhythmic elements (Searcy and Yasukawa, 1996). Second, there is evidence that rhythm in birdsong is disrupted by low-quality aspects of an individual singer (Ballintijn and Ten Cate, 1999). On the one hand, birdsong production depends on features of the brain that easily break down under developmental stress or poor nutrition, thereby making song a revealing handicap (Nowicki, Searcy, and Peters, 2002a). On the other hand, birdsong can also rely on capacities that are condition-dependent: The duration of songs, for instance, depends on energy level, and song repertoire size depends on age and memory skills (Gil and Gahr, 2002). Furthermore, the cognitive systems for song production are extensively studied in songbirds and at least for one species (European starlings), it is known that they are able to actively discriminate rhythmic patterns from arrhythmic patterns (Hulse, Humpal, and Cynx, 1984).
1.3 Position within Cognitive Artificial Intelligence

Research with the label ‘AI’ can be divided into two categories: either, it can be described as modeling intelligence using computers, or on the other hand, as the use of natural computation as an inspiring model for quick computation or decision making in general. This project clearly belongs to the first category. It describes a possible way in which evolution might account for quick decision-making in a relatively simple cognitive system. Moreover, it shows for one very specific kind of behavior how it may be initiated. Constructing an evolutionary model forces a very ‘behaviorist’ view on the subject, because it simplifies the cognitive framework to a great extend. By doing so, however, inferences can be made about the tasks and constraints that might be placed on the brain earlier on in the evolutionary history. Indeed, as Miller (2001) suggests, our language system can be seen as another ingenious and fragile system that breaks down easily under stress, and therefore might have been evolved as an indicator in humans.

Modeling intelligence invokes divergent disciplines. Modeling signal evolution does so too: Recognizing the factors that influenced the design of signals is an interdisciplinarian venture by definition. Since Tinbergen (1963), animal behavior research is divided into four categories, namely questions about mechanism, phylogeny, function, and ontogeny. In this thesis, I will address the different questions with viewpoints from different disciplines. To start with, the question about mechanism (how the signaling system is physically realized) will be addressed following merely a biological/neurophysiological viewpoint. How are rhythmic structures represented in the bird's brain, in such a way that the timing is reflected by the utterances? The ontogenetic question is only indirectly addressed; developmental stress is one of the factors that affects the quality of the signal, but how exactly lies beyond the scope of the project. We focus on the phylogenetic approach, concerning the evolutionary history of the trait under study. Because behavior has fossilized notoriously seldom, other disciplines than paleontology must be called in. Evolutionary modeling is the method chosen in this thesis, depending on informatics. Lastly, for functional answers about rhythmicity the focus shifts to cognition, theoretical biology and communication theory. All disciplines involved here overlap to some extends with cognitive artificial intelligence.

1.4 Status quaestionis

This study thus combines three fields: modeling (evolutionary modeling), theoretical biology (signal theory and evolution theory) and biology (birds and bird song). For each of the specialties, it holds that great progress has been made recently. Although only few theoretical studies from the first two
fields actually have had impact on empirical work, the ones that did, changed the field (see for instance Grafen, 1990a).

One of the big changes in the last decades was the branching from biology into different participating fields. Signal evolution itself received attention from ecologists, mathematicians, linguists and neuroscientists. Recently, signal evolution in the visual domain received a lot of attention by modelers. However, former approaches to signal design have typically focused on aspects other than regularity or rhythmicity, with a few exceptions (Enquist and Arak, 1994). Johnstone (1994) used neural networks to investigate a universal sensory bias for symmetry, proposing that regularity arose as a by-product of the need to recognize objects irrespective of their position (but see Bullock and Cliff (1997) for problems with this and the former approach). In addition to the sensory mechanisms of the receiver, another pressure acting on the signal design is the environment. As Enquist and Arak (1998) state, signals might evolve to become more redundant (which can be instantiated as rhythmicity) and therefore easier to discriminate from a noisy background (p.26). Their appearance may become more 'digital' as some components disappear during evolution, whereas others become more pronounced. The line of reasoning we advocate by linking quality to rhythmicity fits with the great amount of research in the past fifteen years devoted to finding the correlation between irregularity and developmental stress or genetic imperfection (Møller and Swaddle, 1997).

1.5 Overview of the structure

The second chapter discusses evolutionary models and their use in the behavioral sciences. Some evolutionary models are treated in more detail, to be able to address the question which principles might be helpful for the model under construction. Then, some theory about signaling systems and the design of signals is reviewed, along with the most influential models in the respective areas. The third chapter covers the basic aspects of bird song and its production. The important parameters for the evolutionary simulation are regarded, together with some features of birdsong that are especially hard to produce, and therefore may function as ‘honest signals’ of male quality. The neural system behind birdsong is one of the best studied neural circuitries, but still some controversies exist about the temporal representation of songs, e.g. the ‘timing’ of the notes. Indeed birds seem to suffer from irregularities in timing; therefore, the last section is devoted to the various sorts of mistakes that ‘low-quality’ birds make.

Chapter 4 deals with the model itself, the assumptions on which it is based, the simplifications made and their justifications. The second half of the chapter describes the effects different parameters have on the outcomes.

Finally, in chapter 5 the results are presented. An analysis of the female preferences according to the evolutionary model is performed, to test the assumption that rhythmic songs are more discriminating
with regard to the sorts of mistakes that male birds make. Then, the stability of the outcomes of the model is tested.

In chapter 6, I discuss the results and limitations of the simulation in the light of the theoretical background. I point out the overlaps and gaps between the predictions of the model and the ethological data and reflect on possible explanations for them. To conclude, some promising lines of research elaborating on the model are discussed.
2 Evolutionary simulation

Models can be appreciated as visualized thought experiments. However, with visualization comes simplification. This chapter argues that this major pitfall of models is not a fundamental one, but that it must be treated with caution. In the first section, two other conceptual objections against an evolutionary approach are addressed. The validity and use of evolutionary models for questions on signal design will be explored, in order to rephrase our research question in terms of models. The second part of the chapter reviews existing theories and models on signal design. The last section comprises the different approaches and sets the borders for our model.

2.1 Validity of models

The validity of evolutionary modeling depends at least partly on the assumptions of evolutionary psychology. Evolutionary psychology is based on the idea that the mind, at least partly, has evolved to become adapted to the environment (for an introduction, see Cosmides and Tooby, 1997). From this -disputed- assumption, some implications as well as shortcomings arise immediately with respect to the validity of evolutionary modeling of animal behavior. The evolutionary approach namely provides us with a behaviorist view on the subject. Natural selection is based on the selection of the most appropriate behavior in a certain environment. By acting as an optimizer of the behavior, every adaptation that is preserved serves a function. Exactly this narrows the scope of evolutionary models down to phenotypes: the same behavior can be generated in different ways, and the other way around, the same neural structure could account for different output patterns. For this reason, a classic evolutionary model has nothing to offer to the understanding of underlying cognitive structures, as Bolhuis and Macphail (2001) stress. They state that the (evolutionary) function of a particular behavior can never explain the underlying causal mechanisms. Referring to Tinbergen's (1963) typology, Bolhuis and Macphail concur that the function and evolution of signal form are intertwined tightly. Function-related approaches to behavior such as evolutionary models can, on the contrary, never serve as answers to proximate (as opposed to ultimate) questions about mechanism. In other words, an evolutionary model cannot shed light on the processes involved in producing a specific behavior.

Another less fundamental but very impinging implication attendant on evolutionary psychology approach is the presupposition of modularity. For reasons of computational efficiency, evolutionary psychologists claim, every distinctly arising adaptive problem has its own solution: a specialized adaptation. Therefore, the only assumption attached to evolutionary modeling about underlying cognitive structures is their modular nature. For many cognitive functions, this appears to be true
(Cosmides and Tooby, 1997); however, for the main issues in cognitive science, like long-term memory, intelligence, and decision-making, it remains unsolved whether modularity applies to it. Despite this controversy, modularity is one of the main principles of evolutionary psychology and can thus be felt in the modeling approach based on it.

So far, two objections to evolutionary psychology have been raised, the inaccessibility of underlying mechanisms and the claim of modularity. What implications do these two objections have for the construction of an evolutionary model of an indicator? Questions about the function of an indicator cannot be answered without making inferences about the (cognitive) mechanisms involved in displaying the indicator. The connection between the indicator and the corresponding trait exists because of a lower-level mechanism upon which they both draw. The mechanism needs to be taken into account, or at least the assumptions about it need to be specified to a large degree, even for modeling the appearance of the indicator only. There appears to be a conflict between evolutionary models and questions concerning functionality only where a revealing indicator is concerned.

Concerning the second objection, modularity, the functional question asked also seems to clash with the evolutionary approach taken; evolutionary theory as defined above is based on modularity, whereas we want to address the question about the function of rhythm in terms of general viability. As Miller (2000) points out, indicators cannot be modular in the underlying physiological sense. The indicator might be modular in appearance, if the underlying trait however were modular in the way meant by Fodor (1983), that is, at all levels of description, how could it serve as an indicator for general viability? General fitness is the sum of an animal’s various adaptations, therefore indicators of very specialized brain or body functions can only provide limited information. Indeed, all general viability indicators identified so far tap into low-level cognitive functioning, though they might be highly specialized in appearance (Miller, 2000).

This problem seems impossible to solve with an evolutionary approach. A closer look at the nature of the problem however solves the contradiction. Shifting the paradigm from natural selection to sexual selection overcomes both the demand of modularity as well as the inaccessibility of underlying mechanisms. Natural selection might be concerned with phenotypes only, sexual selection works on phenotypes that reveal something about general abilities or even genotypic properties. The classic example is the peacock’s tail: Why would a female peacock base her choice on a useless, but elaborated trait, when it did not reflect the developmental process of the body?

Sexual selection based on indicators works because of the relationship between indicator and general fitness. The mechanism behind the indicator functions as an in-between, but is based on the same low-level processes as the general fitness, and therefore can be addressed (or inferred) by selecting on general fitness in this case.
To answer the question whether evolutionary models are valid tools to study processes that underlie animal behavior, we can conclude that in the case of sexual selection, it might be valid to make inferences about underlying cognitive mechanisms; at least, models based on sexual selection do not suffer from the same constraints as normal evolutionary models do.

### 2.2 What can be concluded from models?

Having stated that some evolutionary models can be valid tools to study animal behavior, still the question remains how constructing these models could be useful. The greatest difference between real time evolution and the computer-based version is in itself the answer to the question about use of simulations: we can watch simulated evolution. Moreover, we can start it over multiple times, in order to try out slightly different assumptions, and look at the effect over generations. Maybe the biggest flaw of evolutionary models is their intricacy. As Bullock (1997) states: ‘…since evolutionary models are tools, they must have tasks’ (p. 149). In general, models serving a scientific rather than a divertive purpose either test the coherence of ideas, or provide testable ideas themselves. Below a classification of possible tasks is given.

#### 2.2.1 Classification of tasks

Perhaps the purpose most widely strived for by means of an evolutionary model is the existence proof or ‘proof of concept’. Showing that something can evolve in principle often comes down to proving of one of the three emergence principles (Miller 2001, p. 163):

1. Many quantitative mutations can form a qualitative change. This is nicely illustrated by Dawkins' simulation of the evolution of an eye (see Todd, 1996): cumulative small changes can result in qualitatively new adaptations.
2. Complex functioning systems, existing of many loose particles, can arise through small steps. Cooperation between individuals forms an example of a class of models illustrating this principle. Evolutionary models of cooperation tend to emphasize the distinctness of the different individuals: for instance Han Young and Miikkulainen (2001), who developed cooperating agents without communication.
3. The innovation buffer can be surmounted (Miller, 2001). Small genetic steps do not yield fitness advantages immediately (they often entail costs), but are needed successively for a qualitative difference. Recently, evidence was found that in stick insects, wings have developed and disappeared again at least four times during the evolutionary course (Whiting, 2003). This was taken for impossible for a long time (how could such a complicated trait develop again and again, with no
advantages for a rudimentary wing?), but had been proposed by more theoretic accounts for evolution (Kimura, 1983).

A shortcoming of ‘proof of principle’ models is that they often fail to connect the existing world with the artificial world, except for the concept proven. Instead of focusing on a plausible constellation of parameters, all assumptions are subordinate to the outcome of the model; this way, as noticed in the 6th European Conference on A-life (Wheeler, Bullock, Di Paolo, et al., 2002), the risk of ‘WYWIWYG’ (what-you-want-is-what-you-get) lurks.

Another, less haughty direction to take is to describe a possible course of evolution (see for instance Todd and Miller, 1991a,b). Models can generate new hypotheses that could in turn be confirmed experimentally. An indirect effect of the modeling process is that all assumptions are explicit, which might be a huge advantage once the hypotheses are actually being tested.

Effects of manipulations can be tested by observing emergent data: Special measurement algorithms can be used as the simulation progresses, to gain insight in the speed of convergence, the amount of evolutionary change, or even how likely some trait is to emerge, given a certain starting point (Cliff and Miller, 1995).

Lastly, experimentally derived data can be obtained. Probably the most interesting use of a model is to investigate what happens when certain parameters are changed. What are the important assumptions and between which borders do certain effects occur? Loosening assumptions, one at the time, is a good way to find the crucial factors. Comparing the final offspring with the start-off-population, for instance, can give insights in the evolutionary process, from which inferences can be drawn about the evolution of existing populations (Houde, 1993). Questions about the stability of a certain trait or the history of emergence can be addressed. The influence of the environment on an individual can be compared with the effects stemming from individually defined parameters on the social behavior of the group.

The most important flaw of evolutionary modeling is that it cannot prove whether the behavior really evolved the way the model proposes. The necessary simplifications and constraints prevent a model from proving the course of evolution; since we know only one ‘real’ instance of evolution, the difference between a plausible and an unavoidable outcome is hard to assess. However, the plausibility of an outcome given some hypothesis can be deducted. History matters in evolution. Therefore, models have to be constructed in consultation with the real world of biological empirical fieldwork, mutually feeding each other new hypotheses.
2.3 Evolutionary models

With respect to the research question on the evolution of a certain signal form, we can discriminate between different classes of evolutionary models. The sort of view they provide us with depends on the class chosen. Therefore, the most suitable models are addressed in detail below.

2.3.1 Basic genetic algorithm

The standard genetic algorithm (Holland, 1975) is based on the four prerequisites for adaptation (Bullock, 1997): reproduction, inheritance, variability, and competition for limited resources. All individuals in a population compete for survival and reproduction. The fitter, more adapted ones manage to reproduce more often. Therefore, in the next generation the good genes will spread, and might be recombined to form even more adapted individuals; over generations, they will cluster around some of the optimal fitness values. GAs can be used for search and optimization problems by clever coding of their genes.

Their biggest shortcoming, aside from their dependency on intelligent coding of the genotype-phenotype mapping, is the static fitness landscape. As the population changes, still one signal or strategy is the winning solution. However, the problem we want to address is intrinsically influenced by its environment; not only that, it depends heavily on its conspecifics and their behavior. As Lewontin (1979) acknowledges, environments might shape organisms, but organisms construct their environment as well.

2.3.2 Lamarckian cultural evolution

To survive in a changing environment, an organism or species needs to adapt. On the level of the organism itself, adaptation can be accomplished by learning.

Lamarck, the earliest proposer of an evolutionary theory, advocated at the beginning of the 19th century (‘Inheritance of Acquired Mechanisms’, 1809) that knowledge acquired during life could be inherited directly in the next generation. If this were true, science would have reached a much higher level of sophistication by now; however, Lamarck’s wrong assumption still proves to be useful. Because it acts on the behavioral level instead of on the genetic level, Lamarckian cultural evolution can proceed much faster than natural evolution, and can therefore be of use for search algorithms. Behavioral plasticity might play a role in evolution after all, but the pure Lamarckian effect has been disproved long ago.

2.3.3 Baldwin effect

Approximately 100 years after Lamarck, Baldwin (1896) proposed that learning on individual level might have an effect on learning on population level, but only in an indirect way.
First, if some information is of importance for the individual to survive, the individual might learn it. For instance, if a constant in the environment has changed so quickly that ‘normal’ evolution cannot get a grip on it, the new inference might be picked up by one individual with plasticity in behavior. The agent without the ability to learn has a disadvantage over the smart agents, and eventually, the population will be filled with learning individuals. Baldwin proposed that the trait for learning could gradually be replaced by the same hard-wired information (instincts), if there are some costs attached to learning. In other words, learning smooths evolution and it might help to overcome the innovation barrier. Hinton and Nowlan (1987) designed an influential formal model to prove this concept. A possible example of this mechanism is the well-known invention of sparrows: one found out how to open a milk bottle. Within a few years, this qualitatively new behavior had spread throughout the international sparrow society. Such an advantageous behavior might end up in the genes.

On the contrary, the Cultural Trap Hypothesis formulates a situation where learning can be stable even though it might not lead to a higher fitness on individual or population level. This is due to the interaction between culture and genes. See Lachlan and Slater (1999) for a model on vocal learning in birds; in their model, song learning is maintained in a population due to an evolutionary trap, even though learning might not be beneficial at all on the level of group selection.

2.3.4 Coevolution

As mentioned above, the biggest flaw of evolutionary modeling is the fitness function that must be imposed by the modeler. In co-evolutionary models however, this problem can be circumvented: the two species that evolve together might influence and assess each other’s fitness, mutually judging each other. This way, the fitness depends on the course of the model itself. Progress is not easy to measure by plotting fitness over generations, for one specified phenotype might have a completely different fitness value five generations later. The influenza virus for instance mutates notoriously frequently, thus leaving last year’s medicine without purpose. One problem with coevolutionary models is that they influence the influence, and it is hard to keep track on who started whom. More sophisticated measures must be used to plot progress in coevolutionary systems.

2.3.5 Arms race

A special instance of coevolution can occur when two species with conflicting interests evolve. An arms race is an example of a ‘Tracking the Red Queen’-situation (Cliff and Miller, 1995): both species have to keep up with each other in developing new tools, because the species lagging behind will become extinct. In other words, they have to keep moving in order to stay the same, just as the running Red Queen in Alice in Wonderland. Examples of arms races are found throughout the animal and human world. One example referring to signaling is found in Krebs and Dawkins (1984), who propose that a sender and a receiver could have conflicting interests as well. For the receiver, it pays to gather as much knowledge about the sender as possible, whereas it might be advantageous for the
receiver to hide his feelings or fitness. This way, the sensory mechanisms involved in sending and receiving signals might develop mutually influencing each other, in an everlasting loop. This might apply to sexual selection, the instance of genetic algorithms we want to use.

2.3.6 Sexual selection

Sexual selection is in turn a special instance of coevolution of individuals with, counter-intuitively, often conflicting interests, and can therefore result in an arms race. The process has been overlooked largely since Darwin proposed it. In modeling terms, sexual selection is a form of coevolution where nothing else but attractiveness defines the fitness function of an individual. The shift in focus as opposed to natural selection has various consequences for the direction and course evolution takes. Whereas natural selection acts as a hill climber, with all individuals searching the adaptive space and gathering at the highest peak, sexual selection merely explores the landscape by escaping from local peaks and even by creating new fitness peaks (Miller and Todd, 1995). The combination of the two principles can, on the one hand, guide evolution towards peaks in the fitness landscape that natural evolution would never have reached on its own; on the other hand, directional sexual selection has a stabilizing effect, when females consistently prefer males that have a high fitness in the adaptive landscape.

2.3.7 Assortative mating

The idea of like seeks like dates from ancient Greek times, and “Birds dwell with their kind” was quoted in the apocryphal book of Ecclesiasticus. The full saying in English, “Birds of a feather flock together” expresses an idea that can be rephrased in evolutionary terms as a special case of non-random mate selection. Assortative mating can take either positive or negative form, that is, mates with an allele (genetic information carrier) similar to the choosing individual are preferred over others, or the other way around. In humans, positive assortative mating preferences have been asserted for level of education and skin color. An example of negative assortative mating was found in women’s preferences, mediated by pheromones, for partners with an immune system complementary to their own (Wedekind, Seebeck, Bettens, et al., 1995). One important effect of non-random mating is that it counterbalances the negative mutation bias; another possible effect is speciation, as emphasized by Todd and Miller (1997).

2.4 Modeling signal design

A big simplification can be made at this stage by eliminating all evolutionary processes that take learning into account. Obviously, learning plays an important role in bird song. A striking every-day example is the ring tone imitation of sparrows: evolution could not possibly keep up with the changing fashion of mobile phone ring tones. However, the basic concern of the model is the design
of the signal, and therefore the learning phase can be easily disregarded, taking up Ockham’s razor and regard a song phrase learned from parents or tutors simply as heritable.

More important choices can be made now. Sexual selection has a role in signal design when indicators are involved, so at least our fitness function depends on other individuals. The evolutionary process may take the form of an arms race, or assortative mating.

As we can now think about the research question in terms of models, two formulations come up:

*How could female preferences give rise to a rhythmic signal as an honest expression of male quality?*

*More generally: Which principle of signal evolution could account for this very phenomenon?*

Darwin thought honesty was the reward for itself. However, from game theory we know that individuals might do better not telling the truth. Still we find many situations in nature where a signal reveals something about the sender, whether voluntarily or not. The demand for honesty has its influences on the design of the signal. Therefore, in the next section a review of strategic design of signals is presented, followed by the efficacy constraints on signals In addition, some important simulation models of communication are reviewed to explore different ways of implementing signal honesty.

### 2.5 The view from the receiver: strategic design of signals

In the former section, it was emphasized that sexual selection has different sources and effects and selects for different traits than normal evolution does. Sexual selection acts on perceivable traits as opposed to practical adaptations. Another word for ‘perceivable trait’ is signal: although it is non-trivial whether all signals are consciously perceived, they are perceivable by definition. Krebs and Dawkins (1984) defined a signal as an action or structure that increases the fitness of the sender by altering the behavior of the receivers, but the authors were criticized because they underestimated the role of the receiver. Receivers, as Bullock (1997) points out, will not increase their survival chances if they believe everything that is told; rather, they quickly try to extract as much information as necessary from a signal, which might be something the sender does not want them to read. So a signal can only exert its strategic influence if the receiver increases its own fitness by altering its behavior according to the signal; that is, there must be a strategic advantage for the receiver in listening.

Nature has found different ways to impose these constraints on signals, which are discussed in detail below.
2.5.1 Sexual selection and signal design

Darwinian evolution seems to advocate efficient, no-nonsense life forms. If some property or function turns out to have no use at all and lowers the viability of a creature, it will eventually disappear. So how could the very same process account for the striking colors in some prey, or for useless and extravagant behavior like ‘conspicuous consumption’ (Veblen, 1899)? Why are birds biding their time singing complex songs, instead of eating and mating?

Darwin proposed, after publishing his most famous work, other coexisting processes acting upon genetic diversity. Besides surviving, there is another task for a sexually reproducing animal to prevent his species from extinction: he has to ensure that his genes are passed on. Sexual selection in different manifestations places constraints on the direction of evolution. Sexy features can become emphasized, without further reason than looking sexy. On the other hand, features that have proven to be adaptive might be regarded as sexy. For this reason, the question of this paragraph is whether there is a difference between looking sexy and being sexy.

2.5.2 Runaway selection

One form of sexual selection, called direct sexual selection, is widely accepted as having a big influence on evolution. The word direct points to the direct benefit the female gets by preferring mates with a particular trait. She might prefer males with a large song repertoire, when they happen to have larger territories as well. The female with the preference will then get more food and healthier offspring then her less dainty neighbor. The process of indirect selection however, where there is nothing in it for the female but genes, remains disputed until now (for an overview, see Andersson 1994).

For a long period, the argument against indirect sexual selection was that it remained unclear how the female preference for a feature originated. It seemed to be not very beneficial for a female to be choosy and stand the risk of not mating at all, if she receives nothing in return for her patience.

Runaway selection, first proposed by Fisher (1930) and formalized by Kirkpatrick (1982), explains one possible way for a preference to develop without a ‘reason’. It is based on the exaggeration of slight variations in a genetic trait such as the size or color of a spot. A female, using her sensory organs for more than discriminating among males, might happen to have a sensory bias that leads to preferring some male trait (for example, the trait being a spot with an ‘appealing’, food resembling color). By choosing mates with the trait, she will get sons having this sexy phenotype property, who are likely to have more offspring themselves. In addition, the choosy female will get daughters with the same preference trait, and grandsons with the male trait, and so on. This way, the preference and the preferred trait become linked and exaggerated by a self-reinforcing process. This process goes on until the exaggerated trait places a burden on the male, thus becoming a costly feature to have.
Central to this notion of sexual evolution is that the traits are initially arbitrary with respect to the male’s viability. There is no pre-existing link between the male’s fitness and the preferred trait, and except for his “street value”, i.e. his attractiveness to other females, the trait does not revail anything to the female about the male’s condition.

2.5.3 Good genes model

On the contrary, in good genes models, the preferred trait is supposed to be an honest reflection of some aspect of male quality. The evolutionary advantage gained by the female for picking a male with the preferred trait is that her children get not only the trait, but also the underlying higher viability.

While the Runaway selection models were faced with the start-off problem of the process, for the good genes models, the major problem is how to maintain the genetic diversity and stop it from eroding, thereby reducing the value of the signal. If in the end all males possess the phenotype property (but not necessarily the associated viability), there is no point any more in preferring it, since the link between the good genes and the conspicuous property is gone. Several models try to explain the abundance of signals in nature in terms of signal honesty.

2.5.4 Handicap principle

Concerning signal honesty, Krebs and Dawkins (1984) discriminate between signals that provide a mutual advantage for both signaler and receiver, and signals between animals of conflicting interests. Obviously, the latter is more vulnerable for evolutionary erosion. Krebs and Dawkins call it manipulation instead of signaling. Counter-intuitively, predator and prey are regarded both as gaining from honest signaling, whereas two potential sexual partners are often regarded as having (evolutionary) conflicting interests. An example of the former would be that strong legs suggest being hard to catch. This can prevent both predator and prey from starting an unavailing and exhaustive chase. If there is a mutual advantage in it for both signaler and receiver, a signal could evolve that reflects the state of the sender. However, if there are no costs associated with this signal, abuse by weaker animals lies just around the corner. More precise, if there are conflicting interests between animals of different sexes, it can be useful for the sender of the signal to exaggerate his quality. In this way the ‘bluffs’ (for a game-theoretic model, see Hasson, 1994) can get all the benefits of the sexy signals, without having the underlying good genes. This results in an unreliable signal, which is of no use for signaler or receiver.

To address this unlikely outcome, Zahavi (1975) proposed the handicap principle. According to this, signals are costly to produce or have a hindering effect on the animal’s fitness in general, thus ensuring that only the ones with a high quality can carry the weight of having it. Taking this literally,
the weight of a long tail slows the animal down, and shows that one can cope with it, escaping predators despite of it. It shows that you are fit enough to stay alive in spite of the handicap.

**GAME THEORY AND THE HANDICAP PRINCIPLE**

Numerous attempts were undertaken to prove or disprove the handicap principle by constructing models for it, both evolutionary as well as game-theoretical models. First, the population-genetic modelers seemed to gather evidence against the handicap principle [for instance Maynard Smith (1976) and Kirkpatrick (1986)], but starting from 1990 with Grafen’s model, some mathematical support was found. Grafen (1990a,b) was among the first to notice that signaling was a communication game and that game-theoretic principles would have a great descriptive power in this field.

In the game theoretic approach, a central notion is Evolutionary Stable Strategy (ESS). This is the applied version of a Nash equilibrium, in which for all players, none will do any better by changing his strategy. The stable strategy used by all players cannot be invaded by another, weaker strategy over generations. This class of models can be applied to proof of principle-tasks. Postulating a concise definition of the handicap principle however took 15 years. As it turned out, Zahavi’s original postulated principle consisted of at least three different principles (Iwasa, Pomiankowski, and Nee, 1991).

Grafen (1990a,b) set the first bounds to signaling games involved in sexual selection, and found that two ESSs can exist: One in which there is no signaling at all, that is, the female takes the first male she encounters, and another strategy that is stable if and only if the cost for the signal is bigger for low quality males than for high quality males. This finding has been refined, broadened and adjusted over time (see Harris and Bullock (2002) for an example), but remained the central notion in handicap models.

**References:**


Krebs, As

2.6 The view from the signaler: efficacy considerations on signal design

Iwasa et al. (1991) unraveled the blurred notion from Zahavi through the use of quantitative genetic models. They argued that part of the initial controversy was caused by the confusion of three different handicap principles. All of them are based on the idea that a handicap imposes a larger cost on a weaker animal than on an animal with high viability. Singing long songs is more costly for birds with a small territory, because they have to spend more time on foraging than their more fortunate neighbors; therefore, song length could form an honest indicator for territory size.

From the three mechanisms they distinguish in the article, the first is 'Zahavi’s model' or the pure epistasis model. This model expresses the idea that there is an epistatrical interaction between the general viability trait and the handicap trait. Thus, the trait places a bigger burden on a lower quality male than it does on a fitter individual.

The second type, known as conditional handicap, reveals the overall condition of the animal by imposing greater costs the larger the trait is. The genes that code for a good condition are of influence here, the bigger trait being merely a consequence of the actual health. (So the size of the expressed trait depends on the state the male is in; this might be an all-or-nothing signal or a gradual diminishing of the size.)

The revealing handicap, lastly, is one that indicates the male’s quality by the state the sexual trait is in. For instance, the size of a tail is encoded in the genes, but due to attacks of parasites (Zuk & Hamilton, 1982) or predators the tail and his owner lose their attractiveness. Songs even can get affected by parasites (see Buchanan, Catchpole, Lewis, et al., 1999).

In some cases, the Fisherian process and Zahavi’s principle will act together, mutually enhancing each other. A trait that honestly expresses viability can be a costly thing to produce and, over generations, be regarded as beautiful. Still the question remains how to overcome the problem of preventing all the males from having the same handicap and quality after a few generations. Andersson (1994) addresses this phenomenon referring to three biases: temporal variation in selection, spatial variation in selection, and a negative mutation pressure (the assumption that mutations have deleterious rather than beneficial effects on fitness), but the degree of honesty remains controversial for most signals, as well as the origin of the preference for it.

2.6 The view from the signaler: efficacy considerations on signal design

As established by now, a signal influences by definition the behavior of the receiver (Dawkins and Krebs, 1978). It will do so in a way that is beneficial to the sender; otherwise, he (the sender) would keep still. The ability to signal effectively, whether honestly or not, therefore changes the survival and
mating perspectives of the sender. Note that within this terminology the notion of manipulation is intrinsically hidden.

However, in this arms race between manipulative signalers and skeptical receivers, as Noble (1998) puts it, the form of the signal is what concerns us. The underlying strategic game as explained above is necessary, but never sufficient in explaining the scattered diversity in signal form and design. For a species with a certain neurological structure that lives in an environment with certain properties, some signal forms are more effective and more likely to develop than others. The signal itself is subject to evolutionary forces, thereby undergoing changes in form, content, and function.

The first known example of a changing environment having its influence on signal design is the famous peppered moths story. In London, the “peppered” moths turned black as a reaction to the darkened trees. Clearly this was a result of the advance of having a protective color. However, the famous example for evolution turned out to be not verifiable (Coyne 1998), which elicited the following statement from the director of the Institute for Creation Research, John Morris: “What a wonderful time to be a creationist, when even the supposed best proof of evolution in action is so flimsy that it cannot stand the test of truth.”

Not only the physical characteristics of the environment influence the efficacy design, also the psychological properties of the senders are part of the environment. Again, we can distinguish between how the signal and the preference are initially evolved, and how they remain noticeable over time. According to Bradbury and Vehrencamp (1998), the evolution of a signal starts with the association between the signal and the condition in the sender. For instance, a sound with low frequency is in general produced by a bigger congenor. In some cases, as Ryan (1990) found, the female preference existed prior to the male trait exploiting it. Arak and Enquist (1993) hypothesized a ‘hidden preference’ in those cases, where the preferred trait lies beyond the range of male traits extant.

Guilford and Dawkins (1991) describe the process of signal evolution differently: they argue that conventional signals occur first, without any link between information and content, only relying on convention (for instance the tie-example from the introduction). They undermine their own honesty and therefore the only signals that are preserved, became reliably connected to some non-fakeble trait during self-enforcing evolution (expensive tie-pins).

Often it works both ways: the form of the signal changes in such a way that the message contains more differentiated information, but in order to perceive this, the receiver has to change his mechanisms for denoting the message too.

---

2 http://www.icr.org/index.html
2.7 Modeling rhythm

As emphasized in the introduction, models on rhythmicity typically focus on the latter of the two factors described above, by evolving signals towards discriminability of the background (effectiveness), or implementing sensory biases and investigating the effects (receiver psychology from the view of the signaler). Models on handicaps and honest signals also rarely overlap with signal design and cannot help us in this respect. Below two relevant models on signal design are discussed, before I close in on the model under construction; though neither of the past approaches applies to our research question, it is useful to consider different approaches first.

The approach Kenward, Wachtmeister, Ghirlanda, and Enquist (2002) take is directed towards repetition itself: they search the literature for possible mechanisms that might have caused spatial repetition to evolve, and test them by co-evolving signal and receiver. The receiver (consisting of a simple feed-forward neural network) essentially evolves to solve a discrimination problem, and the signal evolves towards an efficient stimulation of the network. Kenward et al. manipulate the receiver input by inverting and mirroring the signal, which indeed leads to a symmetrical signal design. This is not the approach we want to pursue, while this class of models fails to address the link between quality and rhythmicity. However, Kenward et al. used a workable definition of rhythm and counted the number of repeated sequences of a certain length (but excluded the cases where the length of the repeated sequence is one). A repetitive pattern is thus composed of a ‘generator’, the reoccurring sub-pattern, and ‘rhythm’ is the number of occurrences times the length of the repeated sequence.

The class of models that is not directed to the stability of a signaling system but directly to the design of signals typically focused on aspects other than regularity or rhythmicity, with a few exceptions, as noted in the introduction. A coevolutionary modeling approach to signal design was used by Werner and Todd (1997) to explore signal design in birds in terms of novelty rather than rhythm. They emphasized neophilia and an evolutionary pressure towards constantly changing signals, as opposed to a female preference for regularity. In their model, females used inherited song preferences along with a “desire for surprise” to select males who evolved over time to perform a wide and ever-changing diversity of songs. This result is complementary to the approach taken here, because the pressure for novelty must be balanced by a pressure for regularity to first form expectations that can then be violated.

What can we learn from former modeling approaches to signal design and to what use are evolutionary models to our research question?
The goal set to the model was to verify whether rhythmicity could evolve as indicator and under what circumstances and parameters this effect can be obtained. Thus, the goal of the model is to describe a possible course of evolution. Since this refers to experimentally derived data, the model has to be robust and simple, in order to be able to adjust or add parameters and still sort clear effects. The useful class of models to consider comprises coevolutionary models with sexual selection, because for an indicator to develop, sexually reproducing individuals are needed. A special property of the individuals, rhythmicity, must reflect overall quality (to be defined later) in the end of the evolutionary run. However, by linking quality to rhythmicity prematurely we would not discover anything about the connection. Therefore song appearance, or structure, cannot be directly linked to quality. On the other hand, if runaway selection would be the only constituting mechanism, rhythmicity could not possibly become linked to quality at all. The theoretical idea we have to apply to the model is the notion of a handicap.

Through game-theoretical modeling, a formal description of the set of conditions for a stable signal was found: rhythmic songs must be harder to produce for lower quality birds for the model to work. Another demand is that the choosing sex must be able to discriminate between mates of different quality, and thus between mates of different rhythmicity.

Now that the mechanisms are addressed that could account for signals as honest indicators of quality, we have to rephrase the question into biological terms. Searching for honest indicators starts with finding a trait, or song type, that is harder to produce for birds of lower quality. That is, it must reveal the general viability, either developmental or heritable. To model such a trait, we need plausible parameters on the bird's life as well: the number of offspring, the degree of heritability, and last but not least: whether some aspect of rhythm indeed reveals something about the sender.
3 Birdsong

As emphasized in chapter 2, biological models can only serve a scientific purpose if they are based on biological data. This chapter provides an overview of existing knowledge and theories about function, design and role of acoustical sexual signaling in birds. In order to keep the model realistic, we need to take into consideration certain parameters of the environmental conditions: facts like the average number of offspring, but also mere social variables matter, like the degree of monogamy, the amount of parental care, etc. The second part of this chapter comprises biological reviews on mate preferences, the function of song and different song features. Lastly, the bird’s cognitive structures and internal parameters should be taken into account as to which simplifications lie within an acceptable range. To what degree are song patterns heritable or learned? What is the nature of the mapping function between brain structures and song patterns? How much do birds differ among themselves with regard to the expression of songs? This chapter comprises all necessary information to base the model on.

3.1 A bird’s life

The average bird stems from a nest with four or five siblings. In most species, both father and mother take care of the nestlings; meanwhile, the chickens might get used to songs from their father and their direct neighbors. After having reached sexual maturity, the males have to occupy a territory and the females start searching for a good partner. In most avian species, the female is the choosy sex; her offspring is limited, whereas a male can pass on his genes whenever possible. The courting season does not last forever, therefore time and space are restricted.

90% of all bird species is predominantly monogamous. From the remaining, some are polygynous and the majority is promiscuous. This high percentage of monogamous birds is not coincidentally the same as the number of parental care giving birds. For altricial birds, it makes sense to choose a faithful partner who can help feeding. However, after DNA paternity analysis was an additional expedient, almost all species that were thought to be monogamous turn out to engage in Extra Pair Copulations (EPCs). In fact, the females use their selective power very calculating: they reject unattractive males, and only get involved in EPCs with males that are more attractive than their own mate (Houtman, 1992). Taken all together, it seems reasonable to conclude that an eligible male has more opportunities to mate than his less attractive rival has, even if the species is merely monogamous (Kirkpatrick, Price, and Arnold, 1990).
Once the female has made up her mind about a partner, they form a pair and start nestling. This process goes on for a few years with usually the same partner. Every year, the average female’s clutch consists of five eggs. The number of viable chickens, however, might vary along with the male’s (paternal) quality. A father who can provide his children with more parental care, like food and protection, will keep his offspring alive, thereby increasing his fecundity. For the female, it pays to find the male with the highest 'quality'. In most bird species, songbirds as well as non-songbirds, the female infers the quality of her potential partners from their songs.

**SONG STRUCTURE**

A birdsong is build up out of strophes (Bergmann and Helb, 1982). One strophe or song type continues for a few seconds, and can be composed of a sequence of one to twenty rapidly repeated phrases. Each phrase or ‘syllable’, which is one uninterrupted piece of sound, consists in turn of a few notes.

A common way to visualize songs is the sonogram, in which three attributes of the sound are depicted: frequency, amplitude and duration. On the x-axis, the duration is indicated, whereas the y-axis indicates the frequency of the note. The amplitude of the notes is reflected by the darkness in the sonogram.

![Sonogram of an adult male canary](From Tramontin and Brenowitz (2000), sonogram of an adult male canary)

Like every language, bird songs both have a phonology (the notes the song is composed of) and syntax (the ordering of the phrases). As we are primarily concerned with rhythm, the latter needs more attention.

Besides the songs, birds also produce calls. A call can be discriminated from a song by its length: a call is a simple, short utterance with a very specific function, for example begging, calling alarm or seeking contact. In this chapter, I will only refer to songs, as I am concerned with signals used for sexual selection and the reverse, the impact of sexual selection on the shape of birdsong.

Reference:
3.1.1 The function of song

The most important function of song, both for birds and for humans, is to attract mates. In most bird species only the male sings, though in some tropical passerine species the female duets with the males (Farabaugh, 1982). Among the passerines (songbirds), oscines have the most elaborated song. At least for them holds that it is not the feathers that make the bird; but also in non-songbirds, the utterances play an important role in male-male competition as well as in mate-attraction (Catchpole and Slater, 1995).

Evidence for the influence of song on mate choice stems from different levels. Birds sing most frequently in the courting season. Once they have found a mate, the song output declines, serving only for the defence of the territory. Other evidence that song serves as a courting signal is that the behaviour of the female influences male song and for many song features it holds that they are correlated with number of offspring (Searcy and Yasukawa, 1996). One striking illustration is that in the part of the male brain that is responsible for song, the Higher Vocal centre (HVc), new neurons sprout forth in the courting season, and disintegrate again in fall (Tramontin and Brenowitz, 2000). It is hard to imagine stronger evidence in favor of the sexual function of bird song.

In the following paragraph the biological constraints for the individual expression of bird song are reviewed, in order to find out which features are hard to produce, and therefore might reflect the singer’s quality. If a feature serves as ‘revealing property’ for the purpose of sexual selection, it will meet certain conditions: First of all, great variation between individuals in one population will exist with respect to the feature to select on. Then, the varying property must reflect or correlate with the overall quality (viability) or another important trait of the bird: for example, it suffices if the perceivable property breaks down easily under developmental stress. At least, in order to maintain honesty, the signal should place a burden on the sender. Lastly, the feature is likely (though not necessary) to be heritable, and the preference for the feature is often (culturally) heritable (see Riebel, 2000).

3.2 Constraints on song

In the past decade, ethologists, behavioral ecologists and neurobiologists have tried to inventory the factors that influence birdsong expression (for an overview, see Gil and Gahr, 2002). Because enormous differences exist between species as well as between individuals, numerous studies have been undertaken to explain them. In the beginning, emphasis was laid on the neurological burden and on the metabolic cost on performance. Which aspects of song expression are suffering from these
constraints? To start with, the performance-related traits, like the total time spent on singing. Besides, within the song, different notes can have varying degrees of difficulty: evidence is found for female preference for ‘sexy syllables’. This already suggests that temporal (respiratory) constraints play a role in the expression of the song, leading us in the direction of rhythm. Then, the variability (repertoire size) might place a neural cost on the singing male, or on the other hand, the stereotypy of the songs. After reviewing all these features that engross the brain or body, we can conclude whether rhythmicity serves a similar role, indicating the quality of a singer.

3.2.1 Performance-related traits

"The power and inclination to sing differ so greatly with birds that although the price of an ordinary male chaffinch is only sixpence, Mr. Weir saw one bird for which the bird-catcher asked three pounds; the test of a really good singer being that it will continue to sing whilst the cage is swung round the owner's head." (Darwin, The descent of man (1871), Chapter 13, section 14).

The performance-related traits (like longer or louder songs, or more time spent singing) affect merely the overall metabolic costs. Moreover, when a bird sings, it cannot eat. The bigger the territory, the easier it is to forage, so this might be a direct benefit for a female choosing males with a high song output. Indeed, as Hasselquist, Bensch, and Von Schwantz (1996) found, song output is correlated with food availability. Song output can be measured in different ways: for instance, Lambrechts and Dhondt (1986) found that the percentage of time spent on singing within a song varies a lot between individuals, and that it declines over time.

Forstmeier, Kempenaers, Meyer, and Leisler (2002) used another performance measure. They found that the amount of time that the song amplitude exceeds a certain threshold corresponds with longevity and male dominance. Above this, in extra-pair copulations, where no direct benefit for the female is present, males that scored high on this measure were chosen more often. Forstmeier also showed that this ‘average loudness’ is a heritable feature. Therefore, performance-related traits fulfil all conditions to function as indicator (heritability, attractiveness and a great variation among individuals), but have nothing to do with temporal rhythmicity.

3.2.2 Sexy syllables

Constraints on the temporal patterning of song components may be an honest reflection of respiratory or physiological constraints of the motor system (see Podos, 1996).

For instance, Vallet, Beme, and Kreutzer (1998) found that one very fast two-note syllable elicited a high level of sexual display in female canaries. They supposed that this high-frequency pattern would
be hard to produce. Therefore, it might provide reliable information to the female about the quality of the neuro-muscular connections needed to breathe.

In most animals, breathing is correlated with the movements of the muscles in the chest, which in turn is influenced by running or flying (Berger, Roy, and Hart, 1970). Until now, it is controversial whether this applies to bird song too; but certainly breathing speed could influence song output, and a fine-tuned neuro-muscular steering system must control both breathing and singing. This view is supported by the findings of Leitner, Voigt, and Gahr (2001), who found that the very same fast frequency syllables are sung more often in the breeding season. Perhaps the birds are exhibiting the precise coordination of their respiratory muscles, which need to be rhythmic in order to function optimal.

![Sonogram of a sexy 2-note syllable, from Vallet, Beme, and Kreutzer (1998)](image)

3.2.3 Repertoire size

Big between-species and inter-species differences exist again as it comes to the number of different syllables. For example, great tits have a mean song type number of three, whereas a robin’s repertoire consists typically of hundreds of different songs. Searcy (1992) stated that despite of these big differences, essentially in all species a large song repertoire is preferred. A first explanation is that a big repertoire prevents the females from becoming habituated to one song type. This was tested by Horn & Falls (1991). They found that high variability indeed elicits more courting behavior. If this were only a matter of sensory bias, selection on song repertoire size would be a runaway form of sexual selection. But the evidence points in a different direction.

Nowicki, Searcy and Peters (2002b) proposed that a large repertoire could reflect that enough food for the birds was present during the crucial learning period. Hungry chicks have more on their mind than choir practice, and well-fed birds make better fathers, as Adam (2000) put it. Another
explanation is that not song acquisition, but memory space is the bottleneck; then memory size, which is also a valuable daily life property, could be inferred directly from repertoire size. These findings seem to interfere with our hypothesis, which assumes that highly standardized song types are more attractive. However, complementary evidence is found for a preference for stable songs.

3.2.4 Stable songs
Seasonal plasticity in the brain is in some species correlated with stereotypy (less song-to-song variability). Tramontin and Brenowitz (2000) found a relation between the degree of song stereotypy and the seasonal growth of brain tissue, thereby implying that a bird needs more brain space to sing less variable songs. This sounds counterintuitive, but Leitner, Voigt and Gahr (2001) also found that more repetitive songs are uttered in the breeding season. Combining the preference for stable songs with the preference for variability, it holds that a high number of standardized or stereotypic song types could indicate quality. Stereotypy is easier to judge in regular signals; that means that a propensity might exist for rhythmic songs to develop.

Ballintijn and Ten Cate (1999) describe deletions of elements in the vocalizations of the collared dove, a non-songbird, and found a negative correlation between the body weight and the occurrence of a deletion of an element. Differences in rhythm were perceived by (male) conspecifics (Slabbekoorn and Ten Cate, 1999).

3.2.5 Shifting rate
Another temporal aspect of song related to song stereotypy is the switching rate between song types. When a male encounters a female, the song switching-rate goes up (Langmore, 1997). Vehrencamp (2000) considers the increasing of the shifting rate just as an amplifier of the repertoire size, so to let the male show off with that sexy feature.

At least two other possible explanations exist for the increase in song switching, which is a general finding among songbirds. The first is that the chance of coming across ‘sexy syllables’ increases with the number of switches. Yet another explanation is that fast switching is difficult for a bird, thereby forming an honest signal of quality by itself (Langmore, 1997). Memory interference is probably a crucial factor that determines the cost of switching between activities or song types.

In all the views listed above, switching makes it harder to produce the signal correctly. Another interpretation is that switching between song types simply makes it easier for the male to sing the song he has in mind. Switching surreptitiously to another song type prevents strained muscles from getting tired. This would yield the same ‘high switching rate’ result as the above explanations, but only for low-quality males. Indeed Lambrechts & Dhondt (1986) found that the males that produce songs with longer strophes are responsible for more viable offspring in the next generation.
These findings (switching rate going up in the presence of a female and at the same time the female preference for longer, unabridged strophes) and their respective hypotheses are contradictive as well as controversial. Therefore, the phenomenon ‘drift’ is addressed in a more detailed way (see the box on ‘Temporal drift’).

**TEMPORAL DRIFT**

Drift is a sign of weariness, which is only observable over time. Towards the end of a bout, the intervals between the phrases become larger and the song output declines. After switching to a new song type, which involves different contraction of the syringeal muscles, the drift is gone. According to this, a big repertoire might be just an adaptation in order to sing longer correct songs. Lambrechts and Dhondt (1988) brought up the question whether drift might indicate a lack of quality by the singer. Their Anti-exhaustion hypothesis reads that if the male is of lower overall quality, singing the same phrase over and over progressively fatigues his syringeal muscles.

Indeed, it has been observed that just before a song switch, the temporal drift is becoming larger (Lambrechts and Dhondt, 1988). Kempenaers, Verheyen, and Dhondt (1997) found that drift correlates with lack of success in gaining paternity. In addition, strophe length with little drift was a good predictor of rank in great tits, and significantly correlated with age. The most convincing finding is, as mentioned before, that the males with the longer strophes had more offspring in the overwintering population (Lambrechts and Dhondt, 1986). This could be explained in a way that not the ‘long songs’ are more attractive to the females, but the lack of drift in the songs, independent of how long they are; in other words, regularity and rhythmicity of songs.

Ballintijn and Ten Cate (1999) describe deletions of elements in the vocalizations of the collared dove, a non-songbird, and found a negative correlation between the body weight and the occurrence of a deletion of an element. It remained unclear whether these signs of physical weakness were perceived as such by conspecifics; but distortions of the internal rhythm, by deleting an element without replacing it with a pause, were perceived by (male) conspecifics (Slabbekoorn and Ten Cate, 1999).

The Anti-exhaustion hypothesis has been fought over by Weary, Krebs, Eddyshaw, et al. (1988). They proposed the motivational hypothesis, which says that the motivation of some males is lower instead of the viability, as the former hypothesis implies. It sounds counterintuitive: why would non-motivated males bother to sing at all? How could this property survive over evolutionary time? All together, at best weak evidence is found for both, and the two theories can exist next to each other: tired from singing and tired of singing. (Weary et al., 1988; Poesel and Kempenaers, 2000)

**References:**


To conclude, it is not investigated whether rhythm as such is hard to produce for a bird. But the evidence on rhythmic sexy syllables, revealing information about the respiratory muscles of the male (Suthers, Goller, and Pytte, 1999), together with the finding that more brain tissue is needed for stereotypic songs, provides some hints towards rhythm or at least regularity. Furthermore, how long a male is able to continue with the same temporal structure (that is, his lack of drift) reveals something about his muscles or brain to the female.

3.3 Neural temporal processing

Partly because of the similarities between bird song and human speech production, the bird’s neural system that coordinates both the fine-grained movements of the syringeal muscles and the respiratory apparatus has been studied exhaustively. Another reason for the tremendous concern about bird song is the regularity of song production, which facilitates the search for corresponding neuronal activity patterns. The next paragraph reviews the important neural structures, to find out the exact relation between neural signals and the actual song output; moreover, we need to know to what extent the output is influenced by distortions in the brain, thereby revealing important information to a female.

3.3.1 Timing in song production

Neurons can represent temporal information either by the average rate of the spike train, the rate coding hypothesis, or by the precise timing of the spikes, the temporal coding hypothesis (as described in DeCharms and Zador, 2000). The latter is at least partly true for temporal information: when stimulated, auditory cortical neurons do change their rate of coordinated spikes, without changing their mean spike rate (DeCharms and Merzenich, 1996). On a population level, the coordinated coding hypothesis (DeCharms and Zador, 2000) suggests that a signal cannot be unraveled by looking at the independent votes of all neurons. This seems to be true for temporal coding by definition, but most of the research done on avian neural systems concentrates on the activity of single neurons.

Indeed Yu and Margoliash (1996) found that the spike burst structure of one single neuron in the nucleus Robustus Archistriatalis (RA) was a good predictor for the occurrence of a specific note. In a follow-up study, Chi and Margoliash (2001) conclude that the neural code for song timing in zebra finches is temporal, operating at the single neuron level. At the same time, they suggest that such temporal precision probably is the result of a tightly connected pattern-generating circuitry.
In other parts of the neural system, the role that single neurons play in the system remains controversial too.

A commentary by Nottebohm (2002) on Hahnloser, Kozhevnikov, and Fee (2002) reveals the division of opinion on the role of neurons in the Higher Vocal centre (HVC). Hahnloser et al. (2002) describe their functioning merely as temporal ‘grandmother-cells’, coding for one unique time in the sequence. Nottebohm introduces the term ‘overseers’. He also assigns the HVC-neurons a leading role in the birdsong’s clockwork, but places emphasis on the distributed representation of timing, whereas Hahnloser et al. (2002) stresses the sparse representation.

More specifically, Hahnloser et al. (2002) found that during a song motif, one specific HVC-neuron fires to the RA area, which in turn projects to the motor circuits. The HVC-neurons only fire at one time step in the song sequence (the grandmother cell analogy), as opposed to the RA-neurons, which merely function as transfer point and can be triggered by several different HVC-neurons. More research needs to be done on groups of RA-projecting neurons in the HVC in order to describe their integrated behavior.

### 3.3.2 Neural discrimination of rhythm

The question remains why female birds, in particular the species in which only the male sings, have such large song control structures. Although the view is widely accepted that the female HVC is used for song perception, little evidence is found. Only very recently, Leitner and Catchpole (2002) found that the size of the HVC of female canaries was correlated with two important properties for sexual selection. Both the amount of discrimination between sexy and non-sexy syllables as well as the total number of copulation solicitation displays depended on the size of the HVC.

How can a female bird be so sensitive to slight deviations of interpulse intervals? Recently, evidence was found that in some anuran species, the female midbrain-neurons are sensitive to one specific interpulse interval and only respond after a fixed number of pulses (Edwards, Alder, and Rose, 2002).
Even the slightest deviation from the rhythmic series caused a reset of the interval-counting neurons. In birds, however, such a neural ‘counting’ mechanism has not been found yet. Some older evidence that female brain structures (specifically, the HVC) play a role in perception was found in redwings. Brenowitz (1982) tested male and female redwings on their discrimination ability between conspecific songs and the mockingbird imitations of it. These imitations only appear to vary with respect to the fine temporal structure. He found that males did not discriminate, whereas the females did. In addition, Konishi and Akutagawa (1985) found that female Bengalese finches, as opposed to males, lack the mapping between the HVC and the RA, which is exactly the link between higher representation of vocal structures and the motor neurons, suggesting that the female HVC functions in roughly the same way as the male, the only difference being the utilization of the neural outcomes.

This is evidence that female birds do have ‘a brain of their own’ (Hamilton, King, Sengelaub, et al., 1997) for the purpose of judging male songs. Does it make sense to judge the qualities of a male by the temporal aspects of his song?

3.3.3 Do temporal aspects of song indicate (neuronal) quality

Several studies address this question, though only indirectly. For instance, Chi and Margoliash (2001) mention some functional implications of temporal drift. But the key point they make is the correlation in timing between RA-neurons and the actual song output. Once they established the same correlation with regard to temporal drift, they aridly state: ‘We demonstrated that the changes in neuronal timing were reflected in concomitant changes in acoustic timing’ (p. 905), disregarding the impact this finding has on the origin of temporal drift. As described before, drift was currently explained in terms of exhaustion or motivation. Now that the correlation between acoustic drift and neural coding has been established, the motivation-theory is losing ground (within-song decrease in ‘motivation’ on the neuronal level seems a strange notion), and the anti-exhaustion explanation must be refined towards a neuronal hypothesis instead of a muscular one. Still, with their findings they establish a direct correlation between neural properties and temporal disorders, so it might be adaptive to base a judgment on the regularity of a song. As it turns out, motivation deviously plays a role in the precision of song timing.

3.3.4 Noise in neural systems

In the same article, Chi and Margoliash (2001) mention unpublished data, in which the few songs that were not directed to females yield less reliable neural patterns. Testing this was instigated by an earlier finding in zebra finches from Hessler and Doupe (1999). They found that the total spike rate
recorded in Area X, one of the anterior forebrain nuclei, drastically decreased when singing was directed to females, and that the distinct pattern associated with different syllables was more evident in the social condition. The authors blame the difference on the amount of background noise (‘unpatterned burstiness’, as they call it), which indeed showed a distinct decrease when the song was directed towards a female. A qualitative difference in song output was not found; this might be because of smaller, or no effects at all, of social condition on the motor nucleus RA. Still, it remains unclear why changes in the RA do not lead to qualitative changes in the song output. It is known that the pathway between the RA and the motor nucleus is developmentally regulated, with the connection to motor nucleus becoming more stable and subtle over development. In well-developed birds, with no noise present in their songs in both of the social conditions, this connection might 'repair' the noisy pattern generated in Area X. This would make the noise present in female-directed songs a good indicator for developmental stability, age and neural noise.

3.3.5 Brain structures and inheritable variation

When proposing sexual selection mechanisms that work on birdsong, a genetic link has to be established between the song feature and the preference for it. So far, evidence is found for the heredity of many song features; for instance, the size of the male HVC has a strong genetic component (Airey et al., 2002a) and is correlated with repertoire size or complexity (Airey et al., 2000b). Moreover, the degree to which female canaries discriminate between songs with different numbers of sexy syllables is correlated with the size of their HVC (Leitner and Catchpole, 2002), and, in addition, it is known that females who prefer longer songs produce children that are more viable. However, the heritability of the respective preference in females is harder to assess and no study has attempted to assess the preferences of the female’s offspring. Moreover, if they would turn out to be heritable, the correlations might still be mediated by song learning capacity. The findings suggest that females are able to pick up some very fine-grained aberrations in the temporal order of a male song. What is the nature of these ‘mistakes’ and which ones are the most obnoxious to her ears?

3.4 Mistakes and female reaction

Drift is just one example of what can go wrong when a bird overestimates his own singing abilities. Overall, ethologists have a hard time telling mistakes from intended variation, whereas note omissions are by nature hard to detect. A good cue for the male’s performance is provided by the response behavior of the listening female, but then again the problem is to detect the nature of the mistake: is it a frequency shift, note omission or timing shift? Because of the complexity, most of the research on this topic is based on manipulations.
It appears that the most common mistakes can be divided in three: insertions, deletions and syllable alterations.

Female zebra finch (left) skeptically listening to a male

http://soma.npa.uiuc.edu/courses/physl490b/models/birdsong_learning/bird_song.html

3.4.1 Inserted pauses

One way to exacerbate well-defined mistakes is to force the male bird to ‘supernormal’ performances, by exposing him to manipulated songs during the learning period. The modification can be constructed in such a way that failures to reproduce the modified tutor song are only due to motor constraints, and are not influenced by memory constraints. Then the amount of female copulation solicitation display (CSD) can be assessed and compared with the non-manipulated song.

This is exactly what Podos (1996) did. He found that young swamp sparrows, when forced to sing syllables at abnormal repetition rates, deliver ‘broken’ songs. Rather than slowing down the phrase, they try to come up with the tutor speed and introduce gaps in between syllables, thus shifting the song one time-step forwards. In a later study, Nowicki, Searcy, Hughes, et al. (2001) assessed female reaction to broken songs. They found that females responded significantly less to broken songs than to normal songs, whereas the male’s aggressive response did not change at all.

3.4.2 Deleted notes

Helekar, Marsh, Viswanath, et al. (2000) performed the most comprehensive study on natural pattern-variation. They found a surprisingly high prevalence of syllable deletions and insertions in a zebra finch colony: both deletions and insertions occurred in songs of more than two-thirds of the finches. Alterations of syllables occur slightly less often. The frequency of mistakes seems to go up at the end
of a song motif, but does not depend on the amount of singing. Unfortunately, the authors refrain from testing the attractiveness of the different song variations on female finches. Such a test might reveal important data: for instance, if the difficult sexy syllables were left out, the song would lose attractiveness. Inserted sexy syllables, on the contrary, would raise the number of female CSDs. The article now focuses on inferring the structure of the motor control pathway and lumps together all pattern variations. Deleting different elements sometimes produces different response effects: as Ballintijn and Ten Cate (1999) describe, removing the third element in a dove-coo without replacing it by silence has a strong effect on the response, but the absence of the second or third element does not alter the reaction of the conspecific.

3.4.3 Repeated syllables

Repetition of syllables is seldom observed. Compared to insertion of arbitrary syllables, repetitions are observed only once in every five occurrences (Helekar et al., 2000).

As with the broken songs, ecologists sought a way to induce syllable repetition experimentally. The only way to force songbirds to repetition was by manipulating the auditory feedback, either by delaying (Leonardo and Konishi, 1999) or, in a more rigid way, by surgically deafening the bird (Woolley and Rubel, 1997). Both methods yield roughly the same results: disturbance of syllable sequences, including deletion and repetition of syllables (‘stuttering’). These phenomena suggest a mechanism of song production that relies on an excitable motor program that can be triggered repetitively by the Higher Vocal centre, which corresponds with the recent findings of Hahnloser et al (2002).

3.5 Rhythm in birdsong

To conclude, evidence can be found that in birds, several sexual selection principles have had their influence on different temporal aspects of bird song (Searcy and Yasukawa, 1996).

It has not been investigated yet whether the underlying neural basis for the production and perception of rhythm is heritable. In order to demonstrate this, the neural structure has to be unraveled further and a correlation has to be established with a brain structure or feature. To some degree, the structure that causes female appreciation and the male ‘rhythmicity’ should be both heritable and genetically linked, but in order to investigate this, more knowledge about female preference needs to be gained.

This can be accomplished by seeking correlations between mating success and rhythmic aspects of songs, and on the other hand, by testing systematically the exact form of the female preference function using playback experiments. If a lack of noise (whether on a developmental, neural or behavioral level) indeed is advantageous for a bird, a rhythmic song might be the peacock’s tail by which a male reveals the low level of noise present in his neural system.
In order to test the plausibility of this hypothesis, in the next chapter a model is described, which emanates from the information and evidence listed above. The most important features needed for implementation are

- Paternal quality (thus, number of offspring) is correlated with the quality of song expression (section 3.2);
- The blueprint of the song is learned or inherited from the parents, but the temporal aspects from the expressed song depend on the quality of the singer (see Temporal Drift);
- Females are able to judge a male by these fine temporal aspects of his song and indeed pick on average males with higher genetic quality (section 3.3.);
- Mistakes in song expression take the form of insertions, deletions or alterations (section 3.4);
- Some parameters on the other aspects of a bird's life, like the average clutch size and degree of monogamy (section 3.1).
4 The model

4.1 Description of the model

The model aims to combine two things. It is based on evolutionary simulations to capture the process of honest signaling of sexual indicators, as described in the second chapter. This phramework will be applied to the “prototype birds” from third chapter. By combining these two, the hypothesis is put forward that rhythmicity might turn out to be an indicator of neural or behavioral noise.

In this paragraph, the choices and simplifications made during the transition from nature to model will be described.

4.1.1 Individuals

The co-evolutionary model involves two types of individuals: male and female ‘birds’. The birds do not have a specific predetermined sex, but can be drawn from the population to serve either as a mom or a dad.

Both of them are born with a basic song template, which they inherit from their parents. This follows the work of Marler (1970), who found evidence for partially inherited birdsong in young birds who utter a crude song even if they are deafened after birth. He postulated an inheritable template that is only later fine-tuned using feedback. This is true for most non-songbirds. Here, we only include inherited song by fixing the template at birth, thus leaving aside learning.

As in nature, the males have to advertise themselves to potentially interested females by singing. In fact, in the model males are reduced to a template with a singing device. The way the template is expressed (i.e., the behavioral phenotype that is produced) depends on the quality of the male, as described below. The females are the choosy sex, because the number of offspring they can have is limited (they can only mate once per generation, whereas a particularly tuneful or lucky male could mate many times). The females consist of a judging device that uses a heritable song template of the same kind as the males. This assumption seems reasonable, because as Leitner and Catchpole (2002) showed, the judging capacity of the female is correlated wit the size of the HVC.

4.1.2 The template

The song template decodes a temporal sequence of a very simple sort, representing what is happening in the song at a sequence of equally-spaced points in time. Here we cut the notion of a “song” down to its simplest binary form: At each instant, an element can either be sung or not, so that the template
can be either on or off (1 or 0). Thus, a template with ten time-steps (the length we typically use here) reflecting the sequence of notes from left to right looks like “0010110010”. Every individual has two templates: one that is expressed when it plays the ‘male’ role, and another female template that is used to judge other males with. The way a template will be expressed, the phenotype, depends on the quality of the male.

Like most sexually selected traits, the expression of the song has both a genetically transmitted and a condition-dependent component (Andersson 1994). Whereas the template is inherited from the parents, the condition or quality (in range 0.0-1.0) is given randomly assigned to males at birth (females do not have condition-dependent traits). Quality has two effects for males. First, the chance that the male bird makes a mistake in singing his template is inversely related to this quality: ‘noisiness’ is 1-quality. Second, quality also determines the relative number of offspring a male can have, if a female chooses him.

4.1.3 Selection

In order to get chosen, the male sings. He tries to copy his template, but relative to his quality, he will make mistakes. For instance, he could introduce a gap and shift the song one-time step forward, or he might forget a note, thereby shifting the song backwards in time, etc. Thus mutated, the final song, including the mistakes (insertion and deletion), is the only thing a female can judge him on. Similarity of genotype is not guaranteed, but she will pick the male with the song that resembles her template most closely.

In the model, females are monogamous. This is a very rough simplification; in fact, in lots of real species the sexual selection goes on even fiercer after nest building. In many passerine species, the number of offspring resulting from extra pair copulations exceeds the eggs from the partner of the female. Neglecting this does not affect the outcomes of the model. The fact that a female can only choose once cannot withhold the males from polygyny, and many females with the same template exist. However, a model from Kirkpatrick et al. (1990) suggests that the outcomes of the sexual selection process might be less conspicuous in monogamous birds.

4.1.4 Judgment

A ‘choir’ of a fixed size is selected randomly, because in the real world one cannot choose his neighbors too. This limited sample reflects the time pressure to mate that real birds face. Every female judges the songs from these males with respect to her own template. We build in a mechanism by which the female makes her choice. The female picks the male in her choir whose performed song (phenotype) is closest—in some sense, defined below—to her own template (genotype) and mates with him, following a best-of-N rule (Real, 1990). Given the possible noise in the male’s singing, similarity between his genotype and phenotype is not guaranteed, but, given low noise rates, each
female will on average pick the male with the song template genotype that is closest (again, in some defined sense) to her own template genotype. Birds of a feather flock together, as they say. How does the female judge the similarity between the male song she hears and her own internal song template? The calculated similarity or distance depends on her preferences, for instance whether she is annoyed by missing notes, or by new notes that she was not expecting, or both. The simplest comparison measure would be the summed difference per time step, combining both of these preference components. Exactly how she judges is captured in her preference table, which defines how the female rates what she hears (in the male song) against what she expected (in her template). These preferences are computed for each time step of the song, and the sum across the whole song is the female’s judgment of that male. Then, Gaussian noise is added to the total score.

4.1.5 Recombination

Once every female has chosen a mate, offspring are made from combinations of the templates of the parents. Here, we face the choice how to define heredity. The simplest way would be to mix the male and female genes (templates): for instance, the father’s male template then could become the child’s female template. As Riebel (2000) holds, not only the sons form their song according to their tutor, but the daughter’s preferences are influenced by the father’s or tutor’s song as well. Another way would be to keep the template types separated by sexes. This vertical transmission might slow down evolution a bit, but essentially the same effects would occur, and no assumptions about cross-sexual inheritance are to be made. The optimal mate, that is, the preference of the daughter, is either equivalent to the female genotype of the mother, or to the female genotype of the father, which he did not use during his life. Then, there is a small chance, defined per bit, that a mutation occurs. Crossover is not implemented in the sense that a new template is constructed from a combination of the parent’s templates: the way the genes code for information does not allow for this. Depending on the father’s quality, some of these ‘imprinted’ eggs hatch and are placed in the final offspring array, which will form the population in the next generation. Because the father’s quality only influences the number of offspring, quality should be regarded as male investment rather than as a heritable feature. The new generation gets again a random, environmentally determined quality. Quality does not reflect ‘good genes’ in this case: it determines only the chance that the genes of this particular couple will be passed on to the next generation.

As emphasized in chapter 2, we cannot define quality as heritable, because the link between rhythmicity and quality is what we want to investigate. If both the templates and the quality would be heritable, there would be too many degrees of freedom: one template would get attached to high quality, just randomly.
4.2 Parameters and their influence

In all genetic models, the parameters that affect the outcomes of the model must be tuned. Therefore, in this section the most important parameters are listed, along with their influences. Because the different ways of singing and judging are the parameters we want to compare, they are addressed apart.

4.2.1 Parameters and convergence

The influence of the parameters is addressed by looking at their effects on the convergence of the population. Convergence describes the degree of similarity between the genotypes and is based on the deviation per bit. Typically, it will diminish over generations until a certain fixed number is reached, depending heavily on the mutation rate. An ideal convergence speed would leave room for better solutions to take over during the beginning (preventing the population from premature convergence) but converge in the end to a reasonably low deviation. All graphs are averaged over 5 runs, with standard parameters.
If the population size is too small, it is unattractive to have a new but slightly better song, because no female has a template matching it. In fact, natural populations of birds are as conservative with regard to song innovations, according to Nowicki, Searcy, Hughes et al. (2001). However, with a starting population that is large and diverse enough and some mutation, the chance that a male deviant would genetically survive is bigger, since he might run into a deviant female who recognizes his unexploited abilities.

Figure 4.1 shows that reasonable size is around 200 individuals. To be sure, some runs with 1000 individuals are done for 10000 generations; this did not change the outcomes.
The number of males that a female can choose from regulates the speed of the selection. The bigger the choir, the better the female can choose, and therefore the greater the effect of her choice is. In nature, there comes a cost for the female with choosiness; it takes time, and with every potential partner she refuses, her chances on the mating market shrink. The graph shows that a big choir converges too quickly, whereas a choir of 2 results in too much deviation. A choir size of 20 individuals is therefore reasonable.

![Quality-offspring Ratio](image)

*Figure 4.3: Quality-offspring Ratio. Convergence (y-axis) against generations (x-axis) for different quality-offspring ratios, averaged over 5 runs.*

The function determining the relation between male quality and the number of offspring essentially controls the speed of convergence. A function that is too steep would let the genes from one pair take over the whole next generation. On the other hand, there must be enough difference between the fertility of low-quality and high-quality males to ensure that evolution is going in the desired direction.

To see what form this function should have, we tried a logistic, two quadratic, and a linear one; the constants do not contribute to the relative number of offspring.

Surprisingly, the simple linear function (blue) has the effect on convergence speed that we want (see figure 4.3): the same high convergence at the end as the two quadratic functions, but with slope that is less steep.
Mutation is implemented per bit. A mutation rate of 0.01 means one out of 10 templates will have one mutation, because the length of the template is 10.

A way to find the optimum mutation rate is by looking at the resulting change after 1000 generations: after initiating the whole starting population with a good template, ideally no changes occur.

4.3 Singing and judging

4.3.1 Singing

From the biological evidence on temporal disorders, whether artificially induced or not, we can infer a way for a simplified male bird to express his inborn template.

We assume that a he tries to reproduce the template at each time step. This way, every time step there is an equally small chance that he does something wrong (which is not true, see for instance Chapter 3: Temporal Drift). He might for instance forget (delete) a note, insert a note or a pause or even sing a different note than planned. Recall that the number of mistakes he makes is inversely related to his quality; for each bit, we define a chance that an ‘error’ will occur. The error has a number of possible instances: The simplest mistakes are a deletion or an insertion of a note (temporally, a shift backwards or forwards). Because repetitions are very rare in nature, every insertion is assumed be an inserted pause. A zero appears on the actual time step and the whole template shifts one bit to the right. The last bit is removed. In the case of a deletion, the male just skips one bit (either a one or a zero) during singing and the whole template is shifted to the left, with on the last bit a zero appearing.

A plain mistake or syllable alteration does not interact with the time scale. Although in real life, an alteration means that something within the syllable was changed, in the model the whole syllable will be replaced: instead of a one in the template, the song will have a zero, and the other way around. This way of making mistakes can be combined with inserting and deleting, by defining for every bit the chance that either one of the three mistakes are made.
Other implementations of mistakes are possible as well: consider repetitions (by taking a number of n notes, ranging from 0 to 3, and doubling the sequence to form a mistake), or a method based on the comparison between the both the template and the song itself (see ‘Transitions’), but we focus on these relatively simple types.

### 4.3.2 Judging

For the model, we assumed that the female judging system is not perfect with regard to the information it receives. As the males are not inviolable and do make mistakes during singing, her perception can be blurred a bit as well. Fewer good poets than good poetry readers exist, so it seems that by definition judgment is easier than production; we chose to add Gaussian noise (sd = 1.5) over the final female judgment instead of defining an error per bit, as was done in the male case. The latter would certainly induce interfering effects and thereby change the outcomes.

There are various ways to compare the male song to the female template for judgment, starting with the Hamming distance. When we take into account only values from the set {-1, -0.5, 0, 0.5, 1}, a reasonable way to keep the number of possible preference tables (5^4, or 625) low is by restricting them to the ones in which the right note at the right place gets rewarded. That is, the rewards for 11 (expected notes) and 00 (expected pauses) must be bigger than the values for unexpected events. This forces the judging device to make a difference between a punishment and a reward. Specifying this in turn, the reward for a note on the right place might be bigger than the reward for an accurate pause. On the other hand, the female’s expectations might be violated more by a sudden note than by a pause, as we know from the biological literature (Ballintijn & Ten Cate, 1999).

We can now construct a preference matrix:

<table>
<thead>
<tr>
<th>Expected</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sung</td>
<td>X1</td>
<td>X2</td>
</tr>
<tr>
<td>1</td>
<td>X3</td>
<td>X4</td>
</tr>
</tbody>
</table>

and set a restriction:

\[ 1 = X4 >= X1 > X2 > = X3 \]

In words: the reward for an expected note (X4) is always equal or bigger than the reward for an expected pause (X1); for each, the reward must be bigger than for an unexpected pause (X2). An unexpected pause must in turn be rewarded higher (or punished less) than an unexpected note (X3).
This restriction filters out all duplications of the preference matrix if we set X4 to 1. Inversions or multiplications of the matrix do not occur in the list. Furthermore, the restriction is based on the idea that for qualitatively the same outcomes (for instance, the table −1 −1 0 0 would yield the same results as 1 1 −1 −1) we show the positive occurrences rather than their negative correlates.

Some qualitative differences exist between the different instances of the table. The symmetrical judging table (e.g. 1001) is a base case. With this matrix, no bias towards more or less ones than zeros exists; a female is basically calculating the Hamming distance between her template and the male song. On the contrary, an asymmetrical judging matrix (0,−1,−1,1), differs from the former one in that there is no reward for producing a zero at the right time, but there is a big punishment for being off the beat. Therefore, a smart low-quality male might reasonably sit still and produce only zeros, thereby hiding his mistakes.

A slight variation, by lowering the punishment on sudden pauses, is (X1-X4) = 0,−1,−1,1. In this case, it matters whether the mistake was just a gap in the song (i.e., a zero appearing where a one was expected), or a surprising new note. This is a way of circumventing the problem that a male could sing nothing at all, not revealing his (ar)hythmicity.

4.4 Expectations

4.4.1 Evolution in detail

We can get a gut-feeling about the model by looking at the first generations in more detail. We focus on a preference table with a reward for the right notes and a indifference for unexpected events, that is, (X1-X4) = 1,0,0,1. For simplicity, we assume that only two types of males exist: high quality (q=1) and low quality (q=0.0000001) males. With this restriction, the genes from the low quality males do not contribute to next generation at all. Furthermore, the choirs are left out of consideration: every female sees every male in the population, and we assume that there is no noise added to her judgment for now. We can look at the first generations in more detail, in order to see what happens on the small scale.

Given these adjustments, we furthermore initialize the whole population (100 individuals) with one template first, [0000000000]. In the first generation, half of the population (50) will not suffer from a shift in their songs at all; so all the ‘perfect’ females (100) will choose them. From their brood, with a mutation rate of 0.01 per bit, 1/10 (10) females will suffer from a mutation in one of the ten bits, and 1/10 males too, half of which happen to receive a high quality. Therefore, 1 out 20 (5) males has a high quality and a copy-error (a flip) in their genes. Their template might look like [0001000000], for
instance. If a female looks the same (as one female does, on average), she will choose him and produce a lot of children (6), of which most have two templates with a one in the sequence.

The average quality of the chosen males with the template [0000000000] will be exactly 0.5, because it cannot be perceived whether a male is perfect (genotype = phenotype) or has made a mistake (a deletion or insertion of a zero). Half of the mutated offspring will extinct, due to their low quality, regardless if they are chosen or not; however, the remaining three will (as female) only choose a similar partner, that is, one of the mutants with a high quality. Mistakes can be noticed by a shift of the 1 [0000100000]. A female with the template [0001000000] thus can discriminate better between high and low quality partners, so the average quality of the males chosen by her will be higher. On the long term, this will result in more and more 'mutants', and the template [0000000000] will not be kept as a solution.

In the other case, when all individuals are set off with an alternating template [1010101010], one shift clearly indicates that the male has low quality. The score the female gives to a shifted template is in the worst case 0: if the mistake is made on the first time step, all notes differ between the female template and the song, they are "out of phase". Therefore, a low quality male will not easily be mistaken for a high quality male, even with the Gaussian noise added to the score.

This noise is a necessary adjustment. If all female judgments were completely accurate, the template [0000000001] would be as good as the rhythmic template, simply because all females would notice a mistake or mutation immediately and never choose a deviant template. With noise added to their scores, however, the difference between scores for a perfect male and a male of lower quality matter. A female with a rhythmic template is on average better at choosing high quality males, because the error is not that large that she would not recognize the man of her dreams. Adding Gaussian noise to each female's appraisal of a male song increases the need for using highly discriminable templates.

Given this framework, we now want to ask whether females can use the male songs they hear to judge male quality and make good mate choices. Moreover, since good mate choices would lead to more offspring, the question is whether appropriate female templates and male songs will evolve to allow for such choices. Our simulations aimed to find out if the female expectations and male performances could evolve to work together. By implementing a way of assortative mating, we expect to see them both converging to one template; we want to explore under what circumstances a rhythmic template would be the outcome of this process. A workable definition of rhythm is found in Kenward, Wachtmeister, Ghirlanda, et al.(2002): Rhythm is regarded as the number of occurrences times the length of the repeated sequence (the 'generator').
5 Results

Figuring out in detail what the model yields is not possible without knowledge of the inbuilt premises. Therefore we did an analysis of the female preferences, in a non-evolutionary way, to examine what influence female choice has on the evolutionary setting. Note however that female choice is not the only evolutionary force acting on the form of the templates; what the females actually prefer during one generation also depends on what the males are doing. However, if we set the female templates all to one template and only let the males evolve, they evolve towards the fixed female template. But what do females prefer?

5.1 Outline

The template we want to find has to be advantageous for the reproduction rate of the female. This will be the case if she can discriminate accurately between high quality and low quality males, using her template as an expectancy pattern. In other words, we first need to find the templates with which the female is more likely to assess male quality correctly. The optimal template differs depending on the female preference tables and the type and amount of male mistakes (section 5.1).

Once we have found these most revealing templates, the next step is to find out how much of help they really are for the females that carry them. If such a template provides a female with a more accurate view on the quality of male birds, then the quality of the males chosen by her will be higher on average. The difference between these average male qualities forms a measure for how much of an attractor the template used for judging is in the ‘seascape’ of the coevolutionary process (section 5.2). Now it needs to be proven to what extend it is useful to have a revealing template. Plotting female templates against offspring would give the direct answer; however, the effect can be tested in the evolutionary setting as well. For instance, a template consisting of only ones will be very revealing too, because whenever a shift occurs, a conspicuous zero is inserted. By setting half the initial population to the most revealing template, and the other half to a local optimum, the question ‘which one wins’ can be addressed in terms of offspring (section 5.3). The next step is to find out how stable a solution is in terms of the evolutionary setting (section 5.4). Do the appropriate female templates and male song evolve together to allow for better mate choices?
5.2 Template discriminability

To find out what the ‘appropriate’ female templates are, we first constructed an optimality model without evolution. For every possible template with length 10 (1024), we created a series of 10000 ‘noisy songs’, each with on average one shiftforward mutation (or with noise = 0.5). This stochastic approach was chosen because it resembles the evolutionary approach better. First, we computed the maximum score for every template (the score the female gives if she encounters a copy of her own template) in order to be able to standardize the outcomes. Otherwise, in the case of an asymmetric score table, there would be an undesired difference between the scores for the templates 1111111111 and 0000000000. This ‘pure’ preference score can vary between [-10,10].

Thus normalized, we can sum the relative preference scores for 10000 mutated templates. The gap between the best score possible with a certain fixed table (whether this means exactly the same or making all mistakes possible) and the mean is a measure for how well a female with this template can tell the males with low quality from the ones with high quality. Recall that the mean number of mistakes is always the same (1). The most discriminative template for a female is at the same time her optimum, since in the evolutionary setting her quality-assessing abilities determine if her genes will be passed on. This way we can analyze her optimum template without regarding the fashion of the day: whether it is indeed the optimum at a certain evolutionary time step depends on the environment that is formed by her conspecifics. In the evolutionary setting, she might be behind the fashion, which influences the impact she has. Because Gaussian noise is added to the female assessments, the more revealing templates will on average result in choosing better males.

The complete results, 25 qualitatively different preference tables with their most revealing template, are listed in Table 4.1; and for one preference table, (X1-X4 = 1001), a list with ‘discriminability scores’ for some of the 1024 possible templates is given (Appendix A), all for on average one deletion or insertion.

5.2.1 Pseudocode discriminability

First, generate the 10^{10} (1024) different templates.

Then, for every possible preference table (systematically increases from –1,-1,-1,-1 to 1,1,1,1 in .5 steps)

---

3 by definition, because the values in the preference tables vary between –1 and 1, and the song has length 10.
For every template
Generate 10000 noisy templates
Print the on average ‘best’ template, that is, the most discriminative

The ‘best’ is computed by
Summing the (preference for the noisy template – preference for the pure template) over 10000 templates
Dividing this by the 10000 in order to get the mean, and dividing it by the preference for the pure template.

5.2.2 Discriminability results

Table 5.1: Most discriminating song template for the restricted value combinations of the preference table (X1-X4). Only four preference tables give a non-rhythmic pattern as the most discriminating one.

<table>
<thead>
<tr>
<th>X1</th>
<th>X2</th>
<th>X3</th>
<th>X4</th>
<th>Template</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.5</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>0010101001</td>
</tr>
<tr>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>0000000001</td>
</tr>
<tr>
<td>0.5</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>0101010101</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>0101010101</td>
</tr>
<tr>
<td>0</td>
<td>-1</td>
<td>-0.5</td>
<td>1</td>
<td>0000000100</td>
</tr>
<tr>
<td>0.5</td>
<td>-1</td>
<td>-0.5</td>
<td>1</td>
<td>0101010101</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-0.5</td>
<td>1</td>
<td>0101010101</td>
</tr>
<tr>
<td>0.5</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>0101010101</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>0101010101</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>0.5</td>
<td>1</td>
<td>0101010101</td>
</tr>
<tr>
<td>0</td>
<td>-0.5</td>
<td>-0.5</td>
<td>1</td>
<td>0000000001</td>
</tr>
<tr>
<td>0.5</td>
<td>-0.5</td>
<td>-0.5</td>
<td>1</td>
<td>0101010101</td>
</tr>
<tr>
<td>1</td>
<td>-0.5</td>
<td>-0.5</td>
<td>1</td>
<td>1010101011</td>
</tr>
<tr>
<td>0.5</td>
<td>-0.5</td>
<td>0</td>
<td>1</td>
<td>0101010101</td>
</tr>
<tr>
<td>1</td>
<td>-0.5</td>
<td>0</td>
<td>1</td>
<td>0101010101</td>
</tr>
<tr>
<td>1</td>
<td>-0.5</td>
<td>0.5</td>
<td>1</td>
<td>0101010101</td>
</tr>
<tr>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0101010101</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0101010101</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0.5</td>
<td>1</td>
<td>0101010101</td>
</tr>
<tr>
<td>1</td>
<td>0.5</td>
<td>0.5</td>
<td>1</td>
<td>1010101011</td>
</tr>
</tbody>
</table>

Across the 20 psychologically plausible preference tables, as shown in Table 5.1, the most discriminating song template was rhythmically alternating (1010101010 or 0101010101) in all but
four cases (with two of the rhythmic templates adding an additional one at the end, but in these cases the fully rhythmic template had very similar discriminability). The four non-rhythmic templates all had X1<0. Thus placing a positive value on producing a pause when it is expected seems essential for rhythmic templates to be the most discriminating. Given this caveat, rhythmic songs templates appear to be the most useful type of signal, for our simulated birds at least, to use to discriminate high quality males from low quality ones.

For many preference tables that yielded the rhythmically alternating template as most discriminating, the difference in discriminability with a monolythic template composed of only ones was not very large (see Appendix A).

5.3 Influence on female choice

From the former analysis, we can pick two templates with highly divergent scores on discriminability. In order to confirm that the rhythmic sequence is also the more adaptive solution, we compare two templates with highly different scores under the same evolutionary circumstances, both for the first generations and for the long-term effects.

5.2.1 First generations

The simplest preference table with the rhythmic result is 1100 (reward for expected notes), with on average one insertion or deletion per song. Mutation rate is 0.001 per bit; population size is 200, choir size is 20.

In the list of preference scores for table 1100, the templates 0101010101 and 1000000000 score respectively D (0101010101) = 0.431 and D(1000000000) = 0.01938. One other template, 0000011111 (D = 0.177) was tested, to make a comparison. Due to the big difference of the preference scores for these templates, the effect on the performance of the females must be clearly visible: if they are able to discriminate better with the template consisting of ones, they will on average choose better males already in the first generations. To compare the different templates, the whole population is initialized with one template. In the table below the three templates and their preference value are listed. The first three generations averaged over 25 evolutionary runs, along with the average quality of the chosen males in that generation. The last row averages all chosen males over the first 100 generations.
As expected, there is as expected a big difference between the two templates with highly diverse discriminability scores in the first generations. Clearly, the template [1000000000] is one of the worst possible; the resulting male quality is only slightly better than average, which must be due solely to the starting digit. The difference for the other two chosen templates is small, but enough to suggest that once the simulation has discovered the rhythmic template, it will be preserved in the population because of the reproductive advantage for the lucky female having this template. On the longer term, the effects of the discriminability differences on the average quality seem to disappear. The templates are apparently subject to mutations; even when the whole population is installed at the optimum (1010101010, with the highest discriminability value), the average template is undergoing changes. We need to test this on the longer term with a lower mutation rate (mutation: 0.0001).

### 5.2.2 Long term

Having done this to watch if it goes well in the first generations, we now have to describe what happens in the long run. How do the templates change? Is quality going up along with the number of alternations? Because we cannot guarantee that the average template remains the same over generations, simply comparing the average quality at the end is not useful. To see the effect of the female choice over generations, quality must be compared to a measure of rhythmicity. We chose to simply count the number of alternations between zero and one notes in all templates in each generation. The maximum for a fully rhythmic template is 9, and the randomly expected mean is 4.5. We can test whether rhythmic templates are much of a help to the females by looking at the average quality of the chosen males per generation. If the females get better in discriminating between males with high and low quality, the average quality will go up.

In figure 5.3 the average quality of chosen males is plotted against generations. Will the average number of changes go up with the same pattern as the average quality, after initializing the population with the template [0000000000]?
Indeed as we can see from figure 5.3, quality goes up with rhythmicity. The graph shows an increase in quality directly after the average template changes to five alternations [0001010011]. The maximum of nine alternations is not reached; moreover, once the average template consists of seven alternations, it changes back to six again. Why is the better solution not stable? To find an answer to this question, we looked pairwise at two templates with different discriminability scores, competing against each other during the same evolutionary run.

5.4 Finding local optima

When the population is filled with two different template types, we expect it to converge to the template with the higher score on the discriminability test. However, from the former analysis, we can conclude that this is not always the case. Which templates are threatening for the rhythmically alternating ones, and how often will the simulation converge to the template with the lower discriminability score?

It would require an imaginary landscape of the solution space to figure this out analytically. For a normal evolutionary model, this is relatively easy, but it is very hard in case of a coevolutionary model. Our derived ‘fitness function’, the discriminability analysis, only applies to the females; the males counterbalance the effects in an unpredictable way, thus contributing to the dynamics of the system.

We can, however, test one of the suspicious templates and see if it forms a deceptive local optimum. From the discriminability analysis (see Appendix A), we know that the discriminability value (D) for the template [1111111111] is 0.165. This is not particularly high compared to the best template.
([0101010101] with \( D = 0.431 \), but in both templates every insertion or deletion is noticed (as it involves a "0") and moreover, the tested template is monolithic. Comparing the two will give us insight in the behavior of two extremes: rhythm against the opposite.

Before we start the evolutionary runs, we compare the computed payoffs from both templates against themselves and against the competing template. The used preference table is still \( X1,X2,X3,X4 = [1,0,0,1] \).

<table>
<thead>
<tr>
<th>Table 5.4: Computed payoffs for two templates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected</td>
</tr>
<tr>
<td>heard</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

The chance for both types of females to choose the other type of males is equally small. Now we can compare the loss of discriminability on the distance of one mutation of the templates. For the template [1111111111], a mutation at one bit (for instance, [1111111110]) is slightly detrimental: the loss in discriminability is 0.005. The same test was done for all direct mutants from the rhythmically alternating template. Dividing the original discriminability by the average discriminability of all mutants yields

\[
D(1010101010) = \frac{0.431}{1.18} = 0.365
\]

\[
D(1111111111) = \frac{0.165}{0.65} = 0.251
\]

For a female with the monolithic template is on average an evolutionary mutation a small gain in discriminability, whereas for the other female every mutation is disadvantageous. The template [1111111111] is apparently not a local optimum. However, consider the following evolutionary path: if the population is initiated with the template [0000000000], and the mutations creep into the template from the beginning of the sequence, the template becomes better than its predecessor with every step. This way, the template [1111111111] can be reached as local maximum: compare \( D[1111111000] = 0.130 \) to \( D[1111111000] = 0.148 \).

Surprisingly, when half of the initial population was initiated with template [1010101010], and the other half with [1111111111], not all evolutionary runs converged to the alternating template: only 7 out of 10 do so. This might be due to other attractors or drift of the population.
5.5 Stability of the solutions

Would the rhythmic templates be stable over time if they predominated the population from the start? To answer this question, we started evolutionary runs using a population filled with alternating templates, both male and female. We used the same parameter settings as we did in chapter two: the population size was 200; choir size was 20, and the mutation probability was 0.0005 per bit. As the population evolves, the average number of alternations goes down gradually, resulting in 7.5 alternations after 1000 generations. Conversely, when the population consists at the beginning only of the constant template [1111111111], it evolves towards an ever more alternating solution. The average after 1000 generations lies above 5 alternations; remember that the randomly expected mean was 4.5. The slope of this line is steeper compared to the first, thereby indicating that the solution is less stable. Convergence does not differ dramatically at the end of both types of runs and is big enough to assume that evolutionary change has not stopped yet.

![Figure 5.5: Stability. Number of alternations (y-axis) plotted against generations (x-axis), each line averaged over five runs. For details: see text above.](image)

5.6 Evolving rhythmicity

Given the above indications,

- rhythmically alternating templates are the most discriminative with regard to the mistakes that we implemented (section 5.2);
- females with such a template are indeed better at discriminating between males with high and low quality (section 5.3);
• this effect is strong enough (although not always) to lead evolution away from the monolithic template, towards a more alternating template (section 5.4), and
• the alternating template is relatively stable against mutations over generations once predominating (section 5.5);

We would expect that a random initialization converges to the optimum template. For this final test, all parameters were again chosen as described in chapter four, with the mutation rate set lower, at 0.001 mutation probability per bit.

![Figure 5.6: Number of alternations in the templates (y-axis) plotted against generations (x-axis) for five different runs](image)

Figure 5.6 shows that, notwithstanding the forces described above, rhythmic results do not evolve unambiguously. After 1000 generations the number of alternations is on average 4.9, which is only slightly better than the randomly expected mean (4.5). None of the runs ended with less than four alternations on average and all evolved away from lower values over time. There does seem to be some indication that females with more alternating templates are able to choose higher-quality males. On the other hand, as the graph shows, not all good solutions could be kept in the population. Overall, we can conclude that although rhythmic signals are relatively stable compared to constant signals, in our framework this effect is not strong enough to drive the evolution in the direction of rhythmic templates. However, we did not find evidence for deceptive local optima, the everchanging seascape in combination with the chosen parameters might have prohibited the rhythmic template from evolving.
6 Conclusion and discussion

6.1 Summary

In this thesis I explored the function of rhythmic behavior as a proximate cue of underlying mate quality. The quality-indicator hypothesis of rhythm was supported by the biological literature. Neural properties of the singing bird are reflected in the temporal pattern of his song, which is in turn influenced by the presence or absence of a listening female; it appears that the females perceive differences in quality through temporal aberrations like temporal drift, and act accordingly to it by preferring rhythmically singing males.

Given this actual situation as an observed instance of an evolutionary process, we raised the question how signals could have evolved to become more rhythmically. In the literature on sexual selection we found some refinements of this question. Sexual selection is regarded by game-theoreticists as concerning a conflict of interests (Krebs and Dawkins, 1994). For the choosy sex, it pays to choose a healthy, fertile partner; on the other hand, for the unhealthy partner, it pays to conceal his state. Despite this conflict, it was shown that indicators of viability or fertility could evolve that could not be faked, if only they are more costly to produce for low quality males than for high quality males (Iwasa et al., 1991).

To translate this notion of handicap into biological terms we looked at bird song traits that vary strikingly between birds, thereby revealing the quality of the singer. In many bird species the process of sexual selection has resulted in complex, elaborated songs, with varying features being preferred in different species: variability, repertoire size, sexy syllables, and duration of songs. These traits have in common that they are harder to produce for lower quality birds. Rhythmic displays in the temporal domain do not impose more costs on the singer than non-rhythmic songs do, but can heighten the salience of neural noise in the singer or other aspects of quality (developmental stability, respiratory fitness, genetic imperfection or current condition). In regularly repeated forms, which can easily be compared with each other, mistakes are clearly visible and a potential partner of low quality will stand little chance to conceal the shape he is in.

We hypothesized that rhythmicity forms a good candidate for an indicator of the general viability of male songbirds, and that it might evolve driven only by the more revealing properties of rhythmic signals. Through the construction and analysis of a coevolutionary model, we investigated the constraints and prerequisites that could have influenced the design of sexual signals in birdsong.

Our model comprises male 'songs' and female 'judging devices', both represented by a bit string. Males produce songs by copying their inborn template with a number of mistakes, insertions and deletions, depending on their quality. The females in turn assess male quality by comparing their own
template to the male song. Since the number of offspring of a female (inheriting the used templates) is relative to the quality of the male, the female choosing a high quality male has an evolutionary advantage. The model shows that rhythmic songs are more discriminating with regard to the type of mistakes that male birds make in comparison to non-rhythmic songs. Moreover, rhythmic songs are more stable against evolutionary drift than monolithic templates. These results indicate that a rhythm-prefering female receives more offspring, thereby implying that rhythmic signals do seem to be useful quality indicators and may have evolved as indicator of mate quality by their conspicuous properties.

6.2 Discussion

Modeling biological phenomena has its limitations per se, as is emphasized in chapter two. Due to unavoidable simplifications a model cannot prove how something has actually evolved in real life. This thesis provides a possible account of the evolutionary process, indicating a set of sufficient assumptions for a trait to develop in silico. The correctness of these assumptions cannot be verified through the use and outcomes of a model; however, by restricting the number of assumptions as much as possible without cramping effects on the outcomes, we can curtail the necessary and sufficient conditions for the phenomenon under investigation, which in turn can lead to verifiable predictions. However, some limitations that led us towards the results obtained might have led us away from reality. As mentioned in chapter 2, the WYWIWYG problem (What You Want Is What You Get) lies just around the corner in evolutionary modeling. Rooting the model in biological data cannot prevent us, eager for rhythmic results, from becoming blind to the artifacts and limitations in our approach.

6.2.1 Limitations

A first and obvious artefact of the model is that the way the females judge the singing procedure is quite harsh. Recall that with the basic, symmetrical judging matrix, they will dislike a male the most if he shifts his template one time step forwards directly at the beginning. This is a strange side-effect: at least to human ears, it would not make that big a difference if someone starts his serenade one time step later. One way to circumvent this would be by simply disregarding zeros at the beginning, and adding pauses at the end. We chose not to do so because an unintended bias towards zeros at the end would occur. Besides, the problem only occurs in some of the possible judging matrices and only when a zero at the beginning is inserted. However, a more general problem is that intuitively it seems that in nature, a mistake occurring later in the sequence would violate the expectation more thoroughly, because the female has had a chance to build up an expectation of the next sequence of notes.
Circumventing the problem would require completely different assumptions on the female judging device. For instance, the implementation of a recording device would solve the problem effectively, starting to record at a note for ten time-steps, and consecutively mapping the recorded template on the rest of the song. This way however we would exclude coevolution and only seek for regular patterns in the male songs.

Another shortcoming of the model is that the initial measure of rhythmicity (number of occurrences times the length of the generator) was replaced by a more prosaic notion: the number of alternations in a pattern. This formed a sensible definition for our purpose, but excluded other forms of rhythm from investigation, for example [1101101101]. Whether these more complicated forms of rhythm would evolve could be investigated as well, for instance by means of different preference tables or different sorts of mistakes, but lies beyond the scope of this thesis.

To leave out song learning completely was a gross simplification, but allowed for simple, heritable templates. The judging procedure is implemented by comparing an inborn template to the song heard, and this resembles the calls and judgments of nonsongbirds better (cf. doves, Ballintijn and Ten Cate, 1999). The largest simplification however was of course to define birds as bare singing devices, notwithstanding peacocks, bowerbirds, and manakins, which show off with different sexy traits like tails, bower capacities or jumping muscles.

On the other hand, some simplifications might have led us away from rhythmic results. By limiting our model to behavioral outcomes, we did not deal with the mechanisms underlying those outcomes and we might have missed some other factors in evolution. Indicators rely on general mechanisms that may have evolved for another purpose, preceding the indicator. One artifact of the model that might have complicated the evolution of the rhythmic signals in our framework is that we overlooked other selective pressures or historical contingencies favoring rhythm. One plausible source of such historical contingencies exists at the neural level, where preexisting central pattern generators originally used for other purposes such as locomotion may have been exapted for new use in creating rhythmic signals. Such rhythmic circuits may also have led to sensory biases in the receivers to prefer regularly repeating displays.

6.2.2 Possible extensions

Most of the suggestions for further research are based on small steps towards more realistic models. There is however always a balance between tractability and realism of the model. A promising direction in which to extend our model would be to build song templates in both males and females on top of a neural substrate in which rhythmic circuits could naturally arise, for instance
continuous-time recurrent neural networks. Some of these have already been explored in the context of music production and perception (see Griffith and Todd (eds.), 1999).

Another interesting extension would be to define the choir spatially, to see if subpopulations arise through assortative mating. Studies on this subject usually work on territory selection (see for instance Ellers and Slabbe Koorn, 2003), but certainly even more intriguing effects could be obtained by including sexual selection in such a spatial model.

Although the model was not based on one particular species but rather on a general 'prototype bird', still some predictions or at least suggestions for research can be made. Evidence for or against the mate-quality-indicator hypothesis for rhythm will come with investigations of the heritability of the underlying neural basis for the production and perception of rhythm. This in turn requires further unraveling of the neural structures leading to rhythmic behavior and of the genetic and environmental factors that can impact on the production of rhythmic signals. Indeed Riebel and Slater (2003) found very recently that temporal variations in song depend on quality of the singer. If a lack of noise (whether on a developmental, neural, or condition level) is indeed correlated with a fitness advantage for a bird, a rhythmic song might be the sonic equivalent of the peacock’s tail through which a male can reveal his mate quality.

A final line of research could extend our findings towards testable predictions about courtship in humans. It would be interesting to see how people behave when they are asked to tap randomly in order to get chosen by the other sex, or whether they use drum computers in order to get chosen for the European song festival.
**Final thought**

When Darwin returned from his Voyages with the Beagle, during which he formulated his natural selection theory as an explanation for the origins of species, he could not stop thinking about the purposeless ornamentation of animals. He was so annoyed by useless peacock tails that they made him nauseous every time he saw them. I must admit that writing this thesis also had its influence on my daily life. Although not that physically, the croaking of a crow still annoys me.

---

4 "The sight of a feather in a peacock's tail, whenever I gaze at it, makes me sick!"

--Charles Darwin, in a letter to botanist Asa Gray, April 3, 1860
6.3 References


Leitner, S., Catchpole, C.K. (2002). Female canaries that respond and discriminate more between male songs of different quality have a larger song control nucleus (HVC) in the brain. J. Neurobiol, 52, 294-301.


Appendix A

The discriminability scores for some of the 1024 possible templates, computed with on average one deletion or insertion.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>0101010101</td>
<td>0.4306899999999996</td>
</tr>
<tr>
<td>1010101011</td>
<td>0.42892</td>
</tr>
<tr>
<td>1101010101</td>
<td>0.41836</td>
</tr>
<tr>
<td>1010101001</td>
<td>0.4134299999999996</td>
</tr>
<tr>
<td>1010101101</td>
<td>0.4126999999999996</td>
</tr>
<tr>
<td>1010110101</td>
<td>0.4124100000000005</td>
</tr>
<tr>
<td>1011010101</td>
<td>0.41208</td>
</tr>
<tr>
<td>1010100101</td>
<td>0.40743</td>
</tr>
<tr>
<td>0101011011</td>
<td>0.40518</td>
</tr>
<tr>
<td>0010101011</td>
<td>0.4042900000000004</td>
</tr>
<tr>
<td>0110101011</td>
<td>0.40278</td>
</tr>
<tr>
<td>0101010011</td>
<td>0.40155</td>
</tr>
<tr>
<td>1001010101</td>
<td>0.4013099999999994</td>
</tr>
<tr>
<td>1010010101</td>
<td>0.4003199999999995</td>
</tr>
<tr>
<td>0101101011</td>
<td>0.3991099999999996</td>
</tr>
<tr>
<td>1101011011</td>
<td>0.39846</td>
</tr>
<tr>
<td>1110101011</td>
<td>0.3953</td>
</tr>
<tr>
<td>1101010011</td>
<td>0.39447</td>
</tr>
<tr>
<td>1010100111</td>
<td>0.39404</td>
</tr>
<tr>
<td>0000011111</td>
<td>0.17682</td>
</tr>
<tr>
<td>1111111111</td>
<td>0.16515</td>
</tr>
<tr>
<td>1111110000</td>
<td>0.12972</td>
</tr>
<tr>
<td>0000000000</td>
<td>0.0</td>
</tr>
</tbody>
</table>