UNIVERSITY OF
PLYMOUTH

# THE DEVELOPMENT AND APPLICATION OF COMPUTATIONAL MULTI-AGENT MODELS FOR INVESTIGATING THE CULTURAL TRANSMISSION AND CULTURAL EVOLUTION OF HUMPBACK WHALE SONG

BY

**MICHAEL MCLOUGHLIN**

A THESIS SUBMITTED TO THE UNIVERSITY OF PLYMOUTH
IN PARTIAL FULFILMENT FOR THE DEGREE OF

**DOCTOR OF PHILOSOPHY**

SCHOOL OF HUMANITIES AND PERFORMING ARTS
UNIVERSITY OF PLYMOUTH

IN COLLABORATION WITH
UNIVERSITY OF ST. ANDREWS, SCOTLAND
AND
QUEENSLAND UNIVERSITY, AUSTRALIA

**SEPTEMBER 2017**

*For my Mother and my Father.*

*Go raibh míle maith agaibh.*

# Authors Declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award without prior agreement of the Doctoral College Quality Sub-Committee.

Work submitted for this research degree at the University of Plymouth has not formed part of any other degree either at the University of Plymouth or at another establishment.

This study was financed with the aid of a studentship form the Leverhulme Trust and carried out in collaboration with University of St. Andrews, Scotland, and Queensland University, Australia.

The following external institutions were visited for consultation purposes: Queensland University, Australia.

Publications:

***Using agent-based models to understand the role of individuals in the song evolution of humpback whales (Megaptera novaeangliae)***

Michael Mcloughlin, Luca Lamoni, Ellen C. Garland, Simon Ingram, Alexis Kirke, Michael J. Noad, Luke Rendell, and Eduardo Miranda

*Music & Science*

**Presentations at conferences:**

***Adapting a Computational Multi Agent Model For Humpback Whale Song Research For Use As a Tool For Algorithmic Composition.***

Mcloughlin, Michael & Lamoni, Luca & Garland, Ellen & Ingram, Simon & Kirke, Alexis & Noad, Michael & Rendell, Luke & Miranda, Eduardo.

*Conference: Conference: Sound and Music Computing 2016, At Hamburg*

***Preliminary Results From A Computational Multi Agent Modelling Approach To Study Humpback Whale Song Cultural Transmission.***

Mcloughlin, Michael & Lamoni, Luca & Garland, Ellen & Ingram, Simon & Kirke, Alexis & Noad, Michael & Rendell, Luke & Miranda, Eduardo.

*The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG11), At New Orleans, LA, USA, Volume: S.G. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Feher & T. Verhoef (eds.)*

***Preliminary results from a computational multi agent modelling approach to study humpback whale song cultural transmission***

Lamoni, Luca & Mcloughlin, Michael & Kirke, Alexis & Ingram, Simon & Garland, Ellen & M, Noad & Miranda, Eduardo & L, Rendell.

*Association for the Study of Animal Behaviour (ASAB) Winter Meeting, At Zoological Society of London, Regents Park, 2015*

**Word count of main body of thesis:** 43,058

Signed ……………………………………

Date ……………………………………

# Acknowledgements

This project spanned three Universities and has given me the chance to travel the world and work with truly amazing people. These people have shown how I should conduct myself not only academically, but also personally.

Professor Eduardo Miranda. Thank you for your support, trust and advice. Your encouragement and confidence really helped. In particular, thank you for being understanding during the more personally difficult times of these few years. Your understanding and patience really helped and shows how academics should conduct themselves. In a time when PhD students seem to be more stressed than ever, your support to your students and their well being has been truly inspiring. Above all, I want to thank you for that.

My second supervisor Dr Alexis Kirke. Thank you for passionate discussions regarding multi agent models, your interest in the project, and being there for any mathematical problems I encountered! Your good humour at lunch breaks was always welcome, as well your policy not to discuss work related matters during break time!

My third supervisor Dr Simon Ingram. Thank you for providing an escape from the world of coding and providing a high level of understanding in regard to all matters relating to cetaceans and culture (and for showing me the best pubs in St Andrews!).

Thank you to my research partner, and closest collaborator, Luca Lamoni. I know I could be challenging to work with, and I thank you for your eternal patience. I'm looking forward to having a beer (or several) with you after you submit. It is an honour to cross the finishing line of formal education with you.

Dr Luke Rendell. Thank you for everything!! Working with you has been a true pleasure. I would particularly like to thank you for your hospitality during my visits to St Andrews. Your work really gives me something to aspire to.

Special thanks need to be given to Dr Ellen Garland. Your insights into humpback whale song were illuminating. I especially want to say thank you for the pep talk you gave me on a train from St Andrews from Edinburgh. Confidence was at an all time low at that point, and your encouragement helped me get back into the fray. Thank you!!!

Professor Mike Noad. Thank you for your hospitality during the field season and my visit to your lab. Your passion for cetacean research is rivalled only by your good humour and kind nature.

Thank you to everyone at the ICCMR. Federico Visi (it was a pleasure to collaborate with you), Rodrigo Schramm, Nuria Bonet, Dr Edward Braund, Pierre-Emmanuel Largeron, Aurelien Antoine, Jared Drayton, Ben Payne, Dr Joel Eaton, and Dr Duncan Williams. Than you for the great lunches, the lovely pints, and fantastic company. It was a pleasure to be in the same lab.

Special thanks to everyone at the Cetacean Ecology and Acoustics Lab at Queensland University. Thank you to Jenny Allen for driving me to Gatton three days a week, giving me a couch to crash on, and most importantly, the stimulating conversations that inspired Chapter 4 of this thesis. Thanks to Dr Ailbhe Kavanagh for making sure this clueless acoustician was prepared for his first day out at sea and hosting the outdoor cinema club, to Aoife Foley for providing countless Foley facts and never ending craic, to Zeus for looking after me while I was at sea and lifts to Gatton, Amy James for her hospitality and kindness (and making sure I wasn't completely surrounded by Irish people!), and Anita Murray for her advice on acoustics and humpback whale song.

Deirdre Hatton. Thank you for everything. I couldn't possibly write it all down here. Montreal doesn't know how lucky it is. Miss you loads.

Liam O'Hare. I don't think I would have gotten through this without your advice. It's reassuring to have you in my corner. Life can be overwhelming at times, but with a friend like you by my side, I think I'll be okay.

Sam Wright. You've always been a great friend. Life can sometimes put blinders on you, and sometimes I felt like I lived only in academia. Being able to catch up with you, laugh, play music, and enjoy each other's company has been more valuable than you can imagine. Thank you.

Nicky and Tony Colleran. You are two of the brightest lights in my life. Thank you for being such great friends, and being a constant reminder that no matter how bleak things look, you can always have the craic. Your support and dedication to your friends is amazing. Thank you.

Thanks to Kevin Brennan and Laurence Durkin. You are absolute stunners! I'll see you in the spooky-verse

Thanks to Ali Hesnan, for brining some much needed Galway cheer whenever she visited Plymouth or I visited Exeter. Thank you for the late nights listening to Planxty and for keeping me sane!

All my pals in Plymouth who I had the pleasure of meeting. Benjahman Hall, Francis Elliot, Ben Wilkonson, Xander Weeks, Jack Jane, Ryan Pullin, Emily Rolley-Parnell, Ellie Hedges, Isla Brown, Christian Toulson, Alex Budgie, Amelia Greenwood, Mike Genner, Ben Greene.

Many of my best friends are scattered around the world, and while I want to dedicate a paragraph to each of you, this thesis is already too long. In no particular order, special thanks to; Mike Hyland, Mark Hatton, Alex Hannon, Alan O'Callaghan, Dylan Greaney, Erris O'Regan, Vinny Hyland, Tara Nic Anna, Tadhg Hanrahan, Tadhg Kelly, Cole Andrews, Nicky Graham, Jessi Milne, Martin Dillon, Conor McCann, Sean McGarrity, Shea Murphy, Barry O'Donnel, Paul McDonagh, Éadaoin Doyle, Emmanuelle Lydon, Paddy O'Loughlin, Aimmee Riordan, Rob

Joyce, Katie Mod, Shauna O’Connor, Claire Howley, Chloe Stafford, Conor Barry, Sally Pelzer, Kit Grier, Dr Alejandro Cervantes, Adriana Contreras, Camillie Longue, Fabian Esqueda, Juno, Charlotte Desvages, Dr Brian Hamilton, Dr Craig Webb, Matt Thomson, Ciarán Macguire, Andreas Miranda, Alex Galler, Alex Paguirigan, Sally Galler, Omer Ockuolglu, Alain Mac Pháidín. If I have missed anyone I’m really sorry!!!

Lots of love to my future sister in law Ciara. Welcome to the family!

I would like to take this opportunity to remember my godmother, Paula O’Donnell, who passed away this year. Thank you for your support all these years and encouraging me to follow my own path. I miss you dearly. Lots of love to Moya and Harry.

Finally, I want to thank the five people in my life who have always supported me. Thanks to my parents, Michael and Anne, for supporting me throughout my life. It is an honour to dedicate this thesis to you. I love you both so much. To my brothers Rory and David, and my sister Maria. A brother could not ask for better siblings. I love you all.

# Abstract

Three different multi-agent models are presented in this thesis, each with a different goal. The first model investigates the possible role migratory routes may have on song evolution and revolution. The second model investigates what social networks could theoretically facilitate song sharing in a population of whales. The third model implements a formal grammar algorithm in order to investigate how the hierarchal structure of the song may affect song evolution. Finally, the thesis attempts to reconnect the models with their origins and discusses how these models could potentially be adapted for composing music. Through the development of these different models, a number of findings are highlighted. The first model reveals that feeding ground sizes may be key locations where song learning from other population may be facilitated. The second model shows that small world social networks facilitate a high degree of agents converging on a single song, similar to what is observed in wild populations. The final model shows that the ability to recognise hierarchy in a sequence coupled with simple production errors, can lead to songs gradually changing over the course of time, while still retaining their hierarchal structure.

## TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

# Chapter 1 – Introduction

Humpback whales produce one of the most complex displays of animal culture in the animal kingdom. They sing long songs with a hierarchical structure. Generally, these songs are population specific, and are performed mainly during the mating season. Over the course of the mating season, humpback whales gradually change their songs. It is also possible for the song of a single population to be replaced by the song of another population. These phenomena are difficult to study due to the long migratory patterns of humpback whales, and the harsh conditions the animals live in. Due to the challenges associated with studying the song, this thesis presents a novel approach to allow researchers to develop hypotheses about how humpbacks may learn songs through the use of multi-agent models. These are models where multiple artificial intelligences (agents) interact with each other and their environment. Multi-agent models have received a significant amount of attention in computer music research, with researchers utilising their emergent properties to study the evolution of music.

## 1.1 Motivation

Cetaceans present some the most striking demonstrations of animal culture. This includes unique dialects for clans of killer whales (Deecke, Ford, & Spong, 2000), and coda click patterns in Sperm whales (L. E. Rendell & Whitehead, 2003). In species such as the killer whale and sperm whale, we can see direct relationships between their vocalisations, social structures, and genetics (Hal Whitehead, 1998). There have also been analyses of dolphin signature whistles that highlight a function for addressing specific individuals in their group (Janik & Slater, 1998). One species that stands out from all other cetaceans though is the Humpback whale. It demonstrates a complex, hierarchal song (R S Payne & McVay, 1971) that changes gradually over time (song evolution) (Payne, K., Tyack, P., & Payne, 1983), but can also undergo a sudden and dramatic change in the form of song revolution, where the song of one population replaces the song of an adjacent population (Noad, Michael J., Cato, Douglas H., Bryden, M. M., Micheline, Jenner, Jenner, 2000). Despite extensive field studies, many questions remain in regard to the song. What is its function? Why does it change over time? And why can a population's song suddenly be replaced with the song of another population? Field studies have examined many of these issues, (Garland et al., 2011; Garland, Noad, et al., 2013; Smith, Goldizen, Dunlop, & Noad, 2008) but due to the difficulty of studying humpback whales in the wild, many of these questions remain unanswered.

These questions are not unlike those associated with music. Why does the music of a culture gradually change over time? How can a new type of music suddenly become dominant in a culture? Why and how did music emerge? Why does music possess a hierarchal structure? Researchers from many disciplines address the issue using different approaches. A musicologist may carry out Schenkerian analysis (Fuß, 2005) in order to examine the elements that make it aesthetically pleasing to a listener. A music psychologist may carry out a series of experiments on different individuals in order to understand what they find pleasing about different types of music. Or a computer scientist may develop an algorithm to identify recurring themes in different pieces of music. Artificial Life (ALife) is an exciting field that that seeks to address natural phenomena in nature by modelling it from the bottom up, incorporating aspects associated with the natural world. One of the methods used in ALife research to achieve this is Agent Based Modelling (ABM), where individual virtual entities (agents) equipped with simple AI's carry out actions and interactions between each other and their environment in order to recreate real world phenomena. Examples of ABM include the development of algorithms that explain how a school of fish or flock of birds manage to move together (Hartman & Beneš, 2006), or how to model how language can change over time (Simon Kirby, 2002a; Steels, 2015). ALife has been used to great success in addressing some of the musical questions mentioned earlier. For example, groups of artificial agents have been combined in order to examine different ways in which music in a society of agents can interact in order to produce different unique pieces of music (Kirke & Miranda, 2009; Eduardo R. Miranda, 2003; Eduardo Reck Miranda, Kirby, & Todd, 2003a). This makes it powerful not only as an analysis tool, but also as a tool for musical composition.

In this thesis, agent based methods used to model the evolution of music are adapted in order to investigate the cultural transmission of humpback whale song. This is achieved through the development of three different models that allow the user to investigate the mechanisms that could potentially be involved in driving these systems. The first model focuses on recreating the migratory patterns of the humpback whale, and their interaction, to, and on, the breeding and feeding grounds. This model is extended to include a learning bias towards novel songs, the incorporation of production errors, and the coupling of production errors and novelty learning bias. The second model uses the same learning algorithms as the first model, but examines agents in social networks generated using the Watts-Strogatz algorithm (Watts & Strogatz, 1998) in order to investigate the potential influence of social structure on facilitating song convergence. The third model implements a formal grammar for learning and producing songs, allowing a researcher to recreate the hierarchal structure of humpback song. This model is then extended to include errors, such as the addition of new ‘sounds’, or the deletion and substitution of sounds. This is used in a vertical cultural transmission model in order to investigate how pattern recognition and production errors could influence the evolution of humpback whale song. Finally, the original migratory model is adapted to investigate how it may used to create music. This ties the project back to its origins and highlights how vocal learning models can be used to create music. It also points the way towards adapting the model for use in investigating how systems associated with vocal learning in animals can be used to investigate cultural transmission in music.

## 1.3 Research Questions

This thesis addresses a number of questions through the development of different types of computational agent based models. They are:

**RQ1:** Can methods used in computer music research be applied to the analysis of animal vocalisations?

There is a rich history of agent based modelling, and multi agent systems in computer music research. One of the goals of this thesis is to take inspiration from this field, and apply some of the methods that have been discussed in the literature to questions regarding the cultural transmission of humpback whale song. This will inform us of the suitability of these methods in animal vocalisation research.

**RQ2:** Can the phenomena of song revolution and evolution be addressed through a spatially explicit model that recreates humpback migratory patterns and fine-scale cultural transmission issues that may occur in individual interactions in humpback song transmission?

Song revolution is a very dramatic display, and there is a strong desire to assume that its cause may be the result of some kind of cognitive bias. Similarly, we may wish to understand song evolution from a similar perspective, with the changes emerging from some kind of cognitive preference. However, before we rush to these conclusions, it is necessary to remove other factors that may be responsible for the changes observed in the song. For this reason, the first model discussed in this thesis focuses on recreating the migratory patterns of humpback whales, so that we can ensure that geographical segregation and key locations for interaction are not responsible for these changes (Chapter 3).

**RQ3:** What different types of social networks could theoretically facilitate cultural transmission phenomena in humpback whales?

While there have been theories regarding the type of social networks humpback whales may possess in regards to song learning (Mauricio Cantor & Whitehead, 2013), we still do not know if there is a social network that facilitates this. Do whales learn from specific whales in a population? Or do they simply learn randomly? In order to investigate this, a model is presented where a social network is synthesised, and agents learn from each other using some of the same methods as the model discussed for RQ2. This research question is addressed in Chapter 4.

**RQ4:** Can formal grammar methods be used in the analysis of whale song, and synthesis of song when coupled with production and learning errors?

This question arose from trying to find a solution to recreating the complex hierarchy observed in humpback whale song. As we will see from the coming chapters in this thesis, methods such as Markov chains are inappropriate when it comes to modelling the hierarchy of the song. This lead to the investigation of learning algorithms that placed an emphasis on pattern recognition in strings. This investigation found the Sequitur algorithm (Nevill-manning & Witten, 1997a). The Sequitur algorithm relies on formal grammars, but it proved to be highly successful in recognising patterns in humpback whale song. A method was then developed to sample from the rules generated by the Seuqitur algorithm in order to reconstruct humpback whale songs. This sampler was then extended to include production and learning errors that may be involved in song evolution. This research question is addressed in Chapter 5 of this thesis.

**RQ5:** How can the methods developed to answer research questions 2 through 4 be used in order to create music?

This research question seeks to bring the previous research questions back to their origins and use them in the creation of music. Music, animal vocalisations, and language, have many similarities, and there are theories such as the musical proto-language theory speculate that music and language evolved beside one another (Darwin, 1871; Fitch, 2013). For this reason, there is not only a creative reasoning for revisiting these models, but also a scientific one. However, there are many technical issues associated with this, and these issues are investigated in Chapter 6 of this thesis.

Chapter 3 and Chapter 6 are partially based on papers submitted to the Evolution of Language Conference, and the Sound and Music Computing conference. Chapter 3 is largely based on a paper that is currently in review. These papers are included in Appendix 4.

## 1.4 Thesis Structure

This thesis is divided into six chapters that seek to address the research questions from the previous section. Following this opening chapter, the chapters investigate the following:

**Chapter 2:** This chapter carries out a literature review on the key concepts and theories required in order to understand the models presented in the thesis. It opens with a discussion on animal culture, defining it and separating it from ideas associated with human culture. Following this, the songs of humpback whales are examined, as well as the theories that believed to be driving the phenomena of evolution and revolution. The differences between the vocalisations of humpbacks and other cetaceans are then examined. Methods for learning, synthesising, and analysing sequences of vocalisations are then presented independent from cultural transmission. Following this review of methods, a number of models and their application for researching cultural transmission in the fields of music, linguistics, and biology are discussed. Finally, the shortcomings and relationship between all these different areas is discussed.

**Chapter 3:** This presents the first agent based model developed for investigating cultural transmission in humpback whales. This is a spatially explicit model that recreates a simplified version of humpback whale migratory patterns, combined with a first order Markov model coupled with a sound transmission loss model (Migratory Model 1). Three extensions to this model are then shown, which includes a bias towards the learning of novel songs (Migratory Model 1.2), the introduction of production errors (Migratory Model 1.3), and the coupling of novelty bias and production errors (Migratory Model 1.4). A parameter exploration of Migratory Model 1 is presented. A series of experiments are then presented that recreate observations from the wild. The results of these experiments are then discussed and compared with observations from the wild. The chapter then proceeds to a discussion on how well the model recreates phenomena, its shortcomings, and insights it provides on the phenomena of evolution and revolution. This chapter seeks to address RQ1 and RQ2.

**Chapter 4:** This chapter presents a method of examining larger populations of agents and song transmission using social networks. This chapter was inspired by discussions with Jenny Allen at the Cetacean Acoustics and Ecology laboratory. The model addresses issues surrounding the model from chapter 2, by allowing the user to create various types of networks using the NetworkX package for Python ("NetworkX," n.d.). The same Markov model from Chapter 3 is used, and the possible ways in which social network structure may facilitate song convergence discussed. This chapter addresses RQ3.

**Chapter 5:** This chapter addresses the issue of the hierarchal structure of humpback song using the Sequitur algorithm (also known as the Nevill-Manning algorithm) (Nevill-manning & Witten, 1997b). The use of the algorithm as an analysis tool is discussed, before moving on to its application as a tool for song synthesis. This is achieved through the development of an algorithm that samples from the rules generated using the Sequitur algorithm. Finally, how the model behaves when coupled with production and learning errors in a simple vertical cultural transmission chain is examined. This chapter addresses RQ4.

**Chapter 6:** This chapter examines applications for the methods discussed in the previous chapters for musical applications. First, a simple adaptation of the model connected to the model for real time performance using the Max4Live and Open Sound Control (OSC) protocols. The shortcomings of this model are discussed and the chapter closes by discussing what improvements need to be made in order for the model to generate musical ideas. This chapter addresses RQ5.

**Chapter 7:** This chapter closes the thesis by highlighting the contributions presented within.

# Chapter 2 - Background

## 2.1 Introduction

In this chapter, we review the literature that is essential to understanding the design process involved in building multi-agent models for the study of humpback whale song cultural transmission.

In the first section, all the biological concepts necessary to build the model described in subsequent are addressed. It opens with a definition of culture, and discusses different aspects associated with it. This section also discusses the songs and behaviours of humpback whales, the vocal abilities of other cetaceans, and the vocal abilities of birds. The second section examines how to analyse and model sequence learning and production for animal vocalisations, human language, and music. These production methods are examined outside of a cultural transmission context. Finally, models that have a strong emphasis on cultural transmission are investigated, and models of cultural transmission for human language, music, and vocal learning in animals are compared. This allows for an understanding of the differences and similarities between these related fields. The final section is a discussion and reflection on the methods reviewed in this chapter.

## 2.2 Culture

Culture is a term that can be misleading, especially within an interdisciplinary context. For this reason, it is very important to be clear and concise about what the term culture refers to in this thesis. It is not unreasonable to assume that when discussing culture, the concepts that come to mind are those that we use to define human culture: food, music, religion, art, rituals, and customs. Here, culture does not refer to human culture.  It is instead used to define a series of shared behaviours that transferred are from one member of an animal society to another. For this thesis, we will use the definition provided by Luke Rendell and Hal Whitehead:

*Culture is information or behaviour – shared within a community – which is acquired from conspecifics through some form of social learning.* (Luke Rendell & Whitehead, 2001)

**Here, conspecifics refers to other members of the same species**. In other words, culture is the behaviours carried out by all individuals in a community that is not a result of genetic and or environmental causation, but rather emerge from learning from another member in their community. To fully understand this definition we must discuss social learning.

### 2.2.1 Social Learning and Cultural Transmission

Social learning is the transmission of new behaviours from one member of a group to another that occurs due to one member of that group teaching another member these behaviours. This can be achieved through a number of ways. For example, it may occur from a chimp observing another member of its social group using a tool (Whiten, Horner, & de Waal, 2005), and copying this behaviour. It is also possible for behaviours and information to be transmitted from a parent to a child, such as a mother polar bear teaching her cubs to hunt (Gilbert, 1999).

When information is shared between individuals in a group, it is generally termed cultural transmission. Broadly speaking, there are two main types of cultural transmission: horizontal cultural transmission, and vertical cultural transmission.

Horizontal cultural transmission is the transmission of behaviours between individuals who are not necessarily related (the chimp copying another member of his group using a tool). Vertical transmission specifically relates to the transmission of behaviours from parents to their offspring (the mother polar bear teaching her cubs to hunt).

### 2.2.2 Social Network Theory

Social Network Theory is the theoretical study of the relationship between individuals and the social structures they belong to through the use of networks and graphs (Krause, Lusseau, & James, 2009). This allows researchers to show the relationship between individuals in the society. Individuals are classified as nodes, and may represent individual animals in a wild population, a group of agents in a computer simulation, or other entities within a network. The connections are termed ties, edges, or links. These show the relationships and or interactions that connect these individuals within the network.

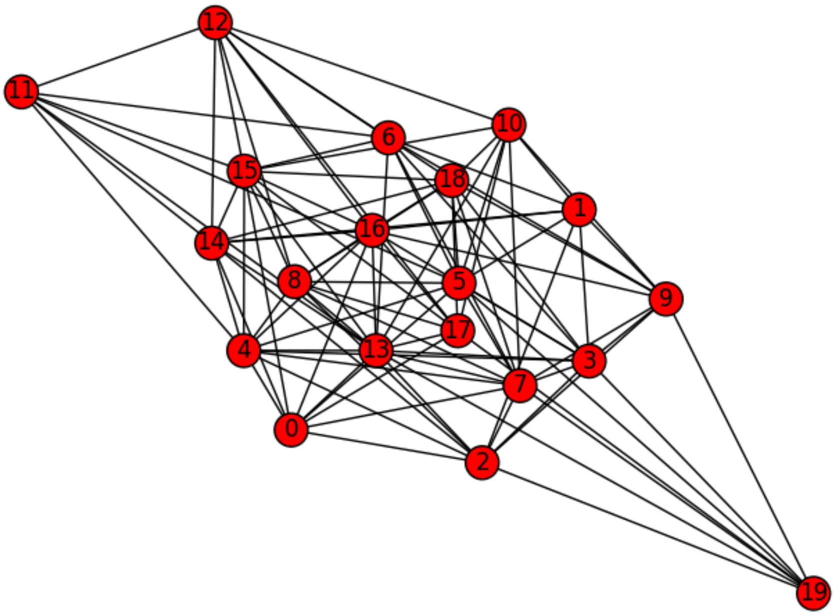


Figure 1: An example of a social network generated using the NetworkX package for Python. Here, any member may learn from another node it is connected to. For example, node 1 can potentially learn from nodes 3, 5, 7, 6, 9, 10 and 16.

In order to understand cultural transmission and social network theory, it is useful to look at an example of social network theory being applied to study an instance of cultural transmission occurring in Cetaceans. One of the most important studies done in social learning and cultural transmission in humpback whales was Allens' application of network-based diffusion analysis that showed that a behaviour, known as lobtail feeding was culturally transmitted (Allen et al., 2013). In lobtail feeding is an extension of the behaviour known as bubble net feeding (Friedlaender et al., 2011; Leighton, T. G., Finfer, D., Grover, E., & White, 2007). In bubble net feeding, a group of humpback whales co-operate in order to feed on a large school of fish. A number of whales (usually four), will swim around the school of fish while expelling air from their blowholes, creating a vortex of bubbles around the school known as a bubble net, which the fish are unable to move past. While the fish are trapped, another humpback whale swims underneath the school and up towards it, releasing a loud vocalisation as they do so. This vocalisation causes the fish to rush to the surface. Once they reach the surface the humpbacks swim up to the surface and feed on the fish. It was the observed that this feeding behaviour had another step added to it, where, after eating their prey, the humpbacks would raise their tails out of the water and slap the surface of the water. The occurrence of this behaviour was initially infrequent and was recorded by Mason Weinreich whenever it was carried out. Weinrich also kept photographs of the underside of the tails that enabled each whale to be individually identified by the unique marks on the underside of its fluke, and also the shape of its tail. This was also coupled with geographic and time and date information. These behaviours were recorded over several years and stored carefully.

Allen was provided the data in order to carry out a social network analysis. She applied a method known as Network Based Diffusion Analysis (NBDA). NBDA is a method that takes into account the order, and timing, in which an individual within a group acquires a behavioural trait. This data is compared with a social network that contains information about potential social learning opportunities. These opportunities may be grooming situations, or in the case of humpback whales feeding opportunities. It may also be that the animals simply spend a large amount of time in close proximity. This allows an emphasis to be placed on socially learned traits that would spread quickly in animal groups that have strong connections in a social network. After applying NBDA to Weinrich's data, it was found that the lobtail feeding behaviour was a culturally transmitted trait.

Allen's study is interesting as it gives us some insight into the social structures of humpback whales. Unfortunately, we do not possess the same amount of information when it comes to song transmission data set used in this thesis. As a result, the resources are not at hand to apply NBDA to the recordings used in this thesis. Even if a whale is being tracked using methods such as passive acoustic monitoring (Cato & McCauley, 2006), it can be difficult to verify that the song recorded belonged to a specific whale that rose to the surface. While this unfortunate, it further highlights the importance of developing multi agent-models for researching humpback whale song. Later in this thesis, we will examine methods for developing multi agent models where specific types of social networks are used to investigate if they have a potential role in song transmission. The transmission of songs in these simulated social networks may give us insight into the social network of humpback whales during the breeding season.

As we see from Allen's study, cultural transmission has a significant impact on the behaviour of animals. One of the most interesting hypotheses is that cultural transmission does not only have an influence on the behaviour of animals, but also the genetic diversity of species. Whitehead investigated how cultural transmission affects four different species of whales with matrilineal social structures(Hal Whitehead, 1998). The term matrilineal refers to the fact that in this species, members tend to spend their entire lives with close female relatives. Members of these species also have a low diversity of mitochondrial DNA (mtDNA), a type of DNA that species inherit solely from their mother. In these species, it was found that culturally transmitted traits including vocal learning behaviours are transmitted matrilineal. Whitehead discusses how these cultural traits 'hitchhike' on these mtDNA alleles and thus decrease genetic diversity. Whitehead reinforced his argument by running multi agent simulations that found that rates of non-matrlineal transmission of dialect must not occur faster than the genetic mutation rate. Whitehead demonstrates an important point that is discussed in other writing (Simon Kirby, 2002b; H Whitehead, Richerson, & Boyd, 2002): That culture can have a direct impact on the genetics of a species. These agent-based models are discussed in more detail in the Cultural Transmission Modelling section of this chapter. An excellent overview of social networks in cetaceans can be found in (Mauricio Cantor & Whitehead, 2013).

## 2.3 Vocal Learning

Vocal learning, in broad terms, is an animals' ability to learn, imitate, and reproduce sounds it hears. Usually, these sounds are produced by other members of its species, but experiments have been carried out that demonstrate that certain animals are capable of learning synthesised sounds (Reiss & McCowan, 1993), and parrots, arguably the most famous vocal learners, are capable of reproducing the speech patterns of humans. Vocal learning is generally viewed as one of the stepping-stones towards the evolution of human language (Berwick, Beckers, Okanoya, & Bolhuis, 2012; Jarvis, 2007), and its study is considered important in evolutionary linguistics and biology. This is due to the fact that it is a useful comparative study to human language. Through the study of vocal learning researchers hope to identify whether this trait evolved independently in the animals that display it via convergent evolution, or if it originates from a trait inherited from some distant common ancestor and was lost over time in other animals via evolution. It is important to note that vocal learning is a type of social learning.

One of the most interesting aspects of vocal learning is that it is a trait displayed by very few animals. The main animals that display vocal learning include songbirds (Berwick, Okanoya, Beckers, & Bolhuis, 2011; Lachlan, van Heijningen, ter Haar, & ten Cate, 2016), cetaceans (Janik & Slater, 1998; King et al., 2013; R S Payne & McVay, 1971), bats, humans, and pinnipeds (Schusterman, 2008). There are different types of vocal learning. (Arriaga & Jarvis, 2013) divide these types of learning into three different categories They are auditory comprehension learning, vocal usage learning, and vocal production learning (Arriaga & Jarvis, 2013). Auditory comprehension learning is where an animal associates a sound with a specific behavioural response to it but does not develop the ability to reproduce the sound it has learned. A common example of this is the ability of a dog to learn the word *sit* and respond with the action of sitting. However, the dog does not learn how to produce this vocalisation. Vocal usage learning is the ability of an animal to associate and produce a particular vocalisation with a change in its environment. The example given by Arriaga and Jarvis is that of vervet monkeys learning to associate, learn, and produce specific vocalisations in response to a particular threat, i.e, different warning calls for different predators. The final type, vocal production learning, is more complex than the previous two types of learning. Specifically, vocal production learning is where an animal must rely on adjusting its vocalisations depending on its experiences. A songbird learning the song of another bird would be an example of this, as it requires the animal to adjust its vocalisations depending on the auditory input. However, there are various types of vocal production learning, and the songbird model should not be seen as the be all and end all model of vocal learning. Humpback whale song is a type of vocal production learning, and for this reason, the discussion on vocal learning will focus on animals that carry out this type of learning.

Song is a common term used to define a sequence of sounds an animal produces in vocal production learning (Kershenbaum & Blumstein, 2014). Individual sounds in a song are sometimes referred to as units, and this terminology will be used throughout the text. Songs may be highly structured like those of the humpback whale, or less structured, like the zebra chaffinch. The entire set of units an animal produces will be referred to as the animals' repertoire.

In the following section, the behaviours and vocal abilities of humpback whales are given special treatment with an in-depth discussion. It begins with a brief description of the migratory patterns that humpback whales follow, as humpback whales have been observed singing primarily on the breeding grounds (Smith, Joshua N, Goldizen, Anne W., Dunlop, Rebecca A., Noad, 2008). However, in recent years there have been observations of song on the breeding grounds (Vu et al., 2012), and this is hypothesised to have an impact on the cultural transmission of the song (Garland, Gedamke, et al., 2013b). Following this the vocal displays of other baleen whales, and mysticetae (toothed) whales, are reviewed. Songbirds are also discussed, as they exhibit similar traits to humpback whale song, including a song that is believed to be involved in sexual interactions, and also possessing a hierarchal structure.

## 2.4 Humpback Whales

### 2.4.1 Migratory Behaviour

Humpback whales undertake one of the longest migrations to their breeding grounds exhibited by any animal (Dawbin, 1966). Humpback whales are highly consistent in their mating and feeding ground locations (Baker et al., 1986). Humpback whales learn their migration pattern from their mothers (Rosenbaum et al., 2009). There are many different populations of Humpback whales across the world, but we will focus only on the migratory route of the populations relevant to the Eastern Australia population. Humpbacks in this population spend their time on feeding grounds in the Antarctic during the summer and migrate towards their respective breeding grounds in the winter.

Schmitt et al carried out a mixed-stock analysis of humpback whales on the Antarctic feeding grounds (Schmitt et al., 2014). The populations that shared these feeding grounds were from Western Australia, Eastern Australia, Oceania and Columbia. They identified areas where the populations mixed in the Antarctic feeding grounds (see

Figure 2). Although there has been no singing recorded while on the shared summer feeding grounds for the south pacific population, singing has been recorded on other feeding grounds (Vu et al., 2012). If singing is present on the Antarctic feeding grounds where the western and eastern Australian populations, it would have strong implications for cultural transmission (Garland, Gedamke, et al., 2013b).

**Figure removed due to Copyright Restrictions**

Figure 2: This figure shows the breeding/migratory populations of the Antarctic and their respective feeding grounds (Areas IV, V, VI, and I). They are divided into pure and mixed stocks. As we can see, the populations on Western Australia and Eastern

Australia mix in Area IV. As we will see in the song section, this has implications in regard to cultural transmission. Taken from (Schmitt et al., 2014).

### 2.4.2 Song

The structure of Humpback whale song was first described by Roger Payne and Scott McVay in 1971 (R S Payne & McVay, 1971). In this paper, Payne and McVay identified the hierarchal structure of humpback whale song using spectrogram analysis. They showed that songs consisted of individual sounds called units. These units are then combined to create phrases. Phrases are then combined to create themes, and themes are sung in a specific order to create a song. Songs last between ten and twenty minutes and may be repeated multiple times to create what is known as a song session. Song sessions may continue for several hours, with longest song sessions observed being 22 hours long (Winn & Winn, 1978). The hierarchal structure of humpback whale song was confirmed by a study (Suzuki, Buck, & Tyack, 2006) that used information theory (Shannon, 2001) to highlight the hierarchal structure of humpback whale song. Songs are performed on both the breeding and feeding grounds, but humpbacks sing more frequently on the breeding grounds than on the feeding grounds (Garland, Gedamke, et al., 2013a; Vu et al., 2012). It should also be noted that certain whales in populations have been observed to perform songs that do not have a structure and do not adhere to the song being performed by the rest of the population. These are termed aberrant songs (Frumhoff, 1983) and appear to be very rare.

Humpback whale song gradually changes over time, in a form of cultural evolution. (Payne, K., Tyack, P., & Payne, 1983; K. Payne & Payne, 1985; Roger S Payne & Mc, 1971) This gradual change is known as song evolution and what drives this is still unknown. Another more dramatic change in song was first observed in 2000, (Noad, Michael J., Cato, Douglas H., Bryden, M. M., Micheline, Jenner, Jenner, 2000), where the song of the western Australian population of humpback whales completely replaced the song of the eastern Australian population over the course of two years. This dramatic change was termed song revolution. It is unique due to the fact that it was not caused by a mass influx of immigrant whales from the western Australian population to the eastern Australian population, and the reasons for its arrival is unclear. Garland noted that since song has been observed on feeding grounds in other populations, the presence of song in the feeding grounds of the Eastern Australian population and its adjacent populations may have a strong influence on the cultural transmission of humpback song (Garland, Gedamke, et al., 2013a). This revolutionary behaviour was not an isolated incident and studies confirmed that the song continued to travel eastward, taking over the songs of the populations in New Caledonia, Tonga, American Samoa, Cook Islands, and French Polynesia. Despite occurring throughout these populations, this behaviour has not been observed outside of the populations in the pacific. Cantor (Mauricio Cantor & Whitehead, 2013) briefly discussed humpback whale social structure in regards to song evolution and revolution, describing the different populations in the pacific as nodes. He highlighted that the reason for a uni-direction flow in regard to song revolution from west to east (eastern Australia to French Polynesia) was uncertain, but suggested that geographical constriction of the North Pacific migration route may connect humpbacks that use separate breeding grounds. This suggests that some type of geographical contraint may be responsible for song revolution. That this behaviour may not have been observed in populations outside the pacific bay be due to these constrictions not being present.

The exact function of humpback whale song is a matter of debate. However, it has been observed that only male humpback whales perform songs (Glockner, 1983). This suggests that it serves some type of sexual function. Whether this is to mediate interaction between males (Smith, Joshua N, Goldizen, Anne W., Dunlop, Rebecca A., Noad, 2008)(Darling, Jones, & Nicklin, 2006), or to serve as a sexual display to attract females (Tyack & Whitehead, 1983) is not clear. It may be possible that it fulfils both these roles, and research presented by Anita Murray suggests that the content of the phrases may distinguish between two different phrase types, simple and complex. A number of acoustic features were used to distinguish between simple and complex phrases, with simple phrases characterised as phrases sung by multiple individuals, with a smaller number of units and unit types, and longer average intervals between units, a lower and narrow range of frequencies, and shorter or longer durations than complex phrases. Complex phrases were more variable, and contained phrases that were unique to the individual. By distinguishing complex and simple phrases, Murray suggests that simple phrases may be used for male-to-male interaction and attracting mates, while the complex phrases would be a courtship song only performed by males for females (Murray, Antunes, Dunlop, & Noad, 2015). This theory is particularly compelling when we consider that low frequency sounds that make up simple phrases travel easily through water, as sound loss transmission is a function of intensity and frequency (*Fundamentals of Underwater Sound*, 2008), with high frequency sounds generally not travelling as far as low frequency sounds. In the section *Other Baleen Whales,* we will see that researchers have suggested a similar dual function for bowhead whales, who generate a complex song, with a similar wide frequency range (O. M. Tervo et al., 2012).

The sound production mechanisms at work in humpback whale song are unclear. It was thought that humpback whales and baleen whales in general had no sound production mechanism that could be compared to the human vocal folds, and the method used to produce sound was not understood. However, autopsies of several baleen whales have identified a structure known as the U-shaped folds (Reidenberg & Laitman, 2007). Reidenberg & Laitman put forward a theory of how these U-shaped folds may vibrate. This theory is currently being reviewed through the creation of computational physical models (Adam et al., 2013) and bio-mechanical models (Cazau, Adam, Aubin, Laitman, & Reidenberg, 2016). In their 2016 paper, Cazau et al. suggested that the non-linearities of the song produced by a whale could be an indication of the singing males body size and physical fitness. This would make these non-linearities an important aspect of the song, and connect it to the theory that song is involved in inter-sexual interactions. It also has interesting implications when we consider the complex/simple phrase types theory put forward by Murray.

### 2.4.3 Social Sounds

Humpback whales have been observed to make sounds that do not appear to be related to songs. These sounds have been referred to as social sounds (Rebecca A Dunlop, Noad, Cato, & Stokes, 2007). In this 2007 study, Dunlop et al. identified 34 types of sounds used in its social repertoire. Of these 34 sounds, 21 appeared in the song from that year, but there were 13 different sounds that were not found in the song and remained consistent over a period of three years.

A behaviour that appears to have a counter-intuitive purpose is that of breaching. Breaching is where a humpback whale rises to the surface of the water and leaps out of it. It is scientifically termed *surface-active behaviour*. Research suggests that it actually functions as a social sound (Kavanagh et al., 2016). Kavanagh suggests that it allows humpbacks to communicated between distant groups, since the probability of observing breaching decreased when the nearest whale group was as far as 4'000 meters. The breaching effectively working as an impulse being sent through the water.

## 2.5 Other Baleen (Mysticeti) Whales

Many Baleen whales produce songs. Here, we review the characteristics of these songs, and what differentiates these songs from those of the humpback whale. Not all baleen whales sing, and only whales that sing are reviewed here. Their social sounds are not discussed.

### 2.5.1 Bowhead Whales

Bowhead whales produce songs that share similarities to humpback whale song. Bowhead whales are much more difficult to study than humpbacks, due to the fact they remain in arctic waters during the winter. Firstly, they are believed to be involved in sexual selection, although, like humpback whale song, it is unclear if this is mediating interactions between males or females. They also have a wide frequency range. (O. M. Tervo et al., 2012) suggested that the reason for this wide frequency range is due to the role of the vocalisations in mating. This broad frequency range is shown in

Figure 3. Tervo also proposes a hypothesis that the wide vocal range of the bowhead whale is due to low frequency and high frequency vocalisations playing a dual role in interactions, with low frequencies being responsible for signalling the whales location to other listening whales, and higher frequencies containing information about the identity of the singer. This hypothesis is similar to that

proposed by Murray discussed in the Section 2.4.2, of this chapter, who also suggests a dual function for the vocalisations.

**Figure removed due to Copyright Restrictions**

Figure 3: The frequency range of different whales (X axis) versus their weight (Y axis). Taken from (O. M. Tervo et al., 2012).

Another similarity between the songs of bowheads and humpbacks is the structure of the song. Like humpbacks, bowheads produce songs that have a hierarchal structure. This means that units are used to created phrases and are repeated in sequence (Clark, 1990). However, bowhead songs last for a much shorter time, and cycles last up to a minute.  Another unusual aspect of the songs is that different songs have been recorded on single breeding grounds. Research suggests that female bowheads have been observed singings songs (O. Tervo, 2011).  This is very different to humpbacks, where only males have been observed singing. Another unusual feature of bowhead whale songs is their ability to produce two sounds simultaneously. (O. M. Tervo, Christoffersen, Parks, Kristensen, & Madsen, 2011). Bowhead whales also demonstrate changes in their song over the breeding season (O. M. Tervo, Parks, & Miller, 2009) and from year to year (O. M. Tervo, Parks, Christoffersen, Miller, & Kristensen, 2011). This is similar to the evolution observed in humpback whale song. Like humpbacks, what drives these changes is still a matter of debate.

### 2.5.2 Blue Whales

Blue whales also sing. Only males have been observed singing, and their singing patterns are very simple, using low-frequency pulsed calls (call types A), or tonal calls (call types B) (Oleson et al., 2007). These calls are repeated in rhythmic, repetitive sequences, or are called out individually intermittently. Researchers have identified different song types in different ocean basins, which suggests that there may be different songs for different populations (McDonald, Mesnick, & Hildebrand, 2006). Due to the simplistic structure of blue whale song, their songs do not exhibit any dramatic evolution over time in terms of song structure and units performed like bowheads and humpbacks. The main change observed over the past several years, has been a decrease in the frequency of the vocalisations (McDonald, Hildebrand, & Mesnick, 2009). While it is possible that this is caused by cultural changes, no solid evidence has been put forward to prove this.

### 2.5.3 Dwarf Minke Whales

Dwarf minke whales also produce songs. They produce an unusual metallic sounding vocalisation that researchers have dubbed the *Star Wars* sound (Gedamke, Costa, & Dunstan, 2001) due to its resemblance to the laser sound from the film of the same name. These vocalisations have a simple song like structure, and do not appear to change much from year to year. They seem to consist of three units (A-B-C) and are repeated in sequences such as (AAABC, AAABC, AAABC). At the time of writing, there is no information on the function of the song or which sex produces them.

### 2.5.4 Fin Whales

Fin whales also exhibit a type of vocalisation that can be classed as song. These are the simplest of any vocalisation created by a baleen whale. They are downwards sweeping tonal calls. Their exact function is unknown, but they are thought to serve a purpose in mating, due to the fact that only male fin whales have been observed creating these calls (Croll et al., 2002). The presence of different songs at different ocean basins (Delarue, Todd, Van Parijs, & Di Iorio, 2009) suggests that culture may be involved in explaining these differences between populations.

## 2.6 Toothed (Odontoceti) Whales

Toothed whales display vocal learning, but do not produce songs in the way that baleen whales do. Reviewing the vocal behaviour of toothed whales is important, as much of the research into cetacean culture has revolved around toothed whales. It is important to note that dolphins are part of this sub-order and their vocal learning abilities will also be briefly discussed here. There are other aspects of these vocalisations that we must acknowledge before we can create any ties between them and humpback whale vocalisations. Firstly, it is important to remember that humpback whales do not use their vocalisations for interaction in matrilineal groups such as those that will described for the killer and sperm whale, and generally live relatively isolated lives. Secondly, there are distinct anatomical differences in the animals, with baleen whales using the U-shaped folds to produce their vocalisations. The vocal system of most toothed whales is used for echolocation (Au, 2004). Despite these differences, it is important to discuss toothed whales as agent based models have been used to study their evolution in respect to cultural transmission (Maurício Cantor et al., 2015; H Whitehead et al., 2002). There are many different members in the sub-order of odontoceti, so for this reason we only examine the three members that have received the most attention in regard to their vocal learning abilities. These are the killer whale, the sperm whale, and the bottlenose dolphin.

### 2.6.1 Killer Whales

Killer Whales are perhaps the most well known of all cetaceans, famous for their black and white bodies. Research has shown that there are different sub-species within killer whales (Morin et al., 2008). Interestingly, culture is thought to play a key part in these species and their evolution. Whitehead noted that low diversity levels of mitochondrial DNA (mtDNA) could not be explained by factors such as low mutation rate (Hal Whitehead, 1998). He also noted that in order for a population to obtain this reduced level of mtDNA variance, it would require a population bottleneck of about 100 animals for 100 generations, or 1000 animals for 1000 generations. These scenarios seemed unlikely to occur given how widespread and long-lived the animals were. Whitehead demonstrated using agents based models that cultural transmission was likely the cause for this low level of genetic diversity due to cultural hitchhiking. A recent study analysed population genomic data from these different killer whale ecotypes, and estimated that the genetic diversity emerged within less than 250’000 years (Foote et al., 2016). The study reconstructed ancestral demographics and revealed that there were bottlenecks for founding events. This genetic analysis showed that there were small founder groups who went into novel niches through plastic behavioural response, and that cultural transmission of these behaviours resulted in the emergence of different species of killer whales. These studies highlight the complex cultural lives of killer whales. Killer whales are well known vocal learners (Deecke et al., 2000; Janik, 2014). Their vocalisations can be broken down into three categories, 1) sonar clicks, 2) pure tone whistles and 3) pulsed call. A pulsed call consists of individual clicks going at high rates. Pulsed calls can consist of many different elements and be extremely complex, with different patters being created. In 1991, the vocalisations of 16 different pods of killer whales in the Canadian waters of British Columbia were recorded. Members of these pods created 7-17 types of discrete calls and individuals appeared to be learning from each other. It was noted that the 16 pods formed four different types of acoustic associations, known as clans, with each clan having its own unique repertoire (Ford, 1991).

### 2.6.2 Sperm Whales

Sperm whales, another large toothed whale, also display interesting vocal displays that are believed to be cultural. They produce specific click patterns that are termed *codas* (Watkins, 1977) that consist of three or more broadband click noises that are carried out in a specific timed order. Like killer whales, sperm whales live in matrilineal societies. Males live fairly solitary lives, while there are several different layers to female social structures and they form groups that are referred to as *units* that possess a stable female membership (Hal Whitehead, Waters, & Lyrholm, 1991). The purpose of sperm whale codas is believed to be used in the identification of individuals, units and clans (Gero, Whitehead, & Rendell, 2016; Oliveira et al., 2016). Recent studies suggest that these social structures can emerge as a result of cultural transmission (Maurício Cantor et al., 2015).

### 2.6.3 Bottlenose Dolphins

Bottlenose dolphins produce a vocal display known as the signature whistle (King, Harley, & Janik, 2014). The signature whistle is a stereotyped vocalisation that is unique to the individual dolphin. One study investigated the hypothesis that signature whistles are used to maintain group cohesion. It achieved this by studying a group of four captive dolphins in two scenarios. First, individual were recorded separate from the group while they all swam in the same pool. Separations would occur when one member of the group swam into another pool. Each dolphin produced one specific stereotyped whistle when it was separated from the group, but non-signature whistles were used whenever they swam together, suggesting that the signature whistle served a purpose to maintain group cohesion (Janik & Slater, 1998).

## 2.7 Birds

Many different species of birds produce songs. The reason for producing songs differs from species to species and can also serve dual functions. In some species it is used to mark out territory, while in others it is used to attract a mate (Brenowitz, Margoliash, & Nordeen, 1997). Due to the large number of species, a full comprehensive review of bird song is not possible, so special attention is paid to species whose songs have similarities to those of the humpback. These are zebra finches, Bengalese finches, and chaffinches.

### 2.7.1 Zebra Finches

Zebra finches are an important comparative study for vocal learning due to the ease with which the animals can be raised in a laboratory environment. Wild populations are also relatively easy to study compared to other species. There are similarities between the songs of the zebra finch and humpback whale. The songs of zebra finches, like the songs of humpback whales, are highly organised. Zebra finches perform what are known as *motifs* that consist of up to six different units of its repertoires, always performed in the exact same order, up to four times. The repeated motifs are preceded by a set of introduction units and followed by a sequence of ending units (Lachlan et al., 2016). Another similarity in zebra finches and humpbacks is that females do not sing (Wade & Arnold, 2004). However, it has been shown that the injection of hormones to enforce masculinity will cause females to sing (Gurney & Konishi, 1980). It appears that male song is one of the cues used for mate choice by females (Hauber, Campbell, & Woolley, 2010; Ritschard, Riebel, & Brumm, 2010). A major difference in zebra finch song is that males learn from tutors during a developmental period, and this serves as a template for their songs in the future. After this developmental period, the song enters what is termed “plastic song”. These plastic songs can contain entire songs from multiple tutors, but eventually form into a single song, in a process that is known as “song crystallization” (Brainard & Doupe, 2002). In the absence of a tutor they will develop a rudimentary version of the song, with abnormally high frequencies, referred to as the “isolate song” (Williams, Kilander, & Sotanski, 1993). Interestingly however, after establishing untutored singers as the tutors of a new generation, it was found that the song stabilised after passing through four generations (Fehér, Wang, Saar, Mitra, & Tchernichovski, 2009). This is different to humpback whales, where the song is constantly updated, and it is not known to undergo a specific developmental period. The only opportunity where this may potentially occur is when humpback calves and their mothers are briefly escorted by a male escort during the mating season (Tyack & Whitehead, 1983). It is noted that Zebra finches are songs are prone to errors, and that these errors may prevent cultural traditions being established that would lead to populations diverging (Lachlan et al., 2016).

### 2.7.2 Bengalese Finches

Bengalese finches also produce songs. The songs have been described using finite-state grammar, using transition probabilities such as Markov chains (Okanoya, 2004). Bengalese finches produce 2 to 5 notes to form a unit. Each unit is produced at a specific state transition. Transition patterns are not fixed as several other possible notes can follow a single note. Certain sequential notes form a chunk and chunks can also be arranged using Markov processes (Suge & Okanoya, 2010). This is quite different to the humpback song, which has a specific hierarchal structure. Only male Bengalese finches sing, and the songs appear to be involved in female mate choice. Female Bengalese finches appear to show a preference to songs they are familiar with (Kato, Hasegawa, & Okanoya, 2010).

### 2.7.3 Chaffinches

Chaffinches have received attention in regard to cultural evolution. Chaffinch song typically consists of a trill of a series of phrases in two, followed by a single broadband signal that is performed only once. They exhibit vocal learning and examples of animal culture. Chaffinch repertoires are small, with the number of songs they are able to learn ranging from 1 to 6, and will sing a succession of songs (Riebel, Lachlan, & Slater, 2015). Thorpe demonstrated that chaffinches raised in the lab away from adult males produced aberrant songs. However, if these birds were exposed to tape recordings of a wild chaffinch, they would eventually produce normal songs matching those found in the recording (Thorpe, 1958, 1961).

## 2.8 Sequence Learning, Synthesis, and Analysis

Now that the vocal learning abilities of different species have been reviewed, we can move on to discuss methods used in learning, synthesising, and analysing animal vocal sequences using computational methods. There are several methods to achieve this, and what method is used is usually determined by the format of the data. For this reason, this section begins with a discussion on the collection and representation of the database used in this project. For this project, it was decided that a symbolic method would be used for representing individual units of humpback whale song. This means that in our data, a Unicode character is used to represent a single unit of song. This limits the number of methods that we are able to apply to our database. As a result, only methods that are applicable to our database are reviewed here.

There are a number of methods in the modelling of sequence learning and production in the fields of biology, linguistics, and music. Many of these methods overlap with each other but have trade-offs that make them more suitable to a particular field. Biological models are more concerned with the accurate reproduction of sequences close to real world systems. Linguistic systems seek to explain phenomena specific to language such as recursion, hierarchal structure, and the mapping of semantics. Musical systems tend to take an approach that emphasises musical aesthetics and creative results over accuracy. Due to the high amount of overlap, each method is approached independently, and its context in each of these fields highlighted. It should be noted that this section examines sequence learning and production independently of cultural transmission. In each of the examples given here, we introduce each model and give an example of it in use. The applications of the models are also discussed in relation to biology, linguistics, and music.

### 2.8.2 Automata Theory

Methods used in sequence learning and production, are tied very closely to theories associated with Automata Theory, and a models capabilities is often used to determine the type of automaton that it may be classed into (Hopcroft, Motwani, & Ullman, 2001). For this reason, we begin this following section with a brief introduction to Automata Theory. This section is dedicated to defining what automata are and classifying the different types of automata.

Automata theory, broadly speaking, deals with automated processes and the application of computation and logic to simple, defined machines known as automata. These machines may be simulated on a computer, theoretical, or even mechanical. Automata theory allows computer scientists to address issues such as what machines are capable of computing, and how it may be computed. Automata theory also allows scientists to theorise and create models of the behaviour of discrete systems.

Generally speaking, an automaton consists of inputs, output, and states. Inputs are generally sequences of symbols from a finite set (E.g. An alphabet). Outputs are sequences of symbols from another finite. States are from a different finite set, Q, and the definition of a state depends on the type of automaton.

There are four major families they are as follows:

- Combinational logic
- Pushdown automaton
- Finite-state automaton
- Turing Machine

This following section gives a description of the Turning Machine, and its most important properties. Following this, a brief definition of the other major families is also given. Later, when discussing the different methods of sequence learning and production, each of the methods discussed will be classified into one of these major families.

Combinational logic is a method where the output of the automatons is dependent only on the inputs of the system. An example of combinational logic would be Boolean algebra, which is the basis of digital computation. Here, the output of the automaton is only based on the present input, and it has not memory in regard to the previous history of the input.

A Finite-state Automaton is a machine that can be in a number of finite states. It can only be in one state at a time and it may transition from one state to another given a certain triggering condition. This is known as a *transition.* Depending of their current state, they may output a symbol. This is known as an *emission.* Unlike combinational logic, a finite-state automaton can have memory, but it is limited by the number of states it has.

A Pushdown automaton differ only from finite-state automatons in that they employ a stack, that they are able to manipulate in order to control their actions. A stack is a memory type with just two type types of operations, push and pop. When you push a stack, you are simply adding an element to its memory. When you pop a stack, you remove the most recently added item from the stack. However, a pushdown automaton is limited in that it may only reference what is at the top of the stack, giving it a limited memory.

The Turing Machine is a concept introduced by British computer scientist Alan Turing in 1936 (Turing, 1936). The simplest Turing Machine is described by Turing as a machine that is equipped with an infinite piece of tape that serves as its memory. This tape is divided into cells (or "squares") where a symbol may be written or read using a tape head. A state register that stores what state the machine is currently in. There is also a finite table of instructions that tells the machine what to do, given the current state the machine is in and the current symbol being read. For example, it may tell the machine to erase or write a symbol, then move the head either left or right. It then assumes the same state or a new state. They are capable of solving problems that the previous automata cannot if they have access to an unlimited amount of memory.

Turing showed that given any algorithm, it is possible to construct a Turing Machine that is capable of simulating that algorithm if the machine has infinite memory. Due to this fact, it is important as a theory for computation, in which a machine is said to be Turing Complete if it is capable of solving any algorithm just as a Turing Machine would. It is important to distinguish Turing Machines from the Church-Turing thesis that states that any algorithm is computable by a human with unlimited resources following an algorithm, only if it is computable by a Turing Machine.

### 2.8.3 Markov Models

In biology, the application of Markov processes to model song production has been prevalent. Markov models are a class of methods that rely on probability to analyse and produce a sequence. They achieve this by calculating the likelihood an observation occurs and using these likelihood calculations in order to synthesise new sequences. Markov models were considered a plausible model for human language (Shannon, 2001) until it was demonstrated that no finite state Markov process was capable of modelling human language (Chomsky, 1956). Although they were proven unsuitable for the modelling of language, they did continue to gain considerable support as a theory for modelling the vocal sequence production in animals (Okanoya, 2004). Markov models are considered to be a type of Finite-state Automata (FSA) where the output triggering condition is controlled by probability.

While the application of Markov chains is beneficial for the analysis of sequence production in animals, it should be noted that they have many shortcomings. A variety of Markov Models and their capability of reproducing animal vocal sequence were examined in (Kershenbaum et al., 2014) using Levenshtein distance analysis (Garland et al., 2012). In order to achieve this, Kershenbaum et al. trained Markov models using a variety of different animal sequences, and synthesised new vocalisations based on them using a variety of sequence synthesis methods. These new sequences were then analysed using Levenshtein distance. His results found that the non-Markovian models returned a lower Levenshtein distance than Markovian models.

Another major problem with Markov models is that they require a large amount of data in order to accurately synthesise complex sequences. In order to capture the hierarchal structure of something as complex as humpback whale song requires a high order Markov chain. These two matters are connected, as the order of the Markov chain increases more data is required in order to train the model. Suzuki et al addressed these issues (Suzuki et al., 2006), and calculated that in order to accurately train a first-order Markov model would require 400-800 units of song. A second order model requires 8000 – 160,000 units. The amount of data required in accurately training these models suggests that Markov models are not an appropriate approach to modelling the hierarchal nature of songs. Given that humpbacks do not require this large amount of data in order to accurately learn a song also points towards it being an unsuitable method at the unit level.

Despite these shortcomings, we will find that Markov models do have many benefits, especially in the domain multi-agent modelling. They are simple models, and computational inexpensive. While they are inappropriate for modelling songs at the unit level, we will see in Chapter 3 that they are adequate for modelling the sequence of themes in a song. Furthermore, they serve as a useful placeholder for more complex models that may be introduced later on.

### 2.8.4 Hidden Markov Models

Hidden Markov models are a form of Markov model in which it is possible to observe the output of the system (the emissions), but not the state the system is currently in. This is best understood by looking at an example of a Hidden Markov Model.

In a hidden Markov model, the system is assumed to have a number of states that cannot be observed (Mukherjee & Mitra, 2005; Rabiner, 1989). These states generally have an effect on what the model outputs.

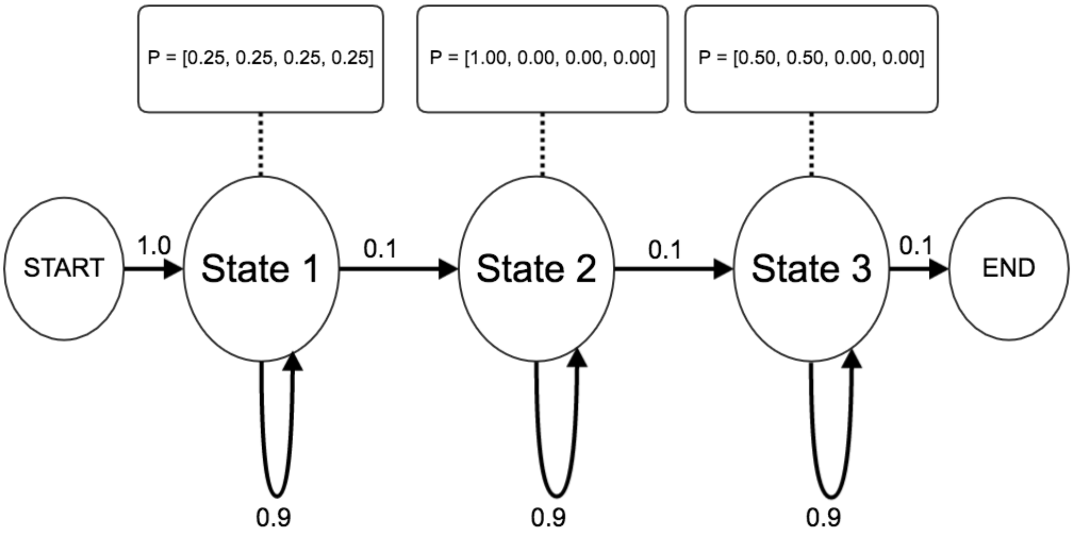


Figure 4: A hidden Markov model, with three hidden states, and probability distributions attached that affect the output of the model. As the model moves through different states, the probability vector, P, that determines what unit is output, changes.

In the example shown in Figure 4, we have three states, which cannot be seen by the observer. Attached to each of these states is a probability distribution that may be sampled from. The model begins in state 1, and samples from this probability distribution to determine what symbol it should output. After doing this, it does a weighted coin toss to determine if it stays in the same state (90% chance of occurring), or move to the next state (10% chance of occurring). As we can see, each state has a different probability distribution attached to it, so depending on the state of the system it will return very different outputs. A uniform distribution for state 1, only the first symbol for state two, and an equal probability of returning symbol 1 and 2 if it is in state 3. By analysing the frequency distribution of a sequence the model produces in chunks, we would be able to determine the state of the system.

Hidden Markov models have been used in biology for gene finding (Eddy, 2004), and also for speech recognition (Rabiner, 1989). They have also been used to for unit identification in humpback whale song (Pace, White, & Adam, 2012) and bird song identification (Kogan, 1998). Research has been presented at conferences that suggest that humpback whales may have complex, and simple phrases, and a hidden Markov model may be used to differentiate between these two different types of phrases. (Murray et al., 2015). Hidden Markov models are probabilistic finite state automatons.

### 2.8.5 Artificial Neural Networks

Artificial neural networks may be used to learn and produce vocalisations. These methods are complex, and while they may produce desirable results, understanding the mechanisms at work in a neural network can be difficult. While a neurological understanding of animal vocalisations is desirable, it should be noted that this field is complicated in itself and understanding the roles of a neural network in vocal learning independent of factors such as cultural transmission is a difficult task, and it can be unclear if the complexity that is arising is a direct result of the processes associated with the black box nature of a neural network or is the result of cultural transmission. This relates to the concept of a 'horse' in machine learning. The term horse refers to a machine-learning algorithm that appears to be achieving a required task, but relies on irregular characteristics in order to achieve its goal (Sturm, 2014). In this regard, it is important to remember that it is theoretically possible to recreate other learning models using neural networks. For this reason, it is often more useful to design a model that follows a strict algorithm, where each step in the learning process and its effect can be examined individually. Following this, the method may be adapted to work in the context of a neural network so that issues such as plasticity of learning may be investigated. There are several different types of neural networks and for this reason we will only focus on the most the basic types. An exhaustive description of all types of neural networks is beyond the scope of this thesis.

Artificial neural networks are a method of learning inspired by the biological processes underlying neuroscience. The building block of the neural network is the artificial neuron, which is inspired by neurons found in the nervous system of humans. Figure 5 shows how an artificial neuron works.

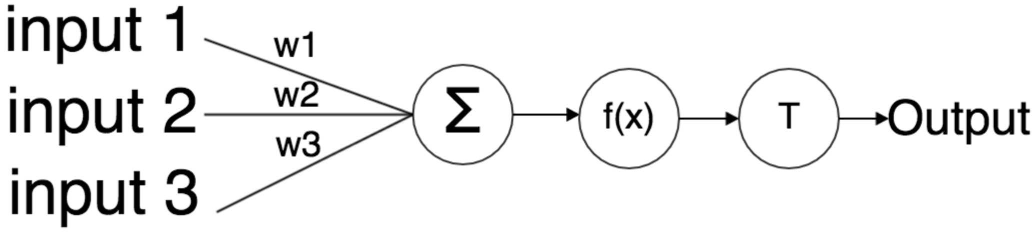


Figure 5: An artificial neuron. The inputs are multiplied by their individual weightings (w1 to w3), and then summed and multiplied by the transfer function f(x). If the resulting value is greater than T, the neuron will output a value.

An artificial neuron has any number of inputs. These inputs are multiplied by their individual weightings, and summed. Following this, the summed value is multiplied by a transfer function and if that value exceeds a threshold value, T, the neuron will output a signal. This could be the summed value itself or simply a binary value of zero or one.

Individual artificial neurons are then combined into network configurations to create the artificial neural networks. Figure 6 shows a simple artificial neural network, a feed forward artificial neural network.

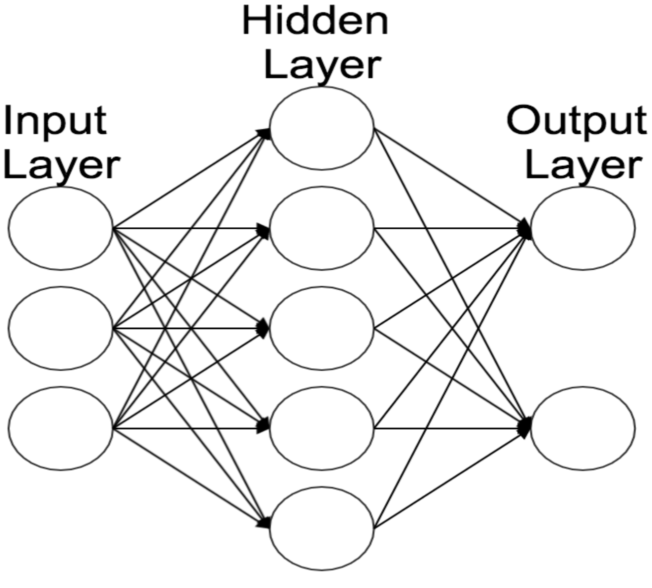


Figure 6: A feed forward artificial neural network.

In order to use a feed-forward artificial neural network, it must first go through a training stage. In this training stage, the weightings of each artificial neuron are adjusted using a form of supervised learning such as back propagation. In simplest terms, this is where data is input into the network, and the output of the neural network is compared to a target set of training data, and the weightings of the neurons are adjusted in order to move the output of the network towards accurately classifying the data. It is important to note that there are other methods of training an artificial neural network, and any decent textbook will explain these methods in detail. Feed forward neural networks have been used for several different tasks in engineering, such as recognising hand written characters (Le Cun Jackel, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, Cun, Denker, & Henderson, 1990). Neurons can be adjusted and transformed to be more complex and create spiking neural networks, where the time of arrival of the data to the input is an important factor in the classification task. These are known as Spiking neural networks and have been used in order to classify different musical instruments (Newton & Smith, 2011).

Self Organising Maps (SOM) (Kohonen, 1997), are a type of neural network that has received the most attention in regard to Humpback whale song. In the simplest SOM, input nodes are represented as points on a Cartesian plane. The input data is also distributed across a Cartesian plane. At each step of training, the algorithm calculates which input node is closest to the training data and moves it closer towards the training data. All other nodes are moved towards the training data as well, only less so. After a certain number of iterations, the input nodes have been organised around the training data. Now, whenever new data is fed in, it is mapped to the input node that it is closest to. These input nodes may then be mapped to an output matrix to visualise the classification task. Ashley Walker developed the first neural networks to analyse humpback whale song that were used by Suzuki in order to classify input units into discrete symbols (Suzuki et al., 2006), and a dictionary of humpback whale song units classified using neural networks is currently being developed (Allen, Garland, Noad, & Dunlop, 2015).

Because of the wide variety of applications artificial neural networks may be applied to, their position in the automata theory hierarchy depends on what system they are currently modelling.

### 2.8.6 Formal Grammars

Formal grammars are a method that describes a set of production rules that allow a string to be generated for a formal language. They are part of the field of Formal Language Theory that was largely developed by the work of Marcel-Paul Schützenberger, although much of the work discussed here was invented and introduced by Noam Chomsky in his paper *Three Models for the Description of Language* (Chomsky, 1956). To begin our introduction to grammars, we begin with by defining the following characteristics for formal grammars:

- **N** the finite set of non-terminal symbols. They do not appear in the final string generated by a grammar.

- $\Sigma$ the finite set of terminal symbols. The output of the grammar. This could be an alphabet or a series of words.

- **P** the finite set of production rules of the form $\alpha \rightarrow \beta$ where $\alpha$ is a string of terminals and non-terminals containing at least one nonterminal, and $\beta$ is a string of terminals and nonterminals.

- **S** the start symbol.

First, we will look at an implementation of a formal grammar, specifically a context free grammar. This will allow us to understand some of the terminology discussed above.

Consider the following example from (Rohrmeier, Zuidema, Wiggins, & Scharff, 2015) that generates a formal grammar known as a context free grammar. First we take the following sentence:

*"Either language came first or music came first"*

We can derive a context free grammar from this sentence in order to generate the following production rules as shown in Table 1.

| **Context-free grammar Production rules** | |
|---|---|
| (1a) | S ➔ either *S* or *S* |
| (1b) | S ➔ NP VP |
| (2a) | NP ➔ language |
| (2b) | NP ➔ music |
| (3) | VP ➔ V ADV |
| (4) | V ➔ came |
| (5) | ADV ➔ first |

Table 1: Formal grammar generated **from** the sentence '*either language came first of music came first'.*

By following the rules in Table 1 we are able to recreate the original sentence by applying the rules in the order of 1a, 1b, 2a, 3, 4, 5, 1b, 2b, 3. The structure of the language can be represented using a tree diagram as shown below in Figure 7.

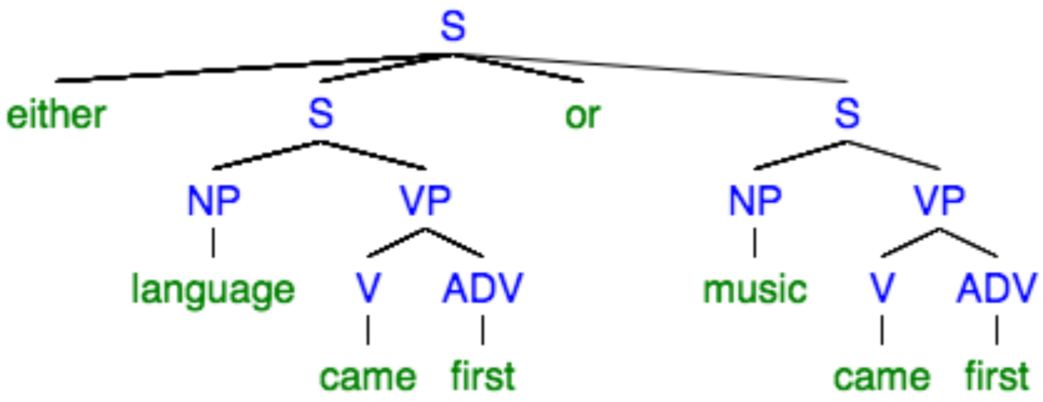


Figure 7: Syntax tree for the sentence '*either language came first or music came first'.*

There are different types of formal grammar, and it is important to distinguish between them. Furthermore, by going through a description of formal grammars in language, we can distinguish them from methods used to produce sequences of animal vocalisations. The Chomsky Hierarchy, also known as the Chomsky-Schützenberger Hierarchy, is a structure that divides formal grammars into a nested hierarchy of four parts. It should be noted that since the Chomsky Hierarchy has been established, researchers have had issues with its structure (Ojima & Okanoya, 2014) and the field itself has progressed. In order to fully understand the Formal Grammars and the methods that surround it, it is necessary to examine the hierarchy as it was first introduced.

Chomsky divides the grammars into Regular Grammars (type-1 grammars), Context Free Grammars (type-2 grammars), Context Sensitive Grammars (type-3 grammars), Recursively Enumerable Grammars (type 4 grammars). The nesting of this hierarchy and its relationship to different types of automatons is shown in

Figure 8.

**Figure removed due to Copyright Restrictions**

Figure 8: The traditional Chomsky hierarchy and its associated automatons. Automata that are capable of reproducing the language are on the right hand side. Taken from (Fitch & Friederici, 2012).

The easiest way to understand the Chomsky hierarchy is examine each aspect of it and define it, so that we can differentiate between each layer of the hierarchy and the relationship between each layer. We do this by examining the production rules (grammar) that allow these families of languages to be created.

### *2.8.6.1 Regular Grammars – Type-1 Grammars*

A language is defined to be regular if there is a finite acceptance to it. This means that it is compatible with any type of Finite-state machine, whether it is deterministic, or non-deterministic. All regular languages are context –free, but not all context-free languages are regular. In Regular grammars, the left hand side is always a nonterminal symbol, and the right hand side is restricted to being a single terminal symbol, a single terminal followed by a nonterminal symbol, or an empty string. An example grammar *G* is shown below. **N** = {**S, V**}, $\Sigma = \{\beta, \alpha\}$, and **P** is the following production rules;

- $\mathbf{S} \rightarrow \beta\mathbf{V}$
- $\mathbf{V} \rightarrow \alpha$
- $\beta\alpha$

Regular grammars have a close connection to regular expressions, but their connection is not addressed here.

#### *2.8.6.2 Context Free Grammars – Type-2 Grammars*

Context Free Grammars are grammars where the left hand side of every production rule consists only of a nonterminal symbol. This may seem like an unimportant fact, but as we will see later, not all languages can be created using context free grammars. Furthermore, another important distinguishing factor of context free grammars is that every regular grammar is context free, but not every context-free grammar is regular.

#### *2.8.6.3 Context Sensitive Grammars – Type-3 Grammars*

In a context sensitive grammar, the set of production rules for both the left and right hand may be surrounded on by terminal and non-terminal symbols. For example, a context sensitive grammar allows for production rules of the form;

- $\alpha \mathbf{A} \beta \rightarrow \alpha \mathbf{Y} \beta$
- $\mathbf{Y} \rightarrow \alpha\alpha$
- $\mathbf{Y} \rightarrow \beta \mathbf{A}$
- $\mathbf{A} \rightarrow \alpha\beta$

#### *2.8.6.4 Unrestricted Grammars – Type-4 Grammars*

Unrestricted grammars, as the name applies, have none of the restrictions that the other grammars previously discussed have. They are Turing complete and capable of generating anything the other grammars are capable of. This means that there are unrestricted grammars that are not context-sensitive, not context-free, and not regular.

#### *2.8.6.5 Formal Grammars in Biology*

Formal grammars in biology are a controversial topic, and generally each animal should be approached individually. However, in regards to the literature there has been debate over whether or not certain song birds are capable of context free grammars (Beckers, Bolhuis, Okanoya, & Berwick, 2012). Formal grammars have also been used to recreate the structure of blue whale song (Kershenbaum & Blumstein, 2014). An example of this is shown in Figure 9. Kershenbaum notes that many researchers argue that these song sequences may be recreated using simpler mechanisms. They were capable of reproducing sequences of blue whale song, these songs are much simpler than humpback whale song, as noted in section 2.5.2. At the time of writing, no methods utilising formal grammars to generate humpback whale song have been found.

**Figure removed due to Copyright Restrictions**

Figure 9: Recreating blue whale song using a formal grammar. This figure shows the production rules in the top left, the grammar tree in the centre, and application of the rules at the bottom. Taken from (Kershenbaum & Blumstein, 2014).

Lindenmeyer Systems (L-systems) are a method of formal grammar that have been used to explain the growth patterns observed in plants and fractals. They are a type of formal grammar that simulates the structure of plants by following a set of rules that guides how to draw the plant. For example, imagine a pen, that has a triplet of variables (x,y,a), where x and y are Cartesian co-ordinates and a represented the angle the pen is travelling in. Every time the pen moves, it will move forward a distance of *d,* and the angle may be incremented by the value of *b.* We then have the following symbols that represent our alphabet:

- F = Move the pen forward a length of *d.* Update the x value so that x' = x + d(cos(a) and y' = y + d*sin(a). We now have a line between the points (x,y) and (x',y').

- f = move forward a length of d, but do not draw a line between the two points.

- \+ = turn right by the angle of b. The pens state is now (x,y,a+b).

- \- = turn left by the amount of b. The pens state is now (x,y,a-b)

This alphabet is then combined with the following set of production rules;

- w = F + F +F + F
- F = F + F – F – FF + F +F – F

These steps should be carried out *n* number of times. They generate the following patterns in Figure 10, for values of n = 0,1,2,3.

**Figure removed due to Copyright Restrictions**

Figure 10: Generating fractals using L systems. Taken from (P Prusinkiewicz & Lindenmayer, 1997).

L-systems can then be extended to generate even more complex models of plants by generating a more complex alphabet rule set and adjusting the production rules. Figure 11 shows an example of modelling plants using L-systems.

**Figure removed due to Copyright Restrictions**

Figure 11: Taken from (P Prusinkiewicz & Lindenmayer, 1997), this shows three computed models of plants generated using L-systems.

#### *2.8.6.7 Application of Formal Grammars in Music*

As noted in (Rohrmeier et al., 2015), Context Free Grammars bear a strong similarity to the method known as Schenkarian analysis found in musicology as shown in Figure 12. However, Rohrmeier notes that the informal nature of Schenkarian analysis means that the complexity it captures may not necessarily be achieved using Context Free Grammars.

**Figure removed due to Copyright Restrictions**

Figure 12: An example of Schenkarian analysis example from (Rohrmeier et al., 2015).

The application of Schenkarian analysis and music has started to receive more attention, and a particularly interesting application and crossover of the two methods was introduced in (Hamanaka, Hirata, & Tojo, 2015), who combined Context Free Grammars for music generated from a computational method of Schenkarian analysis known as Generative Theory of Tonal Music (GTTM). This context free grammar was then combined with a probabilistic model, that determined the likelihood of a production rule being carried out. PCFG's have also been used to compose jazz music (Gillick, Tang, & Keller, 2009). Figure 13 shows an example of a PCFG for use in music.

---

**Figure removed due to Copyright Restrictions**

---

Figure 13: Taken from (Hamanaka et al., 2015), this figure shows a PCFG (Probabilistic Context Free Grammar) where the possibility of enacting a production rule is controlled by probability.

Formal grammars have been applied to music in the form of L-systems. This includes using L-systems to generate scores (Przemyslaw Prusinkiewicz, 1986).

#### *2.8.6.8 Computational Methods for Inferring Formal Grammars*

It is clear that formal grammars and languages are one of the most powerful methods for analysing and producing sequences. The process of generating a formal grammar from a sequence is usually carried out by a human being, who are capable of recognising the complex patterns that may emerge in a sequence due to human's natural ability to easily recognises hierarchy is sequences. This process can be arduous however, and it is desirable to automate it. Not only to save time, but also to quickly implement formal grammars for various sequences and apply computational processes to the production rules.

There are several methods for inferring formal grammars. Some of these methods are reviewed in the following section. Certain methods rely on supervised learning, or search algorithms like genetic algorithms. These methods are difficult to incorporate into a multi agent system and may produce variable results and omitted from the following section for this reason.

#### *2.8.6.9 Sequitur*

The final method in this section is only discussed briefly, as it is used in a multi-agent model that is the focus of Chapter 5. Sequitur (Nevill-manning & Witten, 1997b) is a method for inferring hierarchies from a sequence of symbols and turning them into a set of production rules similar to a formal grammar. Sequitur is unique in that it does not produce a set of production rules directly related to semantics, rather its design allows the creation of production rules to capture the hierarchy of a string, allowing the identification of the hierarchical structure easily. It achieves this by calculating the most frequently used pairs of characters in an input sequence, and replacing them with a non-terminal symbol. It then reiterates the process, looking for frequently repeated non-terminal symbols that can be replaced by another character. It repeats this process iteratively in order to generate rules that capture the hierarchy of a sequence. It is discussed in more detail in Chapter 5.

### 2.8.7 Levenshtein Distance Analysis

Levenshtein distance analysis is not capable of producing sequences. It is however a useful tool for the analysis of sequences, and was important for identifying the cultural revolution that transmitted the song from eastern Australia towards French Polynesia (Garland et al., 2011). Furthermore, it is useful as a tool to determine how accurate a sequence synthesis model is, as shown in (Kershenbaum et al., 2014). Levenshtein distance analysis is a method that determines the similarity of two strings by calculating the minimum number of insertions, deletions, and substitutions are required to transform the second string, into the first. There are a number of ways of computing the Levenshtein distance of two strings. The main method used in this thesis is the Wagner-Fischer algorithm.

### 2.8.8 What About Culture?

From this section we can see that there are several methods to learn, generate, and analyse sequences. Many of them are complex, and it is difficult to see how cultural transmission may affect them. If we used the learning and production models in multiple simple artificial intelligences and had them interact with each other, what effect would they have on the sequences they were initially trained with? How could rules that force agents to interact in a particular way change the songs they produce? In order to understand the effect of culture better, we now turn our attention to multi agent models with an emphasis on cultural transmission.

## 2.9 Multi-Agent models and Cultural Transmission

Here we examine cultural transmission models for music, biology, and linguistics. We begin the discussion by clearing up some terminology. These models are created using Agent Based Modelling and Multi Agent Systems. It is necessary to be clear about the terms used in this thesis. An Agent Based Model (ABM) is a model that attempts to recreate observations from the real world by creating agents that interact with each other and their environment in order to provide an explanation of how a system may work. Multi Agent Systems (MAS) implement agents interacting with each other and their environment in order to develop a solution a problem. This means that generally the term Multi Agent System is used in engineering, while the term Agent Based Model is more common in science. In music technology, the term MAS has been used widely as it seeks to address issues surrounding music composition from an engineering perspective. However, we can also view music as being a natural system, and argue that the term ABM is more appropriate. There are specific scenarios though where a certain MAS may not classify as an ABM, although these instances are not discussed here. The title of this thesis uses the term "Multi Agent Models". This term was chosen as it kept it was the most consistent with the titles and keywords used in papers relating to the original model that inspired this project (Kirke, Freeman, Miranda, & Ingram, 2011b; Kirke, Miranda, Rendell, & Ingram, 2015).

This section opens with a discussion on cultural transmission models inspired by music, as these models served as the genesis for this project. As many of the experiments for cultural transmission in music arose from methods in language evolution, we briefly discuss the Talking Heads experiments by Steels (Steels, 2015), as these served as inspiration for many agent based models in music and biology. Finally, models used for studying cultural transmission in birds and toothed cetaceans are discussed. These models have different goals, and as a result can differ greatly. Models in music are often concerned with a creative output, or a specific question related to musicology. Models in biology can be concerned with how genetics and culture interact.

### 2.9.1 Music

There are numerous multi-agent models that examine various issues surrounding music. Much of this research has focused on Evolutionary Computation and A-life methods, and has been collected in work such as (Eduardo R. Miranda & Biles, 2007; Eduardo R Miranda & Todd, 1999; Eduardo Reck Miranda, 2011). While many of the methods we use are inspired by work carried out in Evolutionary Computer Music, the literature surveyed here has been chosen because of its relevance to the thesis project. Many models discussed in the works mentioned focus on the application of **ALife** and evolutionary computing to systems such as sound synthesis, sound design, work within strictly defined systems such as western music theory, or seek to address a specific task in music such as improvisation. Furthermore, evolutionary models do not need to rely on theories found in cultural transmission. For example, systems such as genetic algorithms may be combined with cultural transmission models, but they do not rely on cultural transmission in order to function. For this reason, this section of the literature review focuses on models that place an emphasis on cultural transmission for the evolution of music.

Peter Todd made a significant contribution to the field of Multi-agent modelling for the purpose of music with his paper *Frankensteinian Methods for Evolutionary Music Composition* (Todd & Werner, 1999). In this paper, Todd & Werner gave a survey of the field for evolutionary computer music and examined how many of the efforts that sought to create an artificial intelligence that could create music and algorithmic methods by composers used to create music has many shortcomings. Todd suggested an alternative model that sought to evolve music using a society of artificial female music critics and male music composers.

Todd & Werner developed a number of models, including a model that utilised neural networks, but found that the analysing the results of these models proved too complex. A simplified model was developed, where male composers consisted of a song of 32 notes long, with initial male agents initialised with a random song. Females were equipped with a transition matrix, which is used to score the songs of the male composers it encounters. Todd & Werner developed three different methods through which females could evaluate and score the songs a group of males (termed a *courting choir*) presented to them using these transition matrices. The female critics would then choose a single male from the courting choir based on this score. This would then produce a new agent from each female into the population, who would share a mixture of the musical traits of both the mother and father. Following this, a certain portion of the population would be killed off, in order to prevent overpopulation and generating a surplus of data. The process is repeated for any desired number of generations.

The first method females used to score songs, called the *local transition preference scoring method*, meant that female agents judged songs produced by males based on how closely they mated their transition matrices. That is to say, a male agent, whose song transition probabilities most closely matched those observed by a female agent, would be chosen as a mate. The second method used by females to score songs, called the *global transition preference scoring method*, meant that females scored songs based on the highest transition probability found in their transition matrices. For example, if a female has a transition value from C-G of .75, she will select songs that have a C-G transition exactly three-fourths of the time. The third and final scoring method used, called the *surprise preference scoring method*, in which females score a males song based on how unexpected the transitions it hears are. Female critics achieve this by analysing the sequence they hear, looking up the likelihood of that transition occurring in its transition matrix, and subtracting this value from the most likely transition it expects. For example, if a female has a transition value of C-G at 0.75, and a transition value of C-A at 0.25, every time it hears the transition C-A, it will score that value as 0.75 – 0.25, giving a value of 0.5 for that particular transition. It carries this out for every transition it hears and sums the results.

Firstly, it was found that using the *local transition preference scoring method* resulted in male agents converging on identical songs that generally consisted of a repeating note, or a song of constantly alternating notes. The *global transition preference scoring method* produced ambiguous results and was not investigated further in this paper. Finally, the *surprise transition preference scoring* resulted in songs that changed rapidly over time, producing evolving songs. Despite the significant result of the surprise-based scoring method, the authors commented that the resulting songs produced were indeed diverse and novel, but sacrificed musical structure. A quote from the Todd & Werner paper sums up the musical results nicely:

*"Each individual male song sounds crappy—but each sounds crappy in a unique way"*

The unimpressive musical achievement aside, Todd & Werner showed a result that is significant outside the realm of music. Using a biologically inspired model based on inter-sexual interactions, it is possible to create an evolving sound system. This is a significant result and the Todd & Werner model will be revisited again later in this document.

The work of Todd is significant and addresses how a series of tunes may evolve in a community of agents, but it is a system that relies heavily on western music theory. Eduardo Miranda sought to address this issue by create a model in which a community of agents equipped with appropriate motor and cognitive skills may evolve an intonation system and repertoire of melodies from scratch. (Eduardo Reck Miranda, 2002)

In Miranda's model, agents have a voice synthesizer, a hearing apparatus, memory, and an enacting script. The voice synthesizer is a physical model of the human vocal system, and agents operate it using three vectors in order to produce sound; *lung_pressure(n) interarytenoid(n)* and *cricothyroid(n)*. The hearing apparatuses works using short-term autocorrelation, which allows an agent to extract a pitch contour from a sound, it hears. The algorithm can adjust the resolution of this hearing apparatus, allowing the sensitivity of an agents' auditory perception to be adjusted.

In order to escape the restrictions enforced using a western music system, Miranda sought to escape these restrictions using what he called a Common

Abstract Representation of Melodic Contour (CARMEC). CARMEC does not utilise western music theory in the sense of tones and semi tones. Agents create a Melodic Unit (MU) by following a series of changes in frequency. Agents initialise this at what is called *p-ini,* which tells the agent whether to start the MU at the middle register (SM), the lower register(SL) or the highest register (SH), after this, as the agent progresses through each time step, they will carry out a decision on which direction to move the melody in. Table 1 shows the changes in pitch an agent may carry out as it moves through its MU. Figure 14 shows a visual representation of a melodic unit.

| | |
|---|---|
| VLSU | Very large step up |
| LSU | Large step up |
| MSU | Medium step up |
| SSU | Small step up |
| RSB | Remain on same band |
| SSD | Small step down |
| MSD | Medium step down |
| LSD | Large step down |
| VLSD | Very large step down |

Table 2: The steps in intonation an agent may move through as they progress through an MU. From (Eduardo Reck Miranda, 2002).

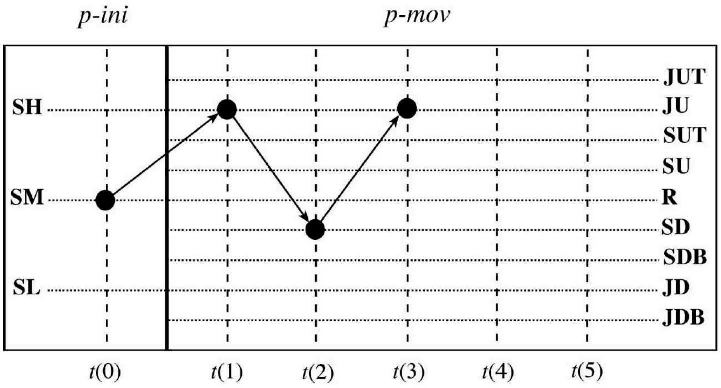


Figure 14: Representation of a melodic unit. From (Eduardo Reck Miranda, 2002).

In the enacting script, agents form pairs and take one of two roles, the *agent-player* or the *agent-Imitator*. The agent player starts interaction by producing a tune, *p1,* which is chosen at random from its repertoire, following this, the agent-imitator searches for a similar tune in its repertoire and produces it. The agent player then analyses this tune *p2* and checks to see if there is no other tune, *pn,* that is more perceptibly similar to *p2* than *p1* is. If another tune, *pn,* that is more perceptibly similar is found, the imitation is deemed unsatisfactory. Otherwise, it is deemed satisfactory. If the tune is satisfactory, the agent-player will send positive feedback to the agent imitator. The agent player will then reinforce *p2* in its memory. Otherwise, it will send negative feedback to the agent-imitator. If the agent imitator receives negative feedback, it will check to see if the song it produced had a high success rate from previous interactions. If it has had a high success rate, then it will not modify the tune it produce, but rather add a new tune to its repertoire that is similar to *p1.*

Miranda's model produces a number of interesting results. As the model progresses through time, the number of melodies in the population increases. Agents also develop similar perceptual repertoires. The interesting aspect of the model however is that there is not a one-to-one mapping between perception and production memory. Even though agents develop near identical perceptual memory, they do not have to develop identical motor-memory in order to enact these tunes.

Miranda and Todd's models were further discussed in an article where they collaborated with linguist Simon Kirby in order to contribute to theories surrounding the development of hierarchical structure in music by addressing issues such as emotion and semantics. They achieved this by adapting a vertical cultural transmission model known as the Iterated Learning Model (ILM)(Eduardo Reck Miranda, Kirby, & Todd, 2003b). The ILM is an important model in evolutionary linguistics, and its importance will be discussed later. For now, we will focus on its importance in relation to music.

In the ILM, there are four components, a signal space, a meaning (semantic) space, agent-teachers (adults), and agent-learners (children). The signal space consists musical sections made up of sequences of short musical passages called *riffs*. The meaning space is made up of emotions and moods, which are combined to create a more complicated hierarchical structure. Agent-teachers use grammars to convey emotions and moods with the music they produce. The agent-learners on the other hand induce these grammars by observing the agent-teachers musical behaviour. Figure 15 shows the structure of the basic iterated learning model. We begin with an agent-teacher and an agent-learner. The agent-teacher performs for the agent-learner for an unspecified amount of time. After this specified amount of time, the agent-teacher is removed, and the agent-learner becomes the new agent-teacher. A new agent-learner is then initialised to learn from the new agent-teacher. Agent-learners are created with absolutely no knowledge of the musical culture they are born into.  This process is repeated for as many times as specified.

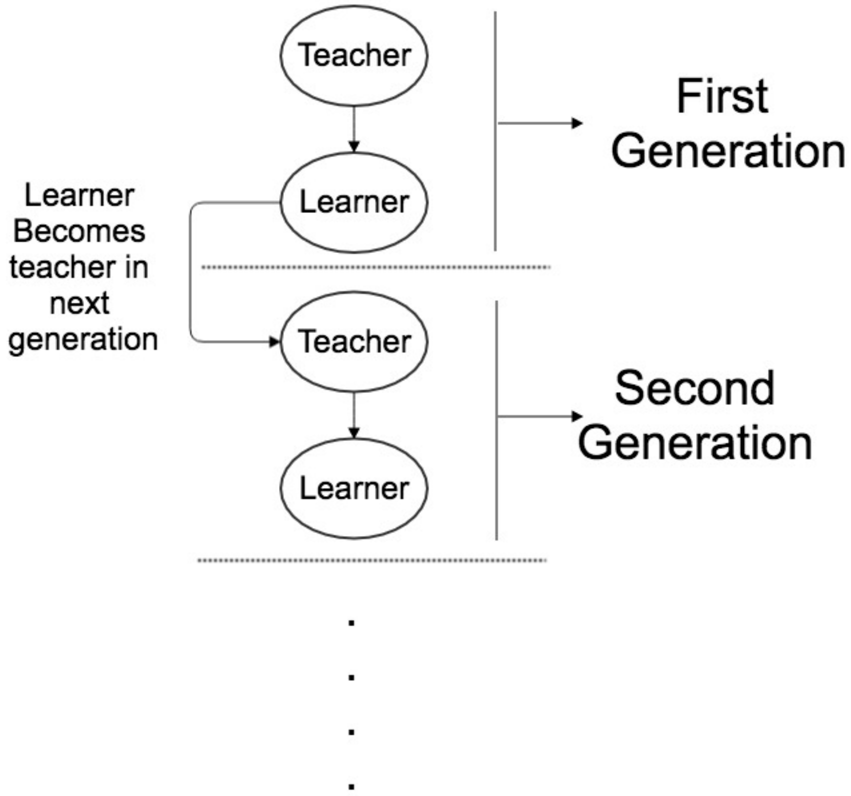


Figure 15: The structure of the iterated learning model. After learning from a teacher for a specified amount of time, the learner agent becomes the teacher agent for the next generation. A new learner agent is initialised with a randomly weighted learning model. This process is repeated for as many generations as specified.

In the ILM for music, agents are equipped with a physical model of a flute, chosen as it is one of the oldest musical instruments in existence. The flute has seven holes and each note corresponds to the notes A5, B5, C6, D6, E6, F#6 and A6. A single component of a signal that an agent produces consists of a note and note duration. Altogether there are six note duration values, very short (vs), short (s), medium (m), long (l), and very long (l). Altogether, there are thirty-five different musical symbols, which consist of combinations of pitches symbols and note duration symbols as follows: A5vs, A5s, A5M, . . ., B5l, B5vl, and so on.

In the ILM, meaning is represented through the combination of riffs and emotions. In total there are 24 emotions that are divided into 8 different groups. It should be noted that these emotions were not chosen arbitrarily but were picked due to the work of a psychologist. The ILM is significant, in that despite the fact it was originally intended for linguistics, its design is easily adapted so that it is possible to investigate specific forms of cultural transmission. In the end of the it can converge on formal grammars with a hierarchal structure. This occurs simply due to the process of cultural transmission.

While all of these models are remarkable for what they demonstrate about the evolution of music, and the methods they employ inspire much of the work in this thesis, there is a caveat that must be considered in that these models rely on ideas that are unique to theories of music. This includes the use strict physical models of the vocal system and flute in Miranda and Kirby's models, the assumption of either an implicit or emergent tuning/musical system, or the reliance on some type of cognitive behaviour associated with human beings. While we will use features of these models in order to move towards models of humpback whale vocalisations, it is necessary to strive to impose a minimum of ideas associated with music theory or human cognition. To do otherwise would be detrimental to the model.

### 2.9.2 Linguistics

Kirby's iterated learning model was originally intended for linguistics and serves as a useful point to segway into discussions on the application of agent based models for linguistics. Agent based models in linguistics were highly influential in developing theories regarding cultural transmission and social learning in language. The Talking Heads experiment (Steels, 2015) is widely credited with starting this trend.

In the Talking Heads experiment, there are two robots, equipped with moveable cameras, audio input and output, and a computer for cognitive processing. There is also a magnetic whiteboard on which different types of coloured shapes (triangles, squares, circles) are pasted. This environment is altered between games, and different types of shapes that have not been seen by agents are added to the games as the experiment progresses. An important point must not be made: this setup was created in multiple locations, in Brussels, Tokyo, Paris, and Antwerp. The agents involved in these experiments technically existed on a server. The agents would be 'teleported to different locations across the globe in order to play the language games designed by Steels. Teleportation here simply meaning that the agents' state at the last game they played at was uploaded on the physical machinery at each location. For this reason, there needs to be an important distinguishment made here; agents are the entities that exist on the server and can control the cameras, computers, and audio devices around the world. These devices that consist of the cameras and audio input/output will be referred as the 'bodies', as in Steels description. Two agents can only interact if they are in the same location.

The agents can be broken down into having different layers that carry out different cognitive capabilities; a perceptual layer that collects data about the images the agent sees, a conceptual layer that categorises the processed image, a lexical layer that is a changing repertoire between images, meanings and words, a syntactic layer that uses a grammar schemata for combining words into larger structures, and a pragmatic layer, that is responsible for the scripts that carry out the language playing games, such as the guessing game. In the guessing game agents choose an image on the board and communicate it to the listening agent (hearer). The hearer then needs to decipher what image was chosen. If the hearer fails, the speakers and hearer adjust their conceptual and lexical layers. This process is repeated and the agents take turns acting as speaker and hearer, and the images on the white board change. This recreates situations in language, where, for example, a hearer may not for example speak the same language as the speaker. If a successful communication system emerges, it shows that the agents have developed some kind of comprehension between their cognitive layers, despite not having direct access to these layers. Steels experiments were important, as they showed that cultural transmission has a direct influence on the learning of vocalisations, and that interaction between conspecific was necessary for these systems to change and evolve over time.

Other agent based models in linguistics have examined theories associated with language. These include Kirby's original ILM to investigate how the how structure can emerge in language (S. Kirby, 2001). Others have examined the evolution of vowel systems in a population of agents (de Boer, 2002). Because of their specificity to language they will not be reviewed here.

### 2.9.3 Biology

Individual based models of cultural transmission in song birds were first demonstrated by Goodfellow and Slater in 1986 (Goodfellow & Slater, 1986; Slater, 1986). These models used agents in a grid environment, with each agent occupying a single cell on the grid. Agents are able to learn songs from other agents in the cells below, above, and beside them. Over time, agents on this grid die, and new agents are introduced to replace them. Cultural evolution would occur through the introduction of a new agent with a new song type, or through errors in copying the song of another agent. The accumulation of these errors led to the development of local dialects, without the presence of any specific convergent learning bias. Copying errors are a process under which cultural evolution may occur. In this process through a mistake made when learning (Slater, 1986). Rendell highlighted that errors in learning can lead to new innovations in a social learning strategies tournament (L Rendell et al., 2010). Research has also been carried out into how copying errors can lead to variation in objects left behind by various cultures (Eerkens & Lipo, 2005).

Grid based models have since become more sophisticated and can be used to answer specific questions following adaption. Ellers and Slaberkoorn investigated how the dispersal of offspring and the learning method used by agents affects the genetic flow of a population (Ellers & Slabbekoorn, 2003). It was found that song was important in creating an incompatibility of genetic divergence and song divergence between populations.

A more recent paper by Rowell and Servedio (Rowell & M.R., 2012) examined how population movement and contact between populations with specific songs affect cultural evolution. Specifically, they wished to know if population specific songs can be maintained in a contact zone or be replaced by a shared song. They also sought to identify the spatial patterns in the distribution of songs that may result from contact. To do this they investigated a number of influences, such as song based mating preferences and movement probabilities, vertical song transmission (paternal) versus horizontal song transmission (non-paternal), and genetic and cultural transmission. They found that population specific songs may be maintained given a number of conditions. The largest of these influences appeared to be when females had a preference for songs from their own population.

A more novel model recently developed by Stowell et Al. (Stowell, Gill, & Clayton, 2016) used a method of recreating temporal patterns of animal call timing using models originally developed for networks of firing neurons. This model was trained using recordings of groups of domesticated zebra finch, and found that the groups of animals have a stable communication network that persists from one day to the next. The model is particularly useful as it includes information about temporal interaction, and how these may affect the time and sequencing of calls. It also can be used as a generative model in order to synthesise call patterns and has information about the relationship and influences between individuals in the network.

In regards to cetacean research, Agent Based Models of cultural transmission have been developed for investigating cultural transmission in matrilineal whales such as the killer whale (Hal Whitehead, 1998) discussed in the section titled Social

Network Theory in this chapter,  and in sperm whales (Maurício Cantor et al., 2015). Since the work of Whitehead has been discussed previously, this section will focus on the work of Cantor. Cantor first carried out an empirical analysis of the vocal dialects and group composition of Sperm whale in the Galapagos, identifying their different clans and vocal codas.

## 2.10 Discussion

From this chapter, we have seen that there are many different things to consider when designing a multi-agent model for investigating humpback whale song. While there have been models designed to investigate cultural transmission in toothed whales, there are many differences at play in terms of biology. This includes differences in anatomy (lack of U-shaped folds in toothed whales), social structure (matrilineal social structures found in toothed whales are absent in humpbacks, as are the pod structures found in bottlenose dolphins), and vocal function; dialects in killer whales, and codas in sperm whales codas are used for clan and unit identification, and signature whistles in bottlenose dolphins are used to maintain group cohesion. This is very different to the vocalisations of humpbacks that are involved in sexual interactions.

From a modelling perspective separate from cultural transmission, we have seen that there are many different types of methods that can be used for learning, synthesising and analysing symbolic sequences of animal vocalisations. These methods are powerful, but when viewed through the lens of cultural transmission, we begin to see that methods such as artificial neural networks are inappropriate due to their black box nature. For this reason, methods used for sequence production and learning in this thesis focus on algorithms that follow a specific process that can be easily broken down, and their independent effects separated from the effects of cultural transmission.

Moving on to the discussion on formal grammars in regards to humpback whale song, James Hurford, a computational linguist at the University of Edinburgh, wrote the following in regards to humpback whale song and birdsong:

> *On a narrow approach ignoring numerical information, the overt patterns of natural bird – and whale song can mostly be adequately described by First order Markov transition tables. Some unusual exceptions require State Chain descriptions. But Phrase Structure grammar is never required, in terms of weak generative capacity, to capture the overt patterns.* (Hurford, 2011, pp. 97)

Hurford strongly criticised the work of Ryuji Suzuki (Suzuki et al., 2006), who demonstrated that Markov chains required an unrealistic amount of information in order to accurately capture the structure at the unit level. Hurford dedicates nine pages of his book to arguing against Suzuki's paper (Hurford, 2011, pp. 63–72), and that only state chain descriptions are required. When we examine the application of Markov chains in a cultural transmission context in Chapter 3, we will see the same problems emerge in regard to the amount of numerical information required. A large amount of information is required in order to accurately capture the hierarchy displayed, and in an agent-based model, agents would be required to generate very long songs in order to accurately recreate their internal transition matrices. We will also see that at a unit level, this transmission results in very short songs that are not representative of humpback song. However, at a theme level, Markov chains are capable of capturing the sequential structure. They are also useful as place holders, serving as a simple method of sequence production and learning, and can be thought of as being modular when used in an agent based model, being replaced by more complex sequence production and learning methods later on. The shortcomings of Markov models should be taken into account however, and, as Suzuki suggests, methods found in formal grammars may prove useful in addressing these shortcomings. In regards to Hurford's comment on the numerical problem, it is worth considering that formal grammars have been used in computer science as a method of data compression (Cleary & Witten, 1984; Nevill-Manning & Nevill-Manning, 1996), and could be seen as a useful method to reduce the amount of incoming information for modelling purposes. We have also seen that formal grammars have been used to generate complex pieces of music and generate complex fractal and plant like structures. These sequences are free from any type of mappings related specifically to language. These musical sequences and pattern generation abilities highlight ways in which we can harness formal grammars in order to synthesise animal vocal sequences. This issue is addressed further in Chapter 5 where we examine the applications of a formal grammar method in regard to humpback whale song analysis, but also to cultural transmission of the song.

A stronger argument has been made against the syntax of birdsong. A study in 2011 suggested that it was possible for song birds to learn a context free grammar (Abe & Watanabe, 2011). This study has proven to be controversial, and was challenged in (Beckers et al., 2012). It is worth noting that the issue of formal grammar and birdsong has received a lot more attention than humpback whale song. This is mainly to do with the fact that songbirds may be kept in captivity for study, and while experiments have been carried out in song playback to humpbacks (Tyack, 1983), experiments to teach songs such as those carried out by Abe and Watanabe would not be possible with any baleen whale. Humpback songs are also significantly different to birds, and there are many aspects of song learning that can be investigated in a laboratory setting, while humpback song research can only be carried out in the field.

Despite the certain structural differences between bird and whale song, we have also seen that there are many other similarities between birdsong and humpback whale song, such as the sexual function that it serves. This would appear to make agent based modelling of birdsong the ideal point to begin our investigation. However, the general approach to modelling birdsong has been inspired by mainly grid-based models. Further complicating this matter are the differences in migration and movement patterns. Birds tend to migrate in order to find food in winter. They're generally not migrating for the very specific purpose of breeding like humpback whales are. Birds also create a nest for when they are giving birth, and this affects their spatial positioning. Furthermore, birds do not produce songs for a sexual function during their migration period. This complex behaviour is something that needs to be considered when creating a model for the cultural transmission of humpback whale song. For this reason, the first model developed and discussed for this thesis implements movement algorithms to recreate some of the phenomena that may influence cultural transmission in humpback whale song. The modelling approach to investigating social networks of Stowell and Cantor are very inspiring. However, Stowell's model requires data that is unavailable to us, and Cantors is specifically related to the emergence of multi-level animal societies, rather than how they influence song transmission. However, in Chapter 4, a model that makes use of social networks for the cultural transmission of humpback whale song is put forward.

The cultural transmission experiments in linguistics also create a number of issues, but also point to interesting areas to investigate. The main problem with the models developed for linguistics is that they generally rely on theories found in language. For example, the mapping between meanings in the Talking Heads of Steels, and the emergence of structure in the Iterated Leaning model of Kirby. We cannot assume that humpback whale vocalisations have strict meaning-symbol mappings like the communication systems that arise from these experiments. When viewed through the lens of evolutionary linguistics, there is very strong argument for vocal learning being a precursor to language, but we cannot attach language like properties onto the vocalisations of animals. These experiments are useful in designing structures for experiments in cultural transmission, and these systems are useful for determining how cultural transmission will affect a communication system equipped with a variety of sequence production and learning rules. They also provide good insight in how to create experiments cognitively. The notion that agents cannot read minds like in the guessing game is one of the key aspects of cultural transmission, and an approach that should be followed whenever designing cultural transmission experiments. Music encounters similar problems as the language models, but it does present an interesting perspective in that we do not need to rely on the mapping of semantics like in the language systems.

These are all things that need to be considered when developing the models. Also, this research was carried out in a true interdisciplinary spirit. There has been significant effort in order to apply the methods from in computer music research to inform the design of these models. Likewise, as we will see in Chapter 6 there is also an effort made to link these models back to music.

# Chapter 3 – Migratory Models and Cultural Transmission

## 3.1 Introduction

This chapter focuses on the design of the first model used to investigate cultural transmission in humpback whale song. The model is built in Python using the SciPy packages (Jones, Eric and Oliphant, T. and Peterson, 2001).

This model is part of a collaboration between colleagues in St Andrews, University of Queensland, and Plymouth University. Specifically, we sought to create an interdisciplinary model, one that would adapt the techniques used for modelling cultural transmission in music, so that a broad range of questions relating to cultural transmission in humpback whale song may be addressed. Through the creation of this model, we are able to recreate specific scenarios that may have a significant impact on the cultural transmission of humpback whale song.

The model presented in this chapter is referred to as the Migratory Model. First, we introduce the most basic migratory model, which is referred to as Migratory Model 1 (MM1). MM1 is the most basic multi-agent model in this chapter. It assumes that sound transmission loss, coupled with a first order Markov model, may be responsible for the phenomena of song evolution and revolution. The model also simulates the migration patterns of humpback whales, and their interactions on the breeding and feeding grounds.

As we will see in this chapter, this is not the case, but this serves as the foundation for our models, and allows us to add extensions. The migratory model is then extended to include bias towards the learning of *novel* songs (MM 1.2), the introduction of production errors without a novel song learning bias (MM1.3), and the coupling of production errors and novelty learning bias (MM. 4).

After the technical description of the migratory model, we investigate how the movement of the agents affects their songs and transition matrices. We also discuss how the model highlights the importance of feeding grounds as a key location for cultural transmission between populations. These opportunities for cultural transmission on the feeding grounds may facilitate song revolution.

## 3.2 Migratory Model 1 Architecture

MM1 has a cyclic structure. After inputting the model parameters, the model begins looping through a series of rules that determine how the agents behave. These instructions can be broken down into three broad categories; Movement rules, song production rules, and song learning rules. Here, we will examine these three categories in detail. The parameters of the model and how they affect its behaviour are discussed later in the chapter.

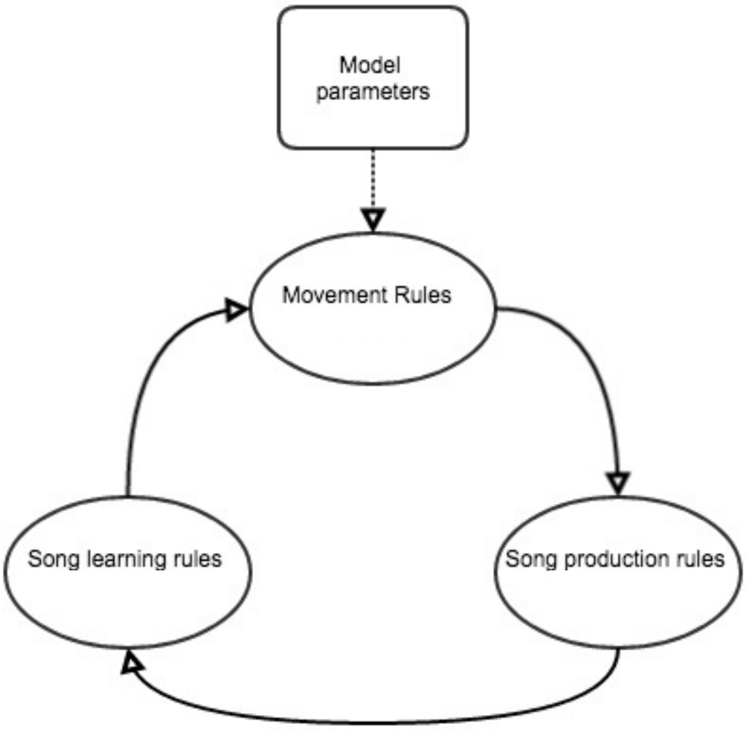


Figure 16: The basic structure of the model, and the order in which the rules are carried out.

### 3.2.1 Movement Rules

In all the Migratory Models presented in this chapter, agents exist on a Cartesian plane. When the model is initialised, the user specifies the size of the area that the agents are spawned in. This area is referred to as the spawning area. The co-ordinates of the spawning area also correspond to the co-ordinates of the feeding grounds. It is important to note that agents are initialised in a square area, but the breeding zone and feeding zone are circular. This was done so that the agents may be initialised in specific patterns that do not necessarily correspond with the size of the feeding grounds. When the model is initialised, agents are assigned random co-ordinates within this spawning area. After agents are spawned, they follow a series of movement rules that controls their movement behaviour.

The movement rules for agents in Model 1 are based on animal movement research carried out by Ian Couzin (Couzin, Krause, James, Ruxton, & Franks, 2002). Couzin's research itself is inspired by the animal flocking algorithms originally developed by Craig Reynolds (Reynolds, 1995). In Reynolds' original algorithm, an agent behaves differently depending on the proximity of other agents. Agents follow three basic rules; collision avoidance, velocity matching, and flock centering. These rules are explained by Couzin through the use of different types of zones surrounding an agent. These are the Zone Of Repulsion (ZOR), the Zone Of Orientation (ZOO), and the Zone Of Attraction (ZOA) and are shown in Figure 17. In Couzin's model, whenever an agent has another agent within its ZOR, it will calculate a new trajectory in order to avoid the other agents within its ZOR. If there are no agents within an agents' ZOR, then the agent will attempt to orientate itself with agents within its ZOO. Agents will also seek to move towards other agents within its ZOA.

**Figure removed due to Copyright Restrictions**

Figure 17: Taken from (Couzin et al., 2002), this figure shows the Zone Of Repulsion (ZOR), , Zone Of Orientation (ZOO), and the Zone Of Attraction (ZOA) surrounding an agent in a three dimension Cartesian environment.

MM1 uses a series of rules inspired by Couzin's work in order to simulate the migration patterns of humpback whales to the breeding grounds, and back to their feeding grounds. Couzin's work is most relevant to flocking animals, and focuses on group sorting, and the role of individuals within that group. Couzin's work also highlights a form a collective memory in a group of flocking algorithms. While humpback whales undertake their migrations at roughly the same time, they do not migrate in flocks. Furthermore, it is unlikely that there is some type of collective memory in the group in regard to their movement. For this reason, the rules were adapted to deal with the fact that, with the exception of mother calf pairs and their escorts, humpback whales do not migrate in groups. Due to this fact there is no ZOO associated with the agents in MM1. The lack of ZOO also helps in reducing the parameter space that needs to be analysed later on. The ZOA rule has been adapted so that the humpback agents will seek out the nearest agent in its ZOA that is singing, but only if they are on the breeding grounds. This behaviour is based on the observation that male humpbacks will approach and interact with other singing males (Smith, Joshua N, Goldizen, Anne W., Dunlop, Rebecca A., Noad, 2008). The ZOR rule remains unchanged, but the method for calculating the new trajectory an agent takes is based on the algorithms described by Shiffman (Shiffman, 2012).

The movement behaviours of agents is also determined by input parameters that specify how long the 'feeding' and 'breeding' season should last. These control a Boolean migration state within each agent. After all agents carry out their movement rules, their singing rules, and their song learning rules, a single iteration of the model is said to be complete. After a specified number of iterations have been carried out, the agents will have their migration state activated. The agents will begin to move towards the breeding grounds, with the size and location specified by the model parameters. Once they are within the breeding grounds, the ZOA rule will be carried out, and they will seek the nearest agent within their ZOA, and also within the breeding grounds. This rule was incorporated as males have been seen to approach other singing males on the breeding ground (Smith, Joshua N, Goldizen, Anne W., Dunlop, Rebecca A., Noad, 2008). If an agent wanders outside of the breeding ground, they will automatically move back towards the centre of it. The duration of the breeding season is a parameter value that can be controlled, and once a breeding season is over agents will begin to migrate back to the feeding grounds. The parameters of the model associated with movement include spawning area size, feeding ground size, breeding ground size, ZOA size, ZOR size, a value to indicate what iteration to begin migrating to the breeding ground, a value to indicate what iteration to begin migrating to the feeding ground, and the number of migrations the agents carry out. During the 'feeding' season, agents carry out a random walk within the feeding grounds.

### 3.2.2 Song Production and Learning Rules

In all the models presented in this chapter, all agents are equipped with a first order Markov transition matrix. The easiest way to understand Markov models is to look at an example of training one and using it to synthesise a sequence. For example, consider a process that produces the following sequence of symbols:

$$S = [A, A, B, B, C, C, D, D, A, A, B, B, C, C, D, D, \dots]$$

We can generate a probability distribution by counting the number of times an individual symbol occurs and dividing it by the total number of units in a sequence. Doing this to the sequence $S$ above would yield the following probability distribution:

$$P = [\frac{4}{16}, \frac{4}{16}, \frac{4}{16}, \frac{4}{16}]$$

Where each element of the array corresponds to the individual symbols as they appear in the sequence, IE. P [1] represents the likelihood of symbol A being output by the process, P [2] represents how likely the process will output B and so on. We may sample from this probability distribution and produce our own symbols using the following formula:

$$x = \sum c \leq U \qquad (1)$$

Where $x$ is our output symbol represented as an integer (A = 0, B = 1, C = 2, D = 3), $c$ is the cumulative summation of the probability distribution $P$, and $U$ is a uniformly distributed random number between 0 and 1. This type of model is known as a zero-order Markov chain. If we were to repeat this process a number of times, we could generate our own sequence. The sequence below, *s2,* is an example of this:

$$s2 = [\, D, A\, , D, C, A, B, C, D\, ]$$

It is easy to see that this model has not captured the structure in the original sequence. The frequency of each symbol is nearly correct, and if we were to continue sampling from our probability distribution we would find that the frequency of each symbol would approach our original distribution. To improve this model though, we can extend it to what is called a first order Markov chain. In a first order Markov chain, rather than simply counting how many times a unit occurs and dividing it by the total number of units, we count how many times a unit follows another unit. For example, how frequently B comes after A, or how many times C comes after B. We could create the probability distribution for the unit A and its subsequent units for the sequence *S.* This would yield the following probability distribution as seen in Table 3.

| | A | B | C | D |
|---|---|---|---|---|
| A | 2/4 | 2/4 | 0 | 0 |

Table 3: This shows a simple probability distribution for which units are likely to follow A. Since A occurs a total of 4 times in the sequence, we simply count how many times each unit follows A. Two times out of four, A was followed by another A, the other two times A occurred it was followed by B.

We may calculate the transition probability distribution for every unit in our sequence *S.* These probability distributions can then be stored in what is called a transition matrix. This is a square matrix of size *N,* where *N* is the number of symbols in our sequence. This is easily seen in table 2. Figure 18 also shows a method of visually representing the information displayed in Table 4.

| | *A* | *B* | *C* | *D* |
|---|---|---|---|---|
| *A* | *2/4* | *2/4* | *0* | *0* |
| *B* | *0* | *2/4* | *2/4* | *0* |
| *C* | *0* | *0* | *2/4* | *2/4* |
| *D* | *2/4* | *0* | *0* | *2/4* |

Table 4: Transition matrix for sequence *S.*

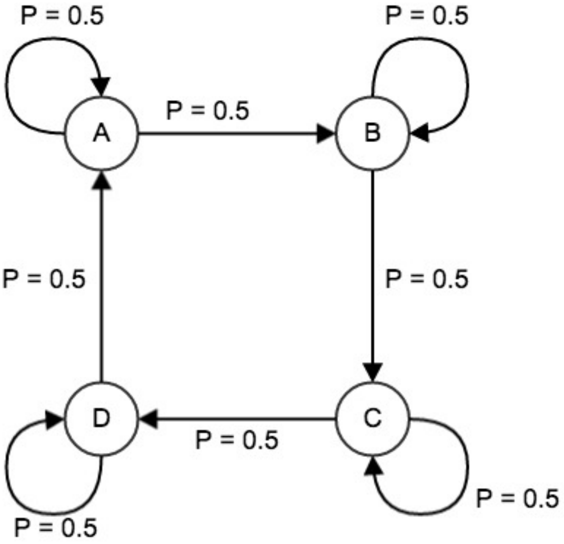


Figure 18: State transition diagram for a first order Markov chain of sequence *S.*

Using Equation 1, we can sample from each row of this transition matrix to generate a new song by sampling from the first row to produce a new unit, storing that unit in an array, and using that output unit to choose which row to sample from next. An example sequence created using this method is given below.

$$s3 = [A, B, B, C, D, A, B, B, B, C, D, \ldots]$$

It is easy to see that *s3* resembles our original sequence *S* clearly is a closer match *s2.* The order of the units is correct, but the number of units is not. We can solve this problem by extending the model to a second order Markov chain. Rather than looking at what follows a single unit in our sequence, we look at combinations of units. For example, which unit follows the combination AA, AB, BB, etc. We can generate yet another transition matrix for *S* as displayed in Table 5.

| | A | B | C | D |
|---|---|---|---|---|
| AA | 0 | 1 | 0 | 0 |
| AB | 0 | 1 | 0 | 0 |
| BB | 0 | 0 | 1 | 0 |
| BC | 0 | 0 | 1 | 0 |
| CC | 0 | 0 | 0 | 1 |
| CD | 0 | 0 | 0 | 1 |
| DD | 1 | 0 | 0 | 0 |
| DA | 1 | 0 | 0 | 0 |

Table 5: Transition matrix for second order Markov chain of sequence *S.*

Using Equation 1, we can sample from the transition matrix in table 3, providing us with the following sequence:

$$s4 = [A, A, B, B, C, C, D, D, A, A, B, B, C, C, D, D, \ldots]$$

Which is equal to our original training sequence *S.* Markov chains can be extended to even higher orders, examining three units at a time and their subsequent units. An important point before moving away from the technical discussion of Markov models is clear up a problem in terminology in regards to *n-gram* models. The concept of an *n-gram* model was introduced by Claude Shannon, and a *n-gram* model is equivalent to an *(n-1)*th order Markov model over the same alphabet (Rohrmeier et al., 2015).

This model is not interested in modelling the hierarchy of the song, so although Markov chains are not adequate for modelling the song at a unit level, they are capable of representing transitions from theme to theme. For this reason, the sequences being produced should be viewed as being at the theme level. The method used to produce and learn songs are almost identical to the method described in the previous section using transition matrices, but with added dimensions that affect how agents learn songs. With every iteration, the agents will 'listen' to the songs being produced by every agent. This means that an agent receives a sequence of symbols produced by another agent, and estimates a transition matrix for this sequence. This estimated transition matrix is referred to as the Song Representation (SR). The number of symbols that agents are capable of producing is controlled by the size of the transition matrix they are assigned at the beginning of the model. All agents have the same size transition matrices. Agents are also equipped with a Boolean state that determines whether or not they will produce a song. This is called the singing state variable. There are a number of factors that affect whether or not an agent is singing. The main factor that controls this state is called the *singing probability vector.* This is a variable that controls the probability of an agent singing depending on how many iterations of the model have passed. Throughout the model run, the probability of an agent singing may be adjusted. The singing probability vector consists of a two-column array. The first column contains a probability value between zero and one. This probability value is used by each agent in a weighted coin toss to determine whether or not they will sing. The second column contains the iteration number which informs the model when to update the probability value on the left hand side to the next row down in the vector. The reasoning for incorporating a probability value relating the likelihood of singing to time is due to the fact that male humpback whales are primarily observed to sing during their migration to the breeding grounds and on the breeding grounds. However, there have been observations of song being performed on feeding grounds (Vu et al., 2012). Generally, model runs are adjusted so that agents have a very low probability of singing while on the feeding grounds, but when the migration trigger is activated, the probability of singing also increases, and increases even further once the agents have reached the breeding grounds, thus recreating the singing frequency patterns observed in the wild. An agent learning a song will not learn the song of a non-singing agent. Agents will stop singing if another agent gets too close to them, as observed in the wild (Smith, Joshua N, Goldizen, Anne W., Dunlop, Rebecca A., Noad, 2008).

When an agent begins to learn a song, it estimates a transition matrix for the first agent in the population. It then calculates the distance between itself and the same agent it calculated the transition matrix for, using the Cartesian distance formula shown in Equation 2.

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{2}$$

Where $d$ is the distance between the two agents, $\boldsymbol{x_1}$ and $y_1$ are the Cartesian co-ordinates for the listening agent, and $x_2$ and $y_2$ are the Cartesian co-ordinates for the singing agent it is currently estimating the transition matrix for. The value of $d$ is then used to calculate what is called the *intensity factor*, or *I,* as shown in Equation 3.

$$I = \frac{1}{d^2} \tag{3}$$

The intensity factor determines how large of an impact the singing agent's song will have on the listening agent's transition matrix. This intensity factor is based on the loss of sound transmission in water (*Fundamentals of Underwater Sound*, 2008). The listening agent's updated transition matrix is calculated using Equation 4 below.

$$T_A = A * (1 - I) + (B * I) \tag{4}$$

Figure 19 shows the full cycle of song production by a singer, all the way through the song being learned by a listener agent. It is important to note that even though a listener agent will calculate the transition matrices for agents that are far from the listener, the use of the intensity factor means that these songs have no influence when they ‘learn’ them. Figure 20 shows the entire cycle the model goes through and summarises the processes described in this section.

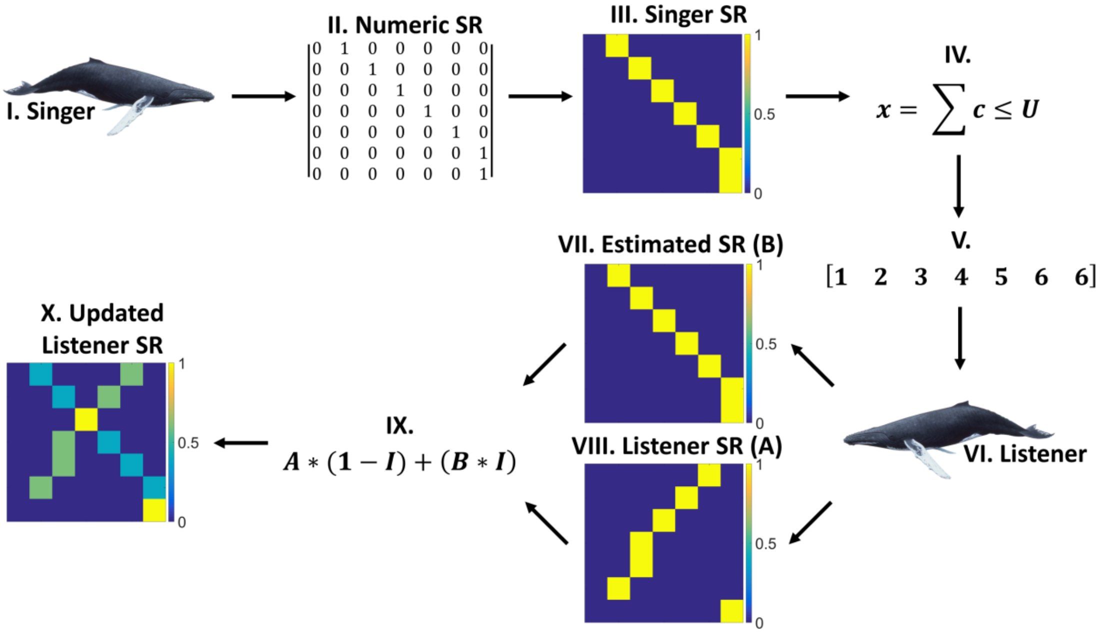


Figure 19: This shows the complete cycle of a singer producing a song using its transition matrix, and the learner agent updating its transition matrix based on the distance between the two agents.

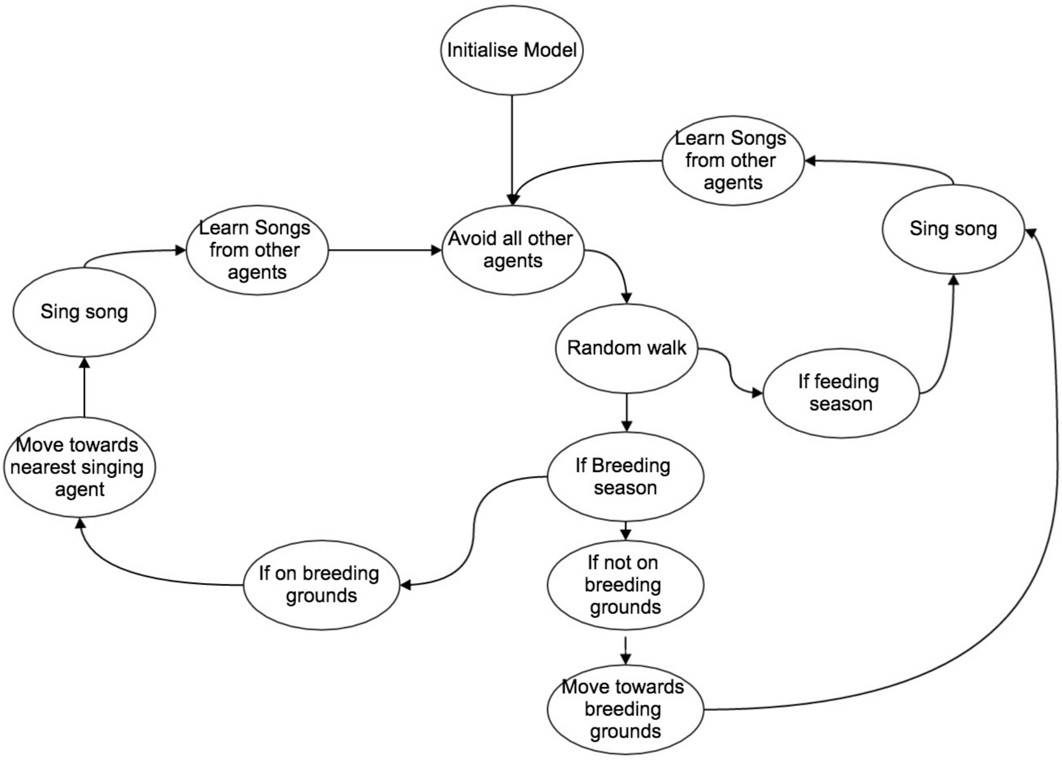


Figure 20: The flow of the model, showing all steps the agents carry out in a single cycle, and how variables such as breeding season and agent location affect model flow. Note, that the “sing song” part of the model cycles are also controlled by a specified probability vector.

A summary of the complete algorithm is given below.

1. Initialise parameters.
2. Spawn in spawning arena.
3. All agents produce one song.
4. Avoid all other agents.
5. Random walk.
6. If feeding season:
    a. If not on feeding ground, move towards it.
    b. Every agent decides whether or not to sing based on probability vector.
    c. Sing if probability vector returns a value of 1.
    d. Learn songs from other singing agents.
    e. Increase iteration number.
    f. Return to step 4.
7. If breeding season:
    a. If not on breeding grounds, move towards it.
    b. If on breeding grounds, move towards nearest agent.
    c. Ever agent decides whether or not to sing based on probability vector.
    d. Sing if probability vector returns a value of 1.
    e. Learn songs from other singing agents.
    f. Increase iteration number.
    g. Return to Step 4.

### 3.2.3 Model Parameters and Interaction

The parameter space for the model is large. This was a deliberate design choice, as the model was intended for investigating scenarios where a specific parameter setup may have a significant effect on the transmission of the song, and allows a researcher to recreate scenarios that facilitate cultural transmission (for example, a bottleneck caused by a small feeding ground). The parameter space is split into two different categories, universal parameters that represent parameters that are universal to all agents (e.g: Breeding ground size), and agent specific parameters, which are assigned to each agent individually using a spreadsheet designed in Microsoft Excel. The full list of parameters is given in Table 6, and whether they are specified in Python or on the Spreadsheet.

| **Parameter Name** | **Universal** | **Spreadsheet** |
|---|---|---|
| Zone of Repulsion | **X** | |
| Spawning Zone Height | **X** | |
| Spawning Zone Width | **X** | |
| Number of Iterations | **X** | |
| Number of Migrations | **X** | |
| Feeding Ground Location | **X** | |
| Migration Trigger | **X** | |
| Return Trigger | **X** | |
| Zone of Attraction | **X** | |
| Singing Probability Vector | **X** | |
| Input File | | **X** |
| Transition Matrix | | **X** |
| Breeding Ground | | **X** |

Table 6: Model 1 Parameters, and whether they specified by model input at Python (Universal) or using a spreadheet.

The input Excel file has a specific structure that allows the unique variables to be assigned to each agent. It also determines the total number of agents in the model. In the Excel file, each row corresponds to an individual agent's breeding ground location, and the transition matrix they are initialised with. All subsequent sheets are specified transition matrices that may be assigned to any agent.

It is important to note that the model is deliberately designed to correspond with distance values on the Cartesian plane, and the sizes of the feedings grounds, breeding grounds, and zones of influence are inspired by real-world ratios rather than by distance metrics such as kilometres.

### 3.3 Migratory Model 1.2 – Adding Novelty

In model 1.2, we sought to address what certain researchers believe to be a factor that may drive changes in the song of humpback whales - novelty (Noad, Michael J., Cato, Douglas H., Bryden, M. M., Micheline, Jenner, Jenner, 2000). As discussed in the literature review, the prevalence of singing on the breeding grounds, and the fact that only male humpbacks have been observed singing, suggests that the song serves some kind of sexual function. This suggests that humpback whale song may serve as a cognitive demonstration to females, and that factors such as novelty may be what drives song change. Novelty can be a confusing term, so we define it here as the transition to an unexpected unit based on an agent's internal SR. In other words, novelty is the inclusion of unexpected transitions from one unit to another. For example, if an agent expects to hear a transition from 1 to 4 based on its internal SR, but instead hears a transition from 1 to 5, and this transition is not accounted for in its SR, then this transition is deemed to be 'novel'.

In order to achieve this definition of novelty, it is calculated using the agent's built-in transition matrix. The novelty algorithm takes inspiration from the work of Todd and Werner (Todd & Werner, 1999), where a transition matrix is used to calculate a novelty score that is used by female agents to decide whether or not they will mate with a male agent. Novelty is calculated as the difference between the transitions an agent expects to hear based on its own SR matrix, and the transitions it actually hears. These differences are then summed, and divided by the total number of transitions observed, in order to create α, the novelty value, which is then used to update the listener's SR matrix as follows:

$$SR_l' = SR_l * \big(1 - (I * \alpha)\big) + SR_s * (I * \alpha) \tag{5}$$

Figure 21 shows the differences between the two learning strategies, and explains them in more depth.

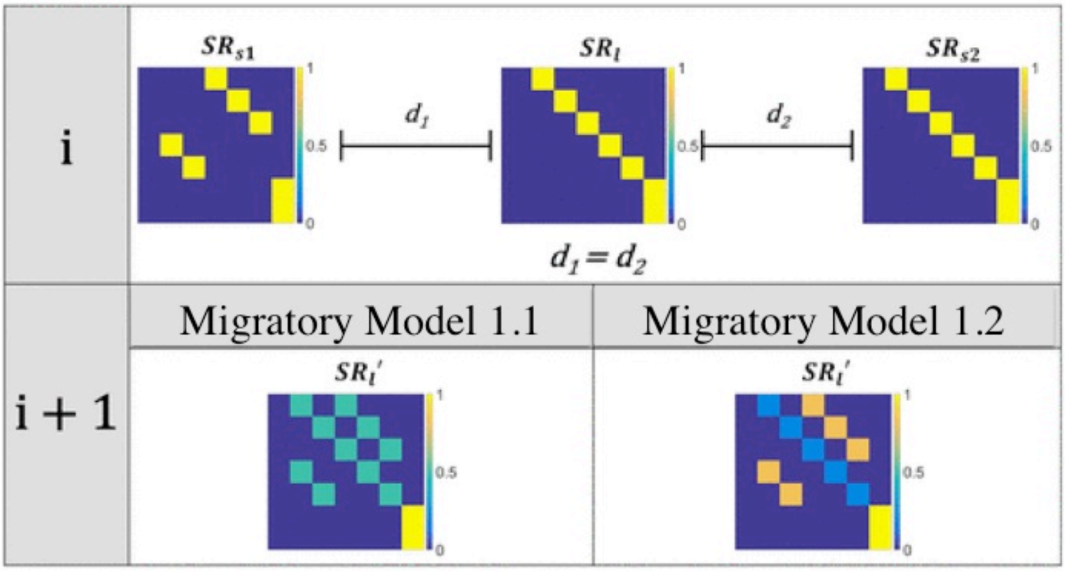


Figure 21: Comparison of the learning processes of Models 1 and 1.2 using a common initial spatial scenario. At iteration *i* the listener hears two equidistant singers. Depending which model is implemented, the listener's song representation ($SR_l'$) at iteration *i*+1 will vary. Using Model 1, the transition probabilities of both singer 1 ($SR_{s1}$) and 2 ($SR_{s2}$) will be equally represented in the resulting listener's SR. Using Model 2, the listener will favour in its resulting SR the more "unexpected" transitions of singer 1.

## 3.4 Migratory Model 1.3 – Adding Production Errors

As discussed in the previous chapter, production errors are a term used to describe when an animal introduces a new element to its vocalisation. These simple mistakes could be one of the things that drive cultural evolution, and may be an important aspect of the song to model. In order to investigate this possible cultural driving force, Migratory Model 1 was extended to include production errors in the absence of the novelty learning bias introduced in Migratory Model 1.2.

In order to model production errors, a weighted edit approach inspired by the Levenshtein distance algorithm is introduced. Using this algorithm, we may control the probability of a theme being inserted, deleted, or substituted for another. Personal observations from both Noad and Garland noted that there appeared to be a higher rate of insertions of themes to the song. However, the probability vector is a variable that can be controlled. The likelihood of a production error being carried out is controlled by a weighted coin toss. The weighting of this coin toss is a model parameter. The production algorithm works in the following manner:

1. Agent produces a song using its SR transition matrix.
2. A production error occurs with the probability $P_e$, a parameter in the model simulation. If that probability is achieved against a random number draw, then select a position in the sequence for editing at random.
3. Insertion, substitution, or deletion is selected based on the probability vector, and performed at the sequence position selected in step 2. In the case of insertions and substitutions, the new theme is chosen at random.

## 3.5 Migratory Model 1.4 – Coupling Production Errors and Novelty

The final model is the most complex, where production errors are coupled with the novelty algorithm introduced in Model 2 and the distance algorithm from model 1. This allows for the investigation of all three mechanisms in a single model, and examine the cultural evolution of the system.

## 3.6 Model Analysis

This section presents a series of experiments designed to investigate how the various parameters of the model affect its behaviour. It begins by discussing the analysis methods used to determine differences between the output of each agent's transition matrices and songs. Following this, each model is discussed independently, examining how distance, novelty, production errors, and the coupling of novelty and production errors affect the model's output.

### 3.6.1 Analysis Methods

Here we discuss the various methods that are used in analysing the model. All analysis was carried out in the Matlab environment.

#### *3.6.1.1 Matrix Dissimilarity Scoring*

In order to analyse the model, a method of measuring an overall dissimilarity between every agent's transition matrices was developed (private communication with partners as St Andrews). Song convergence is measured by calculating the mean transition matrix dissimilarity between pairs of agents within and between breeding populations. The transition matrix dissimilarity between two agents (agent $a$ and agent $b$) is calculated using the following formula:

$$\sum_{i=1}^{n}\sum_{j=1}^{n}|SR_a - SR_b|_{ij} \tag{6}$$

Where $n$ is the size of the transition matrix, $SR_a$ and $SR_b$ are the transition matrices for agents $a$ and $b$ respectively, and $i$ and $j$ are the indices for these transition matrices. These values are averages across pairs of agents in the same breeding populations, and pairs of agents in different populations. This gives us values for transition matrix dissimilarities in every iteration of the model. These values may be plotted over time to give us a chart that shows how similar/dissimilar the two populations' transition matrices are. We can also use this to determine how dissimilar the SRs of the agents in a single population are.

Matrix dissimilarity was chosen as it provides a representation of the differences in internal song representation between agents. Levenshtein distance is another method that could have been used to compare songs between agents. However, analysing the songs does not provide any insight to the probabilities stored in the transition matrices. For example, it is possible for two agents to produce the same song but have fairly different transition matrices.

### 3.6.2 Parameter Exploration

In order to investigate Model 1, a series of experiments were run in which the sizes of the spatial parameters were adjusted. In these experiments, the values of the ZOR, ZOA, FGS and BGS were varied so that an understanding could be gained about how they affect the model. Table 7 shows the parameters and the values they were varied with. All these values were combined so that a total of 625 experiments were created, allowing for every possible combination of ZOR, ZOA, FGS and BGS. The number of agents was limited to 20, only a single breeding ground was used, and all transition matrices were randomly initialised. Each model was run for a total of 12,000 iterations. Here, 1000 iterations roughly correspond to a single month in a year.

| **ZOR** | **ZOA** | **FGS** | **BGS** |
|---|---|---|---|
| 1 | 1 | 50 | 50 |
| 5 | 10 | 100 | 100 |
| 10 | 20 | 200 | 200 |
| 30 | 50 | 300 | 300 |
| 30 | 100 | 600 | 600 |

Table 7: The values used for the parameter exploration in MM1.

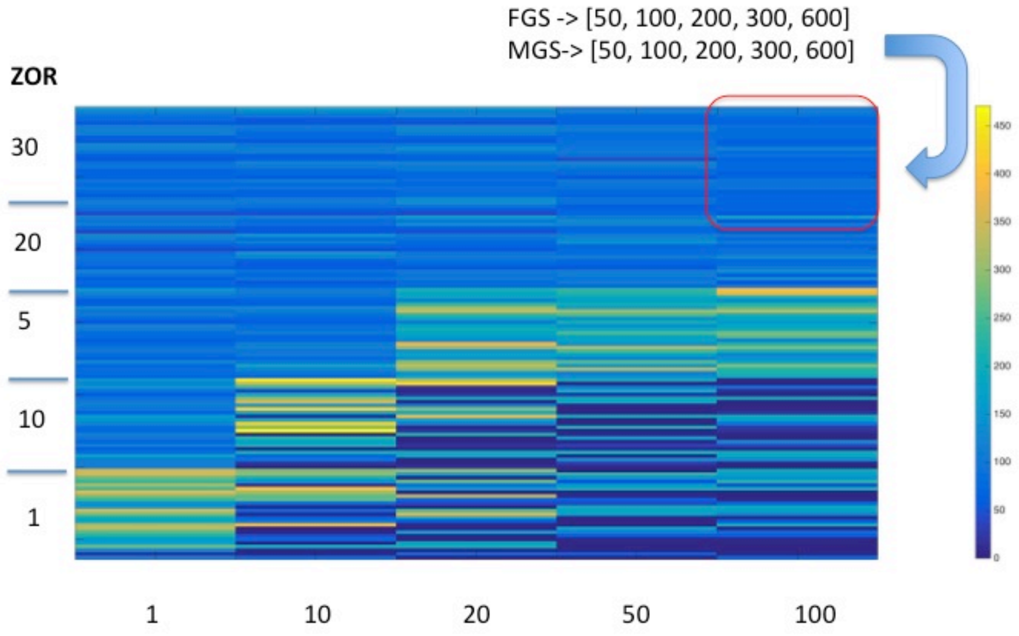


Figure 22: The matrix dissimilarity results for MM1 parameter exploration. Each column corresponds with an increase in the value of ZOA. Within the columns, each row corresponds to different combinations of FGS and BGS, with a specified value of ZOR shown on the Y-axis. In each column, after moving up 25 cells, the ZOR value increases.

Figure 22 shows the summed transition matrix dissimilarity results at the end of each model run. In this figure each cell in the matrix contains the matrix dissimilarity score for that experiment. Each column corresponds to a different ZOA value. In every column, after moving 25 rows up, the value for the ZOR increases. Those 25 cells contain every combination of FGS and BGS for that specific value of ZOR and ZOA. This is demonstrated by the red box at the top right hand corner of Figure 22, which contains the combination of all FGS and BGS sizes, combined with the ZOR value of 30, and the ZOA value of 100. The brighter the colour of the cell the greater the dissimilarity between the transition matrices of the agents within that population. If we take these meta-blocks of 25 cells and calculate the mean average, then Figure 23 is obtained. This figure describes the general behaviour of the model in relation to specific ZOR and ZOA values.

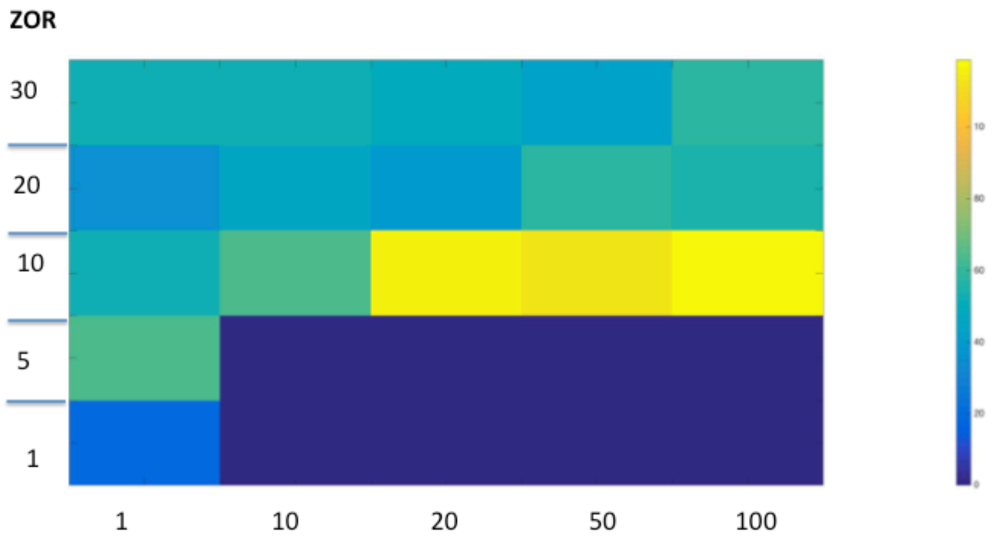


Figure 23: The mean of the cell experiment metablocks.

Figure 23, shows that the ZOR and ZOA have a major influence on the SRs of our agents. If our ZOA value is greater than ten, with a ZOR value lower than 10, our agents converge completely on identical transition matrices. Interestingly, a ZOR of 10 combined with a ZOA value of greater than 20 returns the highest dissimilarity score. This is due to the fact that once our ZOR reaches a high enough value the agents will never be able to enter each others ZOA's, not allowing them to get close enough to learn another agents song. From this diagram, we can learn certain things about the model:

1. A model run with a high ZOA (of 10 or greater), coupled with a ZOR of less than 10, results in all agents in a population converging on the same song.
2. A high ZOR (of 20 or greater) returns a relatively consistent level of dissimilarity across runs, no matter what the size of the ZOA is. This is likely due to agents not being able to get close enough to each other to learn each other's songs.
3. A ZOA of 20 or greater, and a ZOR of 10, results in very high dissimilarity across the population. This is due to the agents trying to approach one another, but only being able to learn a very small portion of another agent's song. Since the matrices are randomly initialised, this returns highly dissimilar matrices.

These insights, while they may seem trivial, inform us about the behaviour of the model. With the exception of scenario 1, most of these situations are not biologically plausible, but they do inform us about the limits of our model and narrow the parameter space that needs to be explored considerably.

In order to explore the parameter space of the Migratory Model 1 further, a series of 96 experiments were created that varied the BGS, FGS, ZOR, ZOA, SR size (number of potential units), and population size. In every experiment, only one breeding ground was used, and the model was carried out over 12,000 iterations. After running these experiments, the SR dissimilarity was calculated for the start of the each model (iteration 1), and the end of each model (iteration 12,000). The difference between these two values was calculated in order to generate the Delta Mean SR Dissimilarity (DMSRD). Due to the large parameter space, a novel plotting approach was required, as shown in Figure 24. This figure shows the delta mean SR dissimilarity on the Y-axis, and the ZOA on the X-axis for each of the four boxes. Each box corresponds with a different sized population and SR. So, for example, the top left hand box corresponds with a population size of 10, a matrix size of 5. To distinguish each model, unique symbols were applied for the different combinations of BGS, FGS, and ZOR, as shown in the table on the far right of the figure. The DMSRD informs us of how our model converged in throughout the entire experiment, so a higher DMSRD corresponds with models converging on similar songs and SR's.

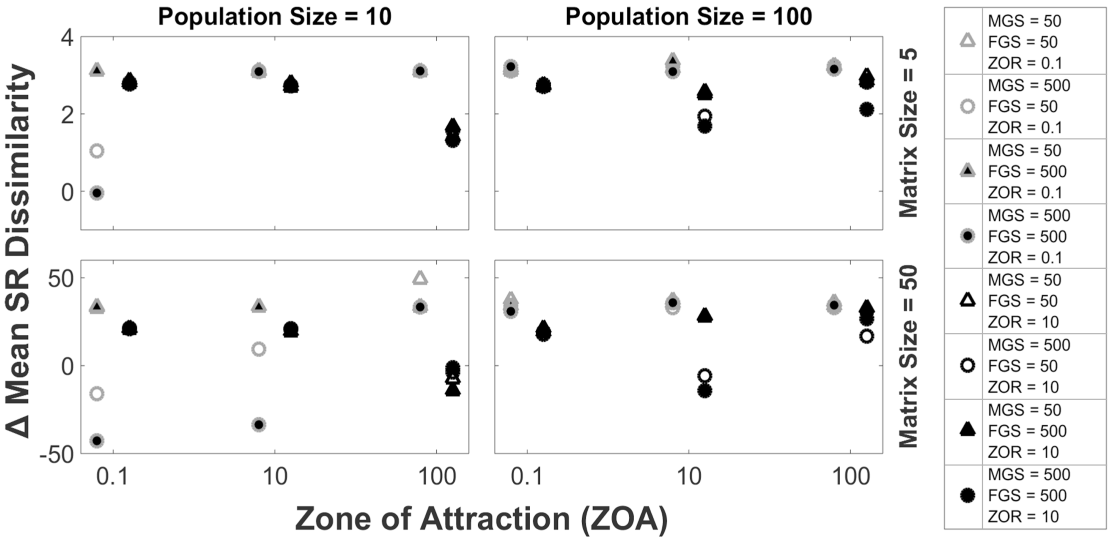


Figure 24: This figure shows the results of the parameter exploration in the model. It shows the delta-mean SR dissimilarity calculated for every experiment carried out for the parameter exploration, a total of 32 experiments.

Figure 24 gives us an idea of how Migratory Model 1 behaves in regard to these various influences. Looking in the top left hand box, we can see that a high FGS, BGS, and low ZOR and ZOA (Symbol 4), results in the agents not converging on the similar SR's. However, if we increase the population size, the agents will converge on more similar SR's since they are in closer proximity to each other. From this we can see that an average distance between the agents is influential in determining whether they will converge on similar songs or not. The key point being that as the concentration of agents within a certain area increases, the DMSRD also increases. This is again reflected in that across these experiments, DMSRD increases as ZOA increases, due to the influence of the attraction rule discussed in the methods section. This is best shown by the cluster of experiments in the top right hand box, with a high ZOA value. This clustering informs us that the model is behaving as expected, and that parameters that have any influence on the distance between the agents will affect the DMSRD.

## 3.7 Results

### 3.7.1 Model 1: Distance

All models were run with the same set of initial parameters in order to make results comparable across the different learning scenarios. All models were run with the three different feeding ground sizes (FGS = 50, 100, and 500) and, unless explicitly stated, agent SR matrices were initialised with each element drawn from a uniform random distribution in the range [0 1] and then normalised. For each of the following modelling scenarios 50 model experiments were run to get a representative view of the model's behaviour. Each model was run with 30 agents, with agents 1-15 being assigned to breeding ground 1, and agents 16 to 30 assigned to breeding ground 2.

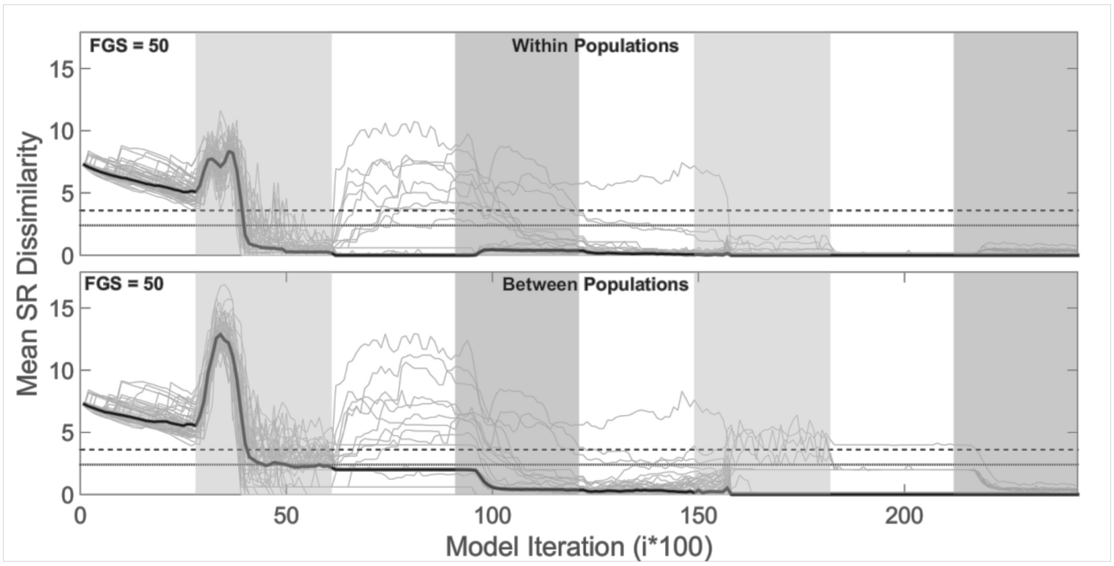


Figure 25: Mean SR dissimilarity calculated every 100th iteration (total number of iterations: 24,000) across the population of agents of model 1. The light grey lines represent 50 modelling experiments carried out in this scenario while the thick black line represents the median of them. It should be noted that because many of experiments tend towards zero later in the run, the median appears distorted at certain instances in these figures. Mean SR dissimilarity within populations is reported in the upper panel while the one between populations is plotted in the lower panel. The dark and light grey areas represent breeding and feeding seasons respectively. In this scenario FGS = 50. The horizontal dashed and dotted lines represent the mean SR dissimilarity estimates calculated respectively in 2002 and 2003, at the end breeding grounds in eastern Australia.

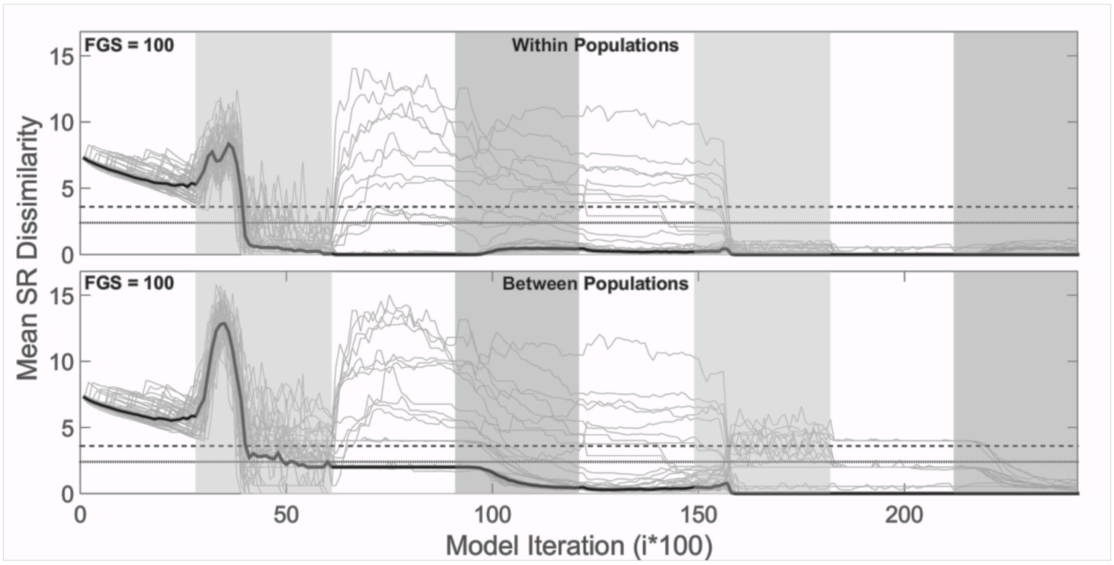


Figure 26: Mean SR dissimilarity. FGS = 100 for model 1.

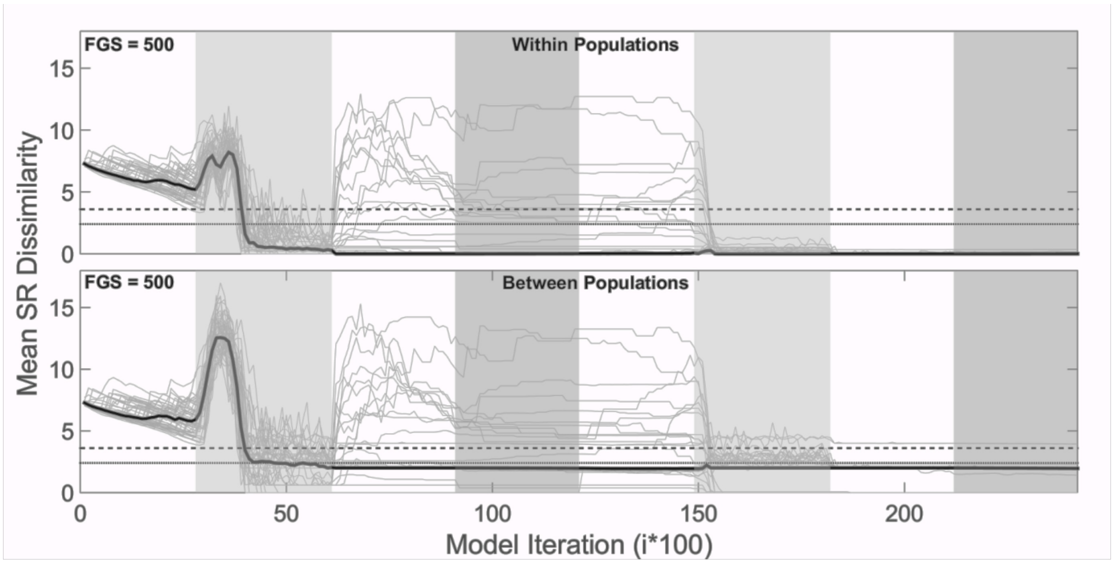


Figure 27: Mean SR dissimilarity over time for FGS = 500 for model 1.

In all runs the mean within-population SR dissimilarity decreased rapidly during the first breeding season (Figure 25, Figure 26, and Figure 27). When FGS was set to 50, once mean SR dissimilarity reached 0 it remained generally low across the remainder of the experiments with the exception of the first feeding season, in which a slight increase was observed in smaller feeding grounds due to the mix of agents from the two breeding populations (with different SRs). Mean between-population SR dissimilarity decreased during the first feeding season as agents returned to a small feeding ground (Figure 25, lower panel). At an intermediate feeding ground size (FGS = 100, Figure 26), between-population dissimilarity decreased less sharply during the first feeding season (compared to FGS = 50,

Figure 25) but then increased a bit more during the second breeding season once the populations separated (Figure 26, lower panel). If the feeding ground was large enough that the two breeding populations never met (FGS = 500), the mean SR dissimilarity between them remained constant across the two migration cycles (Figure 27, lower panel), indicating divergence between populations at the same time as convergence within each population itself. Generally, song sequences produced in all scenarios using model 1 were short. This was due to the agents' convergence on sparse SR matrices with transition probabilities made of 0s and 1s.

### 3.7.2 Model 2: Distance + Novelty bias

The introduction of novelty led to a greater level of SR dissimilarity both within and between populations. During the breeding season, SR dissimilarity between populations increased, showing that the separated populations' SRs diverged (Figure 28). While SR dissimilarity between populations increased, it was generally consistent during the breeding season within populations. It would then decrease within populations before beginning to rise again during the breeding season. During both breeding seasons, the within-population SR dissimilarity converged to roughly the same levels. The behaviour of the novelty algorithm was unusual, and this is discussed further in section 6.3.

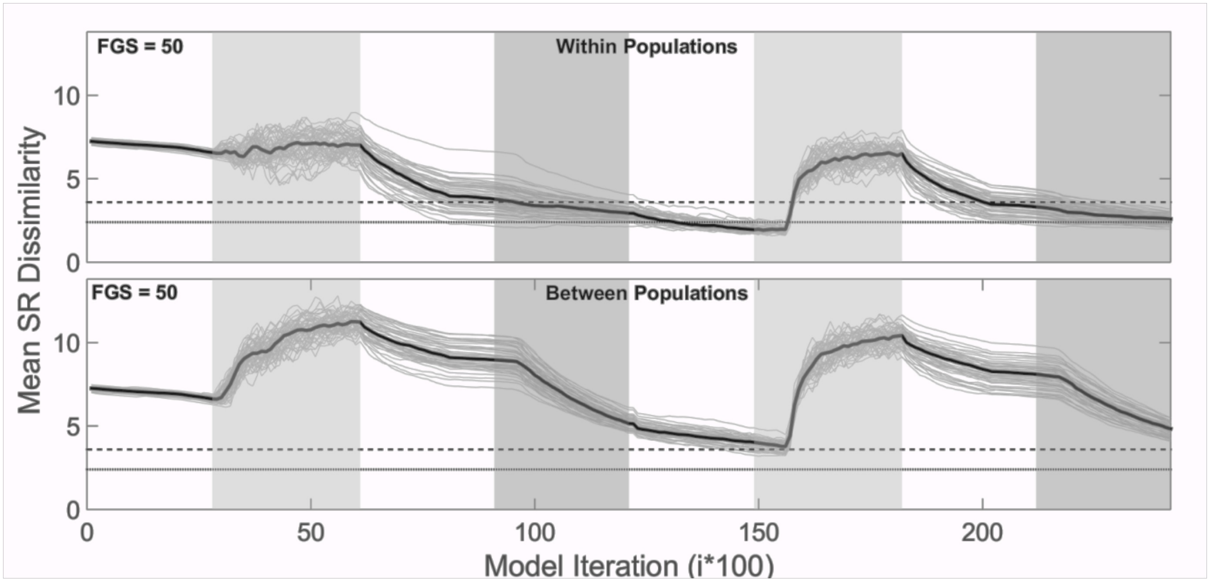


Figure 28: Mean SR dissimilarity over time for FGS = 50 for model 2.

### 3.7.3 Model 3: Distance + Production error

In this scenario the distance algorithm from model 1 was coupled with weighted-edit production errors. Although these models were run with the usual three feeding ground sizes (FGS = 50, 100 and 500), we present here only the results relative to FGS = 50 in order to simplify the presentation of results under the three different edit probabilities. The results of experiments with FGS = 100 and 500 can be found in Appendix 2. The introduction of song production errors triggered more abrupt fluctuations in the mean SR dissimilarity compared to previous results (Figure 29, Figure 30, and Figure 31). Despite different error probabilities, during each feeding season the two breeding populations still converged on similar SRs while diverging between populations. Higher error probabilities ($p_e$ = 0.01, 0.1) however limit the degree of conformity within populations. The introduction of production errors also increased the mean between-population SR dissimilarity during the breeding seasons. Furthermore, higher error probabilities resulted in increased mean SR dissimilarity both between and within populations across the entire experiment (Figure 30 and Figure 31). Compared to model 1, feeding seasons of model 3 present more pronounced SR dissimilarity fluctuations, especially for error probabilities above 0.001 (Figure 30 and Figure 31), even though we still see convergence between the populations in this small feeding ground scenario (FGS=50). To test whether this model scenario gave a genuinely different outcome, as opposed to simply slowing down the trends seen in model 1, we ran a model for 10 migration cycles (FGS=50, p=0.001), and confirmed that production errors kept the populations from achieving complete within-population convergence over these timescales (Figure 32).

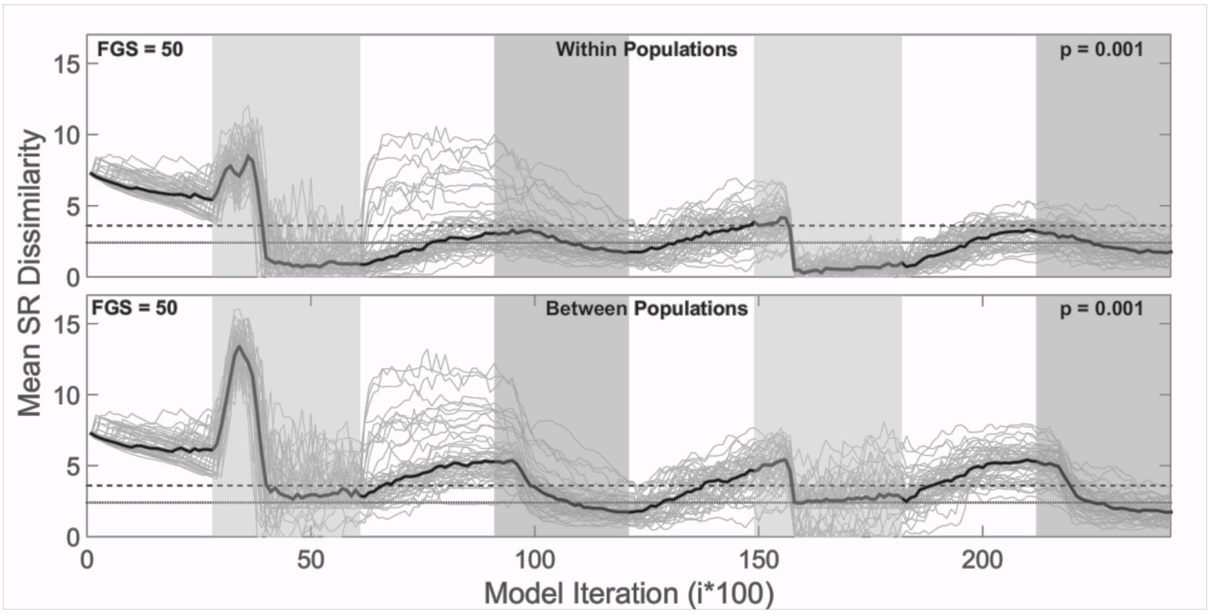


Figure 29: Mean SR dissimilarity over time for FGS = 50 and p = 0.001 for model 3.

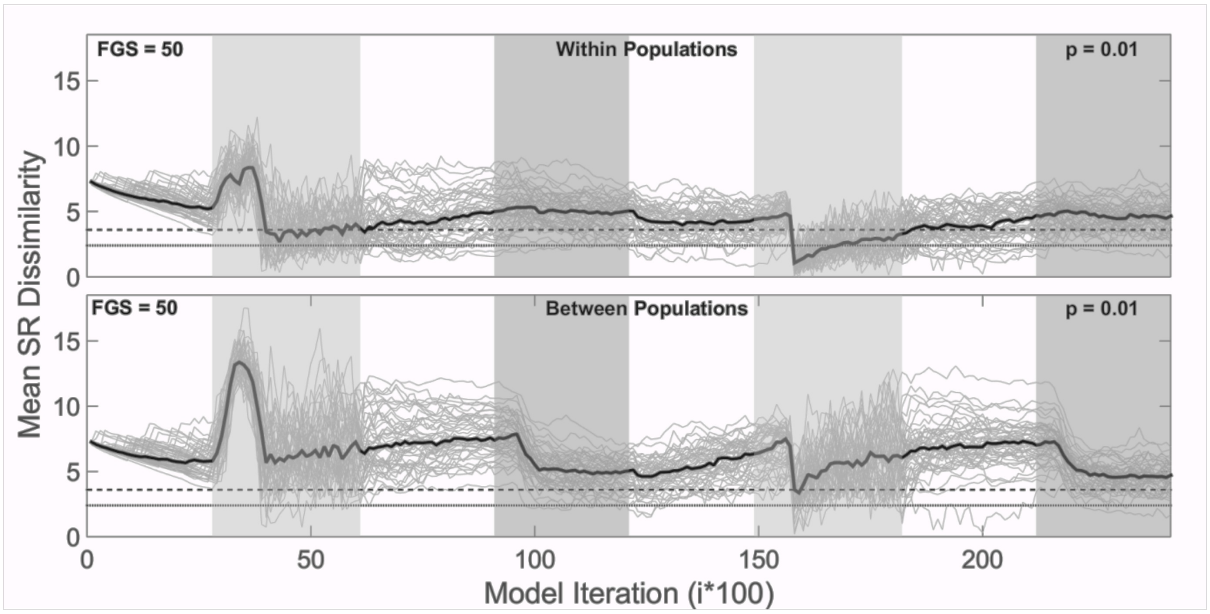


Figure 30: Mean SR dissimilarity over time for FGS = 50 and p = 0.01 for model 3.

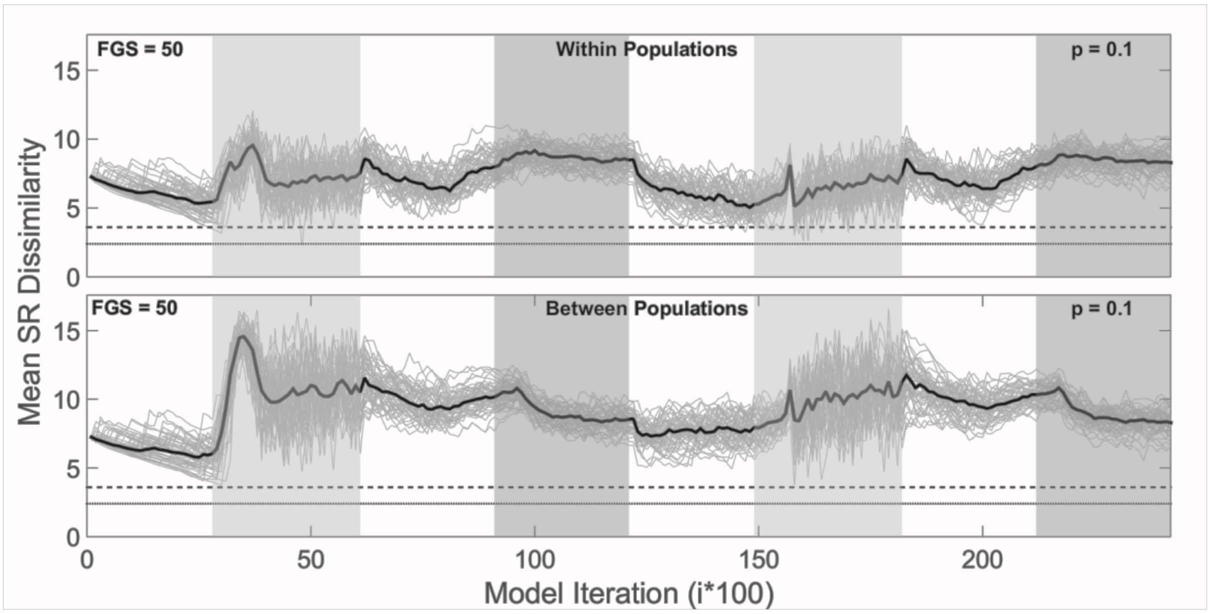


Figure 31: : Mean SR dissimilarity over time for FGS = 50 and p = 0.1 for model 3.

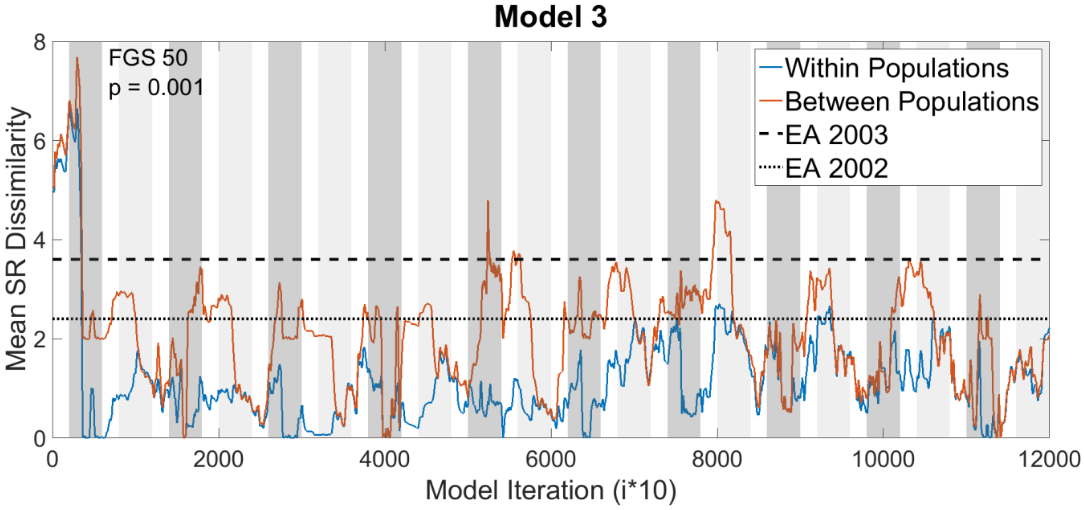


Figure 32: Mean SR dissimilarity calculated for every 100th iteration for 10 migration cycles for model 3.

### 3.7.4 Model 4: Distance + Novelty bias + Production error

In model 4, the design of model 2 was coupled with the weighted-edits algorithm to test how song production errors might alter the effect of novelty bias on the cultural evolution of song. Similarly to model 3, only results from an experiment with a small feeding ground (FGS=50) are presented (the rest are available in Appendix 2). The introduction of song production error did not qualitatively change the impact of novelty bias, as the results obtained were similar to those for model 2 (Figure 28) However, high song production error ($p_e$ = 0.1) lead to a steeper increase in mean SR dissimilarity during the first breeding season of the simulations compared to lower production errors ($p_e$ = 0.001 and 0.01, Figure 33 and Figure 34 respectively) and results of model 2 (Figure 8, 9 and 10). In all cases, there was a lack of divergence between the breeding populations as the meta-population as a whole converged to a low SR dissimilarity condition.

Even with a very low edit probability ($p_e$ = 0.001, Figure 33) both breeding populations never reach complete convergence, i.e., mean SR dissimilarity is always greater than zero. The mean SR dissimilarity trends shown in Figure 33, Figure 34, Figure 35 are also consistent over a larger number of migration cycles. There is a pronounced cyclical pattern of increasing variation (i.e., increasing dissimilarity) during breeding seasons when populations are segregated, which is then erased by the rapid learning of any new variations by the wider meta-population once they are reunited on the feeding grounds (Figure 36).

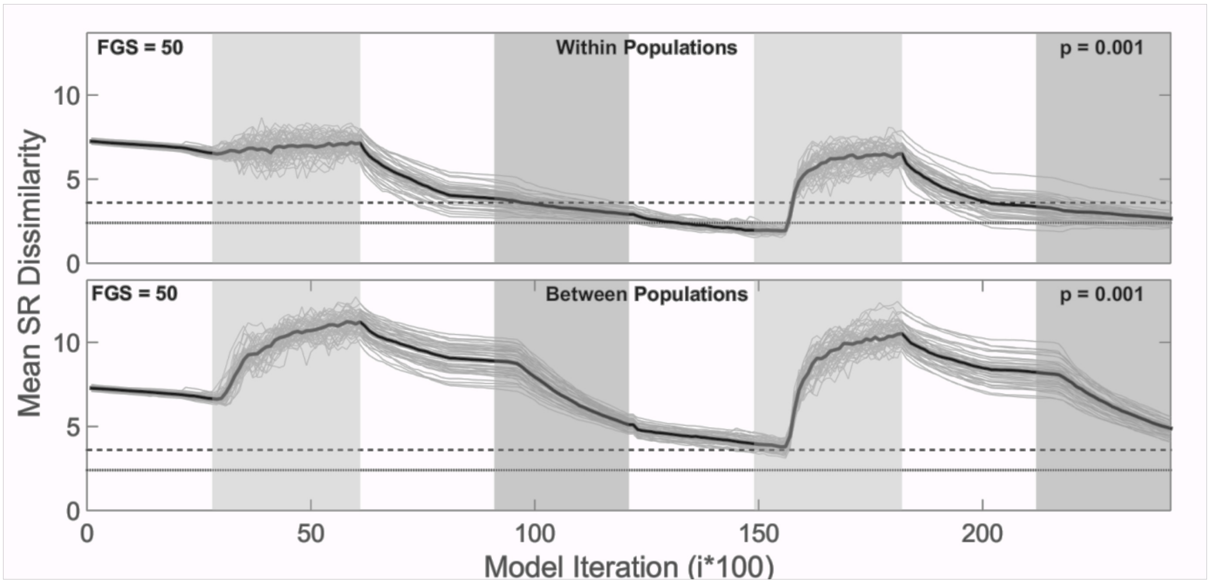


Figure 33: : Mean SR dissimilarity over time for FGS = 50 and p = 0.001 for model 4.

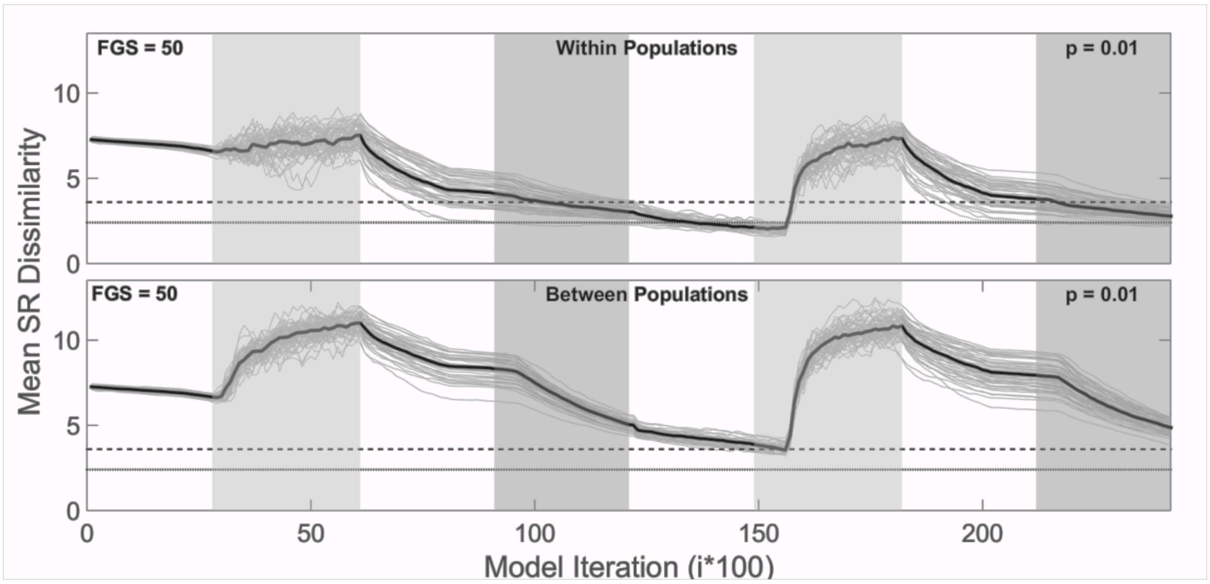


Figure 34: : Mean SR dissimilarity over time for FGS = 50 and p = 0.01 for model 4.

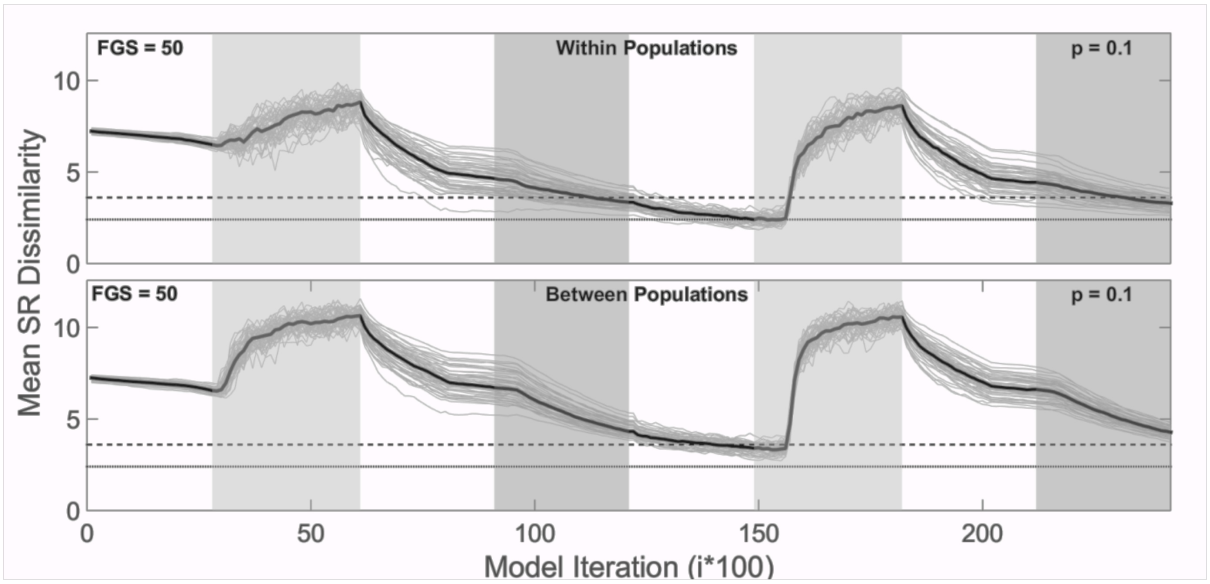


Figure 35: Mean SR dissimilarity over time for FGS = 50 and p = 0.1 for model 4.

.

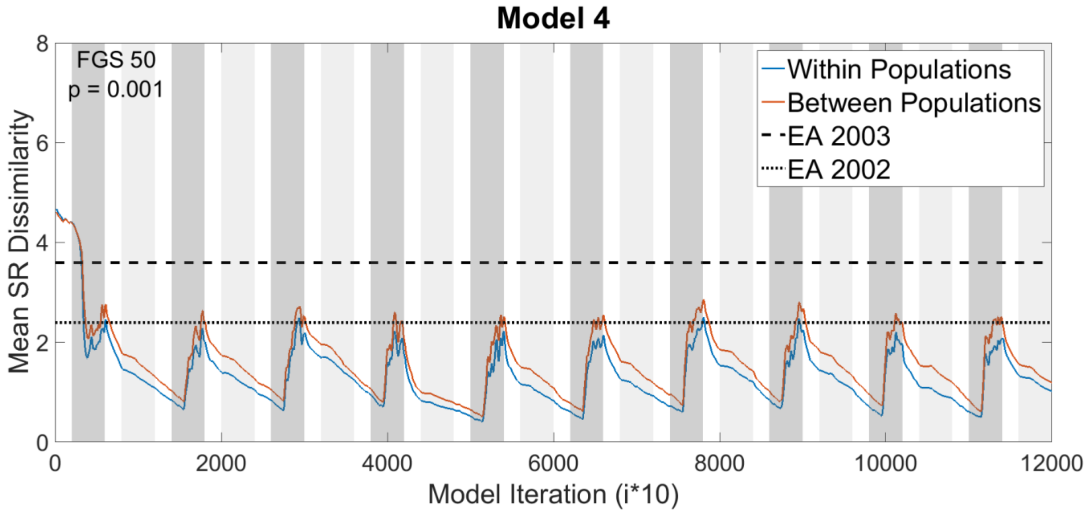


Figure 36: Mean SR dissimilarity every 100th iteration for model 4 carried out over 10 migration cycles.

### 3.8 Discussion

This study was inspired by a broad spectrum of examples of vocal convergence across taxa; these have been highlighted by different disciplines such as biology, linguistics and music. Moreover, the study of animal vocal convergence is relevant in unveiling the characteristics of animal social structures (Tyack, 2008). We aimed to shed light on the kinds of learning mechanisms that could underlie one of the most striking examples of non-human vocal convergence, humpback song. To do this we developed a spatially explicit multi-agent model to study the cultural evolution of humpback whale song. Different learning biases were tested, and their respective results were compared to real song references. Our modelling used realistic assumptions about the spatial structure of migrating populations to arrive at a number of new insights about how these processes affect song behaviour in humpback whales.

First, our model in which the only factor controlling song learning was distance from the singer, based on empirically realistic transmission loss, produced extreme convergence within breeding populations. These are unrealistically extreme when compared to empirical measures of convergence from the eastern Australian humpback population. Varying the feeding ground size, and thus the extent to which members of the two populations were exposed to each other's song during the feeding season, dramatically altered the extent of between-population divergence – small feeding grounds, on which the populations were forced to mix, prevented divergence between populations, while large feeding grounds, where mixing was much rarer, showed strong between-population divergence. Thus this simplest of our models showed how the spatial arrangement of feeding and breeding grounds can produce quite different cultural evolution outcomes even when the underlying learning mechanisms are the same. This confirms the predictions of feeding grounds and migratory routes as key locations for song transmission (Garland et al., 2011; Garland et al., 2013). Contrary to observations in the wild, however, the length of songs produced by this model decreased drastically across the model experiment, and by the end of the model runs agents showed a high degree of song conformity on very short songs. In cultural transmission, signals may decrease in length. For example, the range of movement in an invented sign language decreased over multiple generations of an iterated learning model. In Kirby's iterated learning models (Kirby, 2002), the number of rules in a language may decrease over multiple generations of the model and the system approaches an optimal level where the minimum number of rules required in order for a language to be expressive is reached. However, it seems more likely that the decrease in length in our model is a by-product of the learning algorithm used here. Songs do not evolve within this scenario, because when complete song convergence is reached, the population's song system stagnates, fixed on purely 1/0 SR transition matrices, unless a new song is introduced (which can happen when two breeding populations with different songs mix on the feeding ground).

Second, given that a simple distance parameter did not lead to song evolution through time in a population, it was necessary to add a new component to the model to try and understand how a population of agents could show song evolution by the simplest mechanisms possible. Song revolutions recorded in eastern Australia (Noad et al., 2000) indicate that males might be preferentially attracted to novel or different songs introduced by immigrant conspecifics from western Australia, so we introduced a novelty bias in song learning. The novelty model did not result in the dramatic convergence observed in Model 1. Rather agents maintained a relatively constant level of SR dissimilarity within in their populations. It also resulted in a larger degree of SR dissimilarity between populations. This dramatic difference shows that while novelty does not lead to a gradually changing song, it does result in populations diverging their songs, as observed in the wild. This shows that when a cognitive bias is included in song learning, the SRs of the populations will diverge. These results matched well with the observed situation in the South Pacific, where there is clear divergence between breeding populations (Garland et al., 2011). Issues surrounding the novelty algorithm are discussed in more detail in Chapter 6 in the section titled *Developing a New Novelty Algorithm*.

Neither model 1 (distance) nor model 2 (distance + novelty bias) produced gradually evolving songs, so were not sufficient to explain observed song variation. To produce continual evolutionary song change, some mechanism was required to prevent populations 'fixing' on purely 1/0 SR matrices from which no variation could occur. In order to address this we introduced the assumption of song production errors, based on a weighted-edits algorithm. Informed by humpback whale song literature describing within-population song variation (Payne et al., 1983; Cerchio et al., 2001) we assigned a high probability of theme addition, with theme substitution and deletion being possible, but significantly less likely. The addition of production errors significantly changed the song evolution dynamics in the model. Rather than agents converging on identical transition matrices, they instead maintained a level of dissimilarity that oscillated to varying degrees depending on the probability of production errors. Model 4 appeared to show that novelty mitigated the induced production errors, producing results very similar to model 2.

All models are thought experiments that force scientists to abstract out many real-world details, but the model we have presented here, while no different, has been closely informed where possible by empirical observations to produce an informed model of how the cultural evolution of humpback whale song might emerge from spatial structure and simple learning and production rules. This type of complex modelling is necessary when we consider the cost and difficulty associated with studying these behaviours in the field. Modelling for the purpose of studying vocal convergence is not a new idea. It has been used in several fields such as biology, linguistics and music (Todd and Werner, 1999; Kirby, 2001; de Boer, 2002; Miranda et al., 2010). While these models study vocal convergence, they do so in strictly defined systems. The specificity of these systems is one of the factors that led to the choice of first order Markov models as a song learning/production substrate in our model. It presents a simple song production and learning system that makes the least assumptions about the cognitive capabilities of humpback whales whilst also allowing us to incorporate other influential factors that may impact song learning. Our models suggest that migratory movements and spatial factors play an important role in the transmission of song, but when combined with a simple learning system like a first order Markov model, these factors do not sufficiently explain the patterns of both evolution and revolution. This is not unexpected, given the concerns raised by other researchers in regards to animal vocalisations and Markov models(Kershenbaum et al., 2014; Suzuki et al., 2006). However, we consider these Markov models as placeholders that should ultimately be replaced by a way of modelling song production that is more closely informed by data from real humpback songs, once they become available.

Humpback whale social learning is, of course, a biological system and will be subject to variance in many ways that have not been captured in the current model. For instance, variance among listeners in the rate of song learning in general, and uptake of novelties in particular, are likely to generate asymmetries that may be important in preventing convergence among populations. Nonetheless, our model produces a number of interesting suggestions by modelling the interaction between humpback whales on the breeding ground, how migratory movements influence song learning, how the size of feeding grounds may impact transmission, and how the acoustic loss in transmission of song over distance, among other factors, might influence song learning. It also highlights that even when song occurrence on the feeding grounds is low, it can have a strong influence on the songs of geographically separated populations as long as they meet on the feeding grounds, resulting in between-population song convergence. However, these behaviours are not necessarily the same as the revolutionary behaviour observed in the wild. For example it may more closely resemble the behaviour of song mixing of two populations.

By using methods inspired by computational research into the origin of music and music composition, we have developed a multi-agent model that simulates the migratory movements, interactions and singing behaviour of humpback whales. Incorporating a sound transmission loss factor into our model allowed the simulation of song convergence within separate breeding populations and simultaneous divergence between populations. Song convergence between populations also occurred when feeding ground sizes were small enough, highlighting the importance of the feeding grounds of humpback whales as being a key location for song cultural transmission, as hypothesised by Garland (Garland et al., 2011; Garland et al., 2013). A novelty bias was found to result in strong divergence in geographically separated populations, suggesting that some form of cognitive bias is responsible for these populations' diverging songs. Induced production errors mitigated the rapid song convergence observed in distance-only model (MM1), increasing the level of song dissimilarity within a population. Finally the coupling of novelty and production errors resulted in similar results to the novelty-only model (MM2), with separated populations diverging to different songs. These results are similar to those observed in the wild and point towards scenarios where cultural revolution may take place.

# Chapter 4 - Social Network Model

## 4.1 Introduction

This chapter discusses the development of an agent-based model to investigate the cultural transmission of humpback whale song in different social networks. It is inspired by conversations with Jenny Allen at the Cetacean Ecology and Acoustics Laboratory, at Queensland University. During our conversations, Allen and I discussed potential networks that would facilitate song convergence in a population of whales, discussing in particular small world networks. This model investigates lattice, small world, and random networks in order to see how well they facilitate convergence in a population. Grid based models are not discussed since there have been previous models that investigate bird song transmission in grids in order to determine which agents will learn from each other (Goodfellow & Slater, 1986; Slater, 1986).

Models used to investigate cultural transmission in sperm whales leads to complex hierarchies emerging due to cultural transmission (Maurício Cantor et al., 2015). These models however focus on the emergence of specific, hierarchical social networks emerging, rather than being imposed from the very beginning. In a single population of humpback whales, all whales will sing the same song at a given time, with the main differences likely being caused by copying errors. Since a grid social network would lead to the emergence of multiple dialects, it was determined that this type of social network inappropriate for facilitating song conformity across the population. Models **that** utilise vertical cultural transmission are useful for identifying changes that can occur from one individual to another (S. Kirby, 2001; Eduardo Reck Miranda et al., 2003a), but these systems deal with changes from one generation to another. The research of Cantor also shows that complex social structures can emerge from cultural transmission. This type of model allows us to generate theories on what type of social structures male humpback whales may have once they are in place.

Since there is no real world data to compare it to, here two questions are asked in regards to social networks and song transmission in humpbacks. 1) What is the minimum number of conspecifics whales need to learn from in order for a high degree of song convergence? 2) Which social network structure best facilitates song convergence?

## 4.2 Generating The Network

In this model, the small world network is generated using the algorithm known as the Watts-Strogattz algorithm (Watts & Strogatz, 1998). The Watts-Strogatz model generates a random network graph with small world properties. These properties focus on short average path lengths and high clustering in the social network. A short average path length means that we can get from one node in the network to another in very few steps (less than five), and a high clustering coefficient means that agents are connected to in neighbourhoods (IE, an agent is connected to a certain number of its neighbours). The Watts-Strogatz algorithm was chosen as it is the most common way to model small-world networks, and is readily available in most packages offering graph generation.

Given $N$ nodes, the mean degree (number of connections a node has to other nodes) $K$, the special parameter $\beta$, and the conditions that $0 \leq \beta \leq 1$ and $N >> K >> ln(N) >> 1$, the Watts-Strogattz algorithm creates an undirected graph with $N$ nodes and $\frac{NK}{2}$ edges using the following steps:

1. Generate a regular ring lattice with N nodes each connected to $k$ neighbours.
2. For every node, given the probability parameter, $p$, there is a possibility of removing a connection between that node and another random node, and connecting it to another random node in the network.

In the Watts-Strogatz algorithm, as the value $p$ increases, the network becomes more random, with a network of $p = 1$ generating the most random networks for that value of $k$.

**Figure removed due to Copyright Restrictions**

Figure 37: How varying the value $p$ affects the structure of a social network generated with the Watts-Strogatz algorithm. Taken from (Watts & Strogatz, 1998)

Relating the networks generated using the Watts-Strogatz algorithm to humpback whale social structure, we can think of the parameter *k* as referring to the number of conspecifics a whale may learn a song from. In doing this, we can begin to gain a rough estimate of how many conspecifics a whale needs to learn from in order for a song to converge in a population. The parameter $p$ can be thought of increasing the randomness of what whale learns from each other. In reality, it is unlikely that every whale in a population learns from a few conspecifics in their proximity. The whales move about and are able to interact with others in their population. However, the sizes of the eastern Australian population are much larger than the population sizes investigated in this chapter. For this reason, the investigation of song convergence in this chapter should be considered a sub-sample of a single population at a single point along their migratory route, as opposed to a large population over the course of their migration.

## 4.3 The Model

Agents in the model are equipped with a first order Markov model, enabled using a transition matrix. These Transition Matrices are termed Song Representation (SR), using the same terminology as that used for the Migratory model presented in Chapter 3. SR's are randomly synthesised for each agent. In order for an agent to learn a song, the following equation is used:

$$SR_l' = SR_l * (1 - l) + (SR_s * lr) \qquad (7)$$

Where $SR_l{}'$is the updated SR for the agent learning the song, $SR_l$ is the learning agents SR, and *lr* is the learning rate where *lr* ≤ 1.

In this model, the small world network represents the relationship between agents, and which agents may learn from each other. Every node represents an agent, and every edge (connection) represents which agent that agent may learn from and teach to. This small world network is used to generate chains of cultural transmission, inspired by the work in cultural transmission models of Simon Kirby (Simon Kirby, 2002a, 2002b). Here a chain is defined as the process through which a song spreads from agent to agent. As an example, consider the following social network in Figure 38.

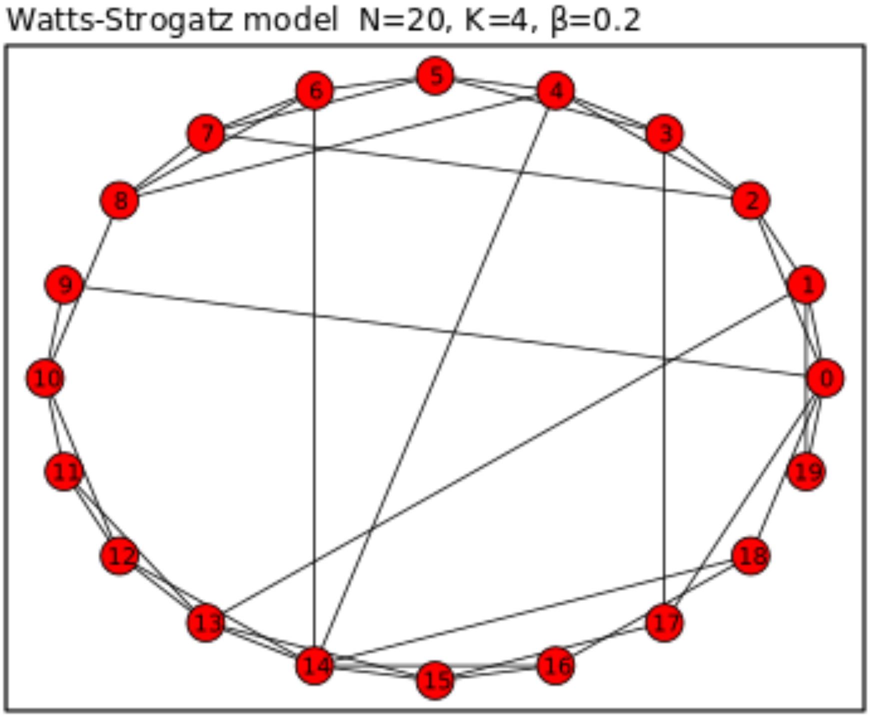


Figure 38: A small world network generated using the Watts-Strogatz algorithm.

Using this social network, we can generate a chain of cultural transmission by selecting two agents at random and calculating the shortest possible path between them. As an example, consider agent 6 and 16. The shortest path between these two agents is 6 -> 14 -> 15 -> 16. This path represents our cultural transmission chain. Here, agent 6 will produce a song, which will be learned by agent 14. Agent 6 produces a song that is learned by agent 14. Agent 14 then produces a song that will be learned by agent 15, who will finally teach the song to agent 16. A parameter called the Exposure Time parameter determines the amount of time agents spend learning a song. The reason for designing the model this way is to remove some of the random processes that may occur. In theory, there is nothing wrong with the idea of a humpback whale learning a song from a conspecific and teaching it another one, even if the conspecific it learned the song from never interacts with the conspecific it teaches the song to. This model is deliberately designed in such a way to model this type of potential cultural transmission.

## 4.4 Model Parameter Exploration

| K | P | Shortest Average Path (estimated) | Clustering Coefficient (estimated) |
|---|---|---|---|
| **3** | 0.1 | 19.08 | 0.0004 |
| | 0.4 | 12.44 | 0.0013 |
| | 0.7 | 10.86 | 0.0015 |
| | 1 | 10.48 | 0.1231 |
| **4** | 0.1 | 4.97 | 0.3671 |
| | 0.4 | 3.63 | 0.1352 |
| | 0.7 | 3.43 | 0.0449 |
| | 1 | 3.40 | 0.0324 |
| **5** | 0.1 | 4.97 | 0.2682 |
| | 0.4 | 3.64 | 0.1317 |
| | 0.7 | 3.44 | 0.0464 |
| | 1 | 3.41 | 0.0322 |
| **6** | 0.1 | 3.63 | 0.4471 |
| | 0.4 | 2.83 | 0.1645 |
| | 0.7 | 2.71 | 0.0628 |
| | 1 | 2.70 | 0.0520 |

Table 8: This table shows the results from parameter exploration.

While the Watts-Strogatz algorithm can generate small world algorithms, not every combination of *k* and *p* will produce a small world network. In order to identify networks that displayed small world network properties, multiple networks (100 for each parameter combination) with 100 nodes were generated using varying values of *k* (3 to 6), and *p* (0.1 to 1, step size equal to 0.3). For each combination of *k* and *p,* 100 networks are generated, and their shortest average path values and clustering coefficients summed and averaged. The result are summarised in Table 8. This range of values was chosen to help identify what parameters are necessary for small networks to emerge. Specifically, it was to help show what synthesised networks allow for a high average clustering coefficient, and low average shortest path length. From this table, we can see that not every combination of $k$ and $p$ will return a combination of low shortest average path lengths and high clustering coefficient. The network that performed the best was the combination of $k = 6$ and $p = 0.1$, making this parameter combination particularly noteworthy.

After calculating these averages, the model was run using the values of $k$ and $p$ shown in  Table 8. The models were run with parameters shown in Table 9. Each model was run 20 times so that the results could be averaged. This gives us a total of 320 experiments. Agents were initialised with random SR's.

| | |
|---|---|
| *Number of Agents* | *100* |
| *Model Iterations* | *10'000* |
| *Exposure Time* | *1* |
| *Lr* | *1* |
| *maxSongLength* | *30* |
| *minSongLength* | *10* |
| *Number of Units* | *17* |

Table 9: List of parameters for Social-Network Model

After running the experiments, the song that was most popular in the population was calculated, as well as the number of agents singing the most popular song. These results were plotted in Figure 39. In figure four, each experiment run was plotted using scatter plots, with the percentage of agents singing the most popular song on the Y-axis, increase in the value of $p$ at the bottom of the X-axis, and increase in the value of $k$ shown on the top of the X-axis. The bars are the summed and averaged results of each experiment with that particular value of $k$ and $p$.

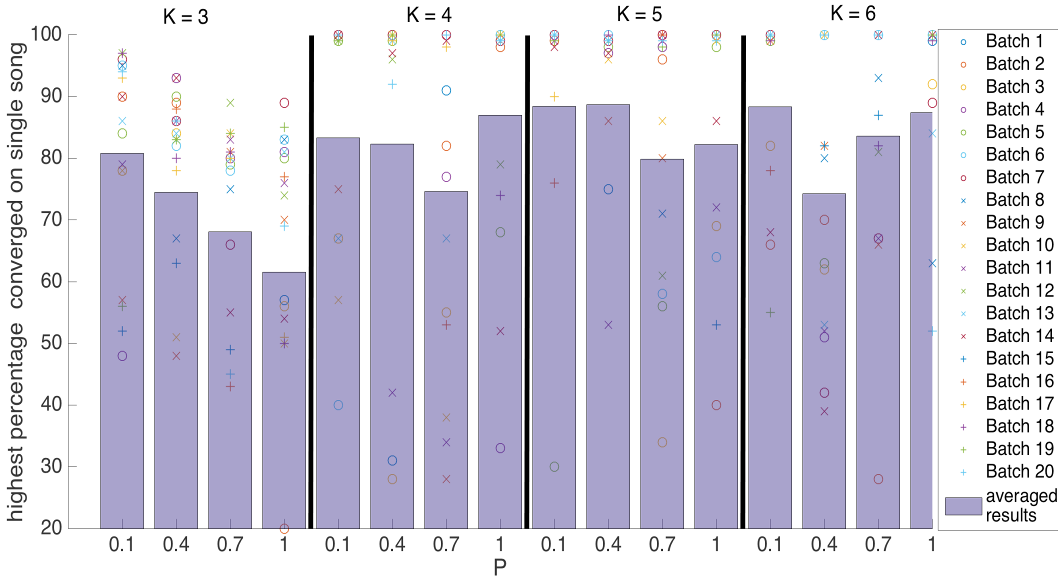


Figure 39: Results for various p values (0.1 to 1), and increasing values of k (3 to 6).

As we can see from Figure 39, it is difficult to see a clear pattern between the different combination of *k* and *p.* The only exception to this is *k = 3,* where we can see a general decrease in the number of agents singing a single song as *p* increased, and the results of individual models show some degree of clustering. However, these other results do not show any direct correlation between varying the values of *k* and *p.* This is not unexpected, and as stated earlier in this section, what is more important are the properties that define a small world network (a low average shortest path length, and a high clustering coefficient). As mentioned earlier in this section, the parameter combination of *k* = 6 and *p = 0.1* was noteworthy, as this combination displayed the best tradeoff between low average shortest path length and high clustering coefficient. We can see that this experiment returned multiple results where all agents in the population converged on the same song and the bar plot shows it outperforms the majority of other models in this series of experiments. This parameter exploration suggests that small world models outperform non-small world social networks the majority of times, even when these social networks have low average shortest path length. The parameter exploration also shows that increasing values of *p* tends to lower the clustering coefficient.

After analysing the results from parameter exploration, it became clear that it would be necessary to identify a social network with a higher clustering coefficient and low average shortest path. Since increasing *p* would lead the networks becoming more random, it was decided to increase the value of *k.* An initial value of *k = 8* was chosen as the starting point for this study, as it was the lowest value of *k* resulted in consistent small world network properties. Other values for *k* that were tested include are 10, 20, and 30. These parameter combinations and their estimated average shortest path length and clustering coefficients are shown in Table 10.

| **K** | **Average Shortest Path Length** | **Clustering Coefficient** |
|---|---|---|
| 8 | 4.92 | 0.6204 |
| 10 | 4.12 | 0.6436 |
| 20 | 2.44 | 0.6853 |
| 30 | 1.98 | 0.6998 |

Table 10: Clustering and average shortest path length estimates for social networks of varying *k* and *p = 0.1.*

After validating that all these networks had low average shortest path length and high clustering coefficients, the model was run with these parameters, a *p* value of 0.1 and the parameter combinations shown in Table 9. The results showed that 85% (68 of the 80 experiments run) of the experiments resulted in the entire population converging on the same song (68 of the 80 experiments run). The results are plotted in a bar chart shown in Figure 40.

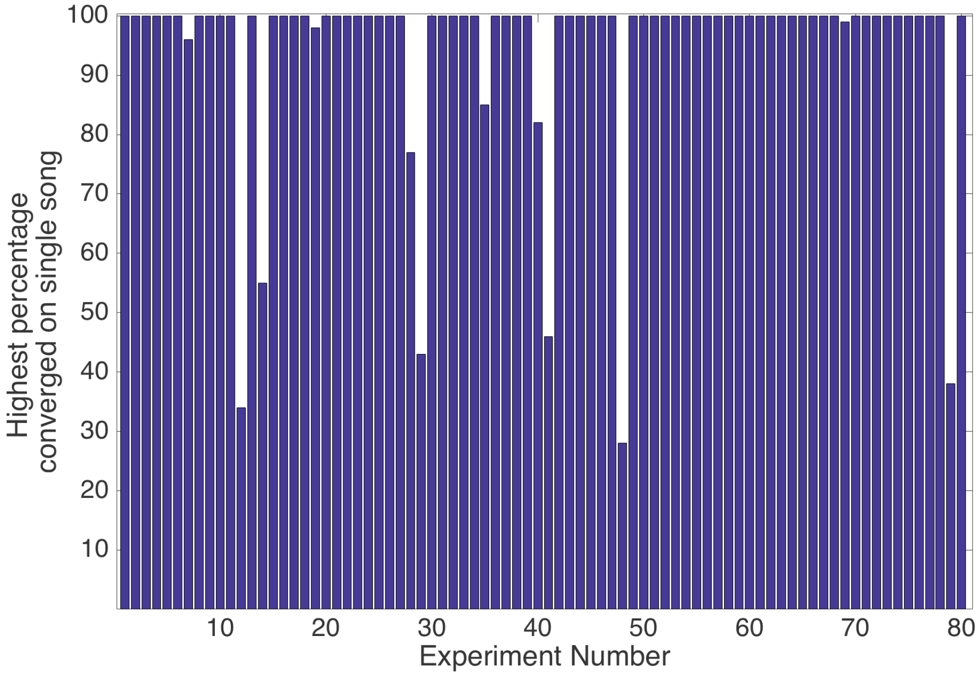


Figure 40: Results for small world experiments.

The experiments and their respective *k* values were averaged in order to produce the bar graph shown in Figure 41.

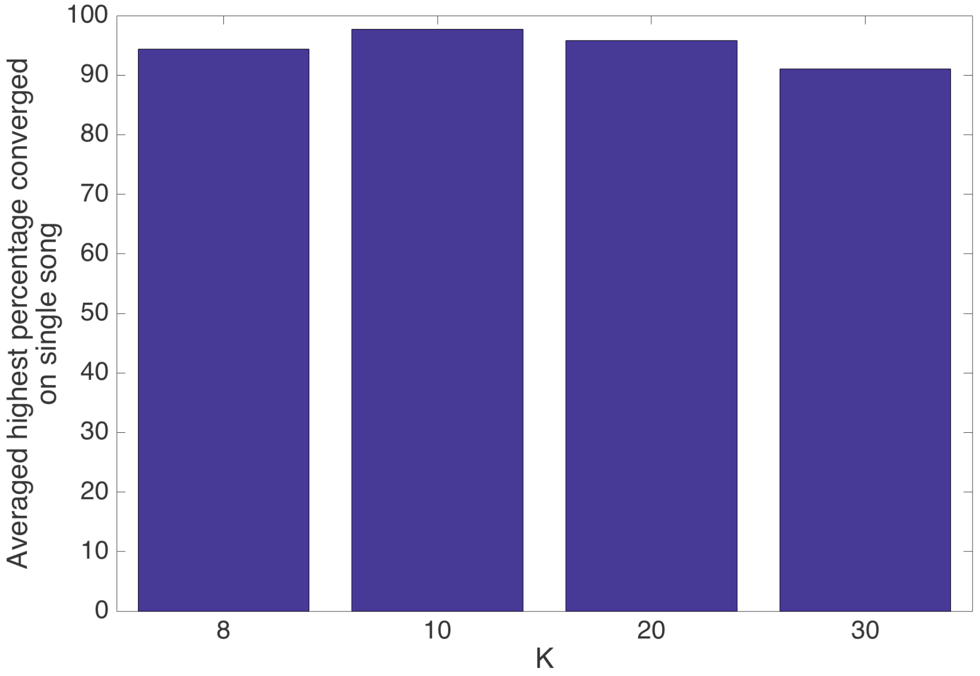


Figure 41: Averaged results. The experiments and their respective *k* value were summed and divided by 20 in order to give an average result.

Interestingly we see that the average level of song convergence increased from when *k = 10,* before beginning to decrease as *k* increased. The experiments with the paramateres *k = 30* having the lowest average levels of song convergence across the population.

## 4.5 Statistical Analysis

Three experiments were designed in order to verify that the shape of a social network will influence song convergence in a population of agents. Three social networks were chosen: a lattice network, a small world network, and a random network. These networks we synthesised using the Watts-Stogratz algorithm. The parameters used to synthesis each social network are shown in Table 11.

| Parameter | Lattice | Small World | Random |
|---|---|---|---|
| **Average Clustering Co-efficient** | **0.66** | **0.64** | **0.09** |
| **Average Shortest Path Length** | **5.45** | **4.02** | **2.2** |
| **P** | **0** | **0.01** | **1** |
| **Connected nodes** | **10** | | |

Table 11: The parameters used to synthesis each social network structure.

Each experiment was run 40 times, giving a total of 120 experiments. For each experiment, the highest number of agents converged on a single song was calculated. In other words it is a measure of how many agents in population are singing the most popular song.

### 4.5.1 Results

The results for these experiments are shown in Figure 42.

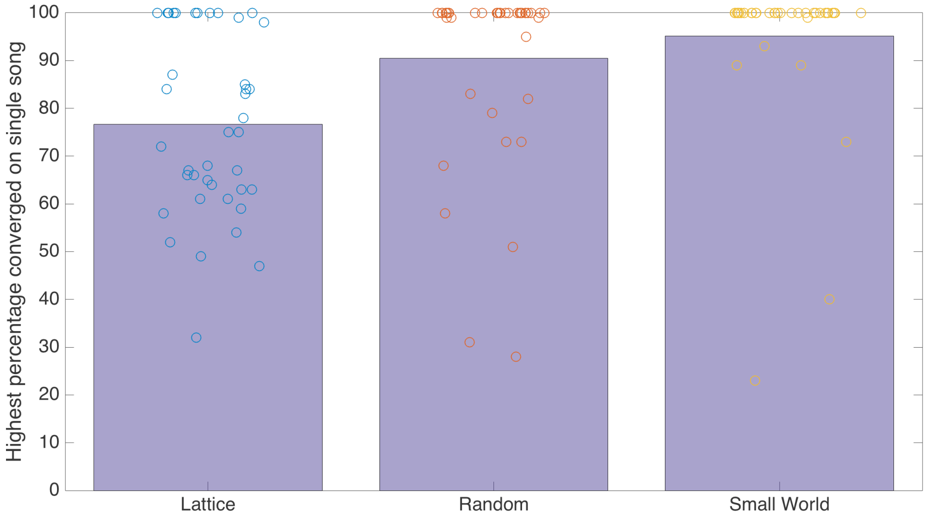


Figure 42: The results of each experiment. The dots represent the results of instance of experiment run, and the bars represent the mean of each experiments collected results.

From these results, we can see that it is possible for Lattice models to facilitate high levels of song convergence in a population of agents. However, this result is not consistent. Random models in general tended to return high levels of song convergence in a population. Small world networks outperformed all other models, with a mean of 95%, and only 5 experiments not returning total song convergence.

### 4.5.2 Analysis

In order to verify that there is a significant difference between all three models, an Analysis of Variance (ANOVA) was calculated. The result of each experiment was grouped according to their respective models, and the *p-value* that the mean of all the groups are equal was calculated. The results are shown in Table 12.

| **Source** | **Sum of Squares** | **Degrees of Freedom** | **Mean Squared Error** | **F-Statistic** | **Prob>F (*P-value*)** |
|---|---|---|---|---|---|
| **Between Groups Variation** | **6886.6** | **2** | **3443.28** | **10.38** | **0.000071** |
| **Within-groups variation** | **38491.2** | **116** | **331.81** | | |
| **Total** | **45376.8** | **118** | | | |

Table 12: ANOVA results for the three models.

In order to investigate the difference between each the results of each model, an independent two sample Student's t-test was carried out. The results of this analysis are shown in Table 13. Our t-test here is examining the null hypothesis that the pairwise difference between the data vectors has a mean equal to zero.

| **t-test name** | ***p-value*** | **tstat** | **Degrees of Freedom** | **Standard Deviation** |
|---|---|---|---|---|
| **Lattice** | **0.0016** | **3.26** | **78** | **18.91** |
| **Small World** | **0.33** | **-0.978** | **78** | **17.87** |

Table 13: t-test results comparing the random model with the lattice and small world networks.

From the results seen in Table 13, the t-test rejects the null hypothesis that the mean difference between the data for the lattice and random networks is equal to zero. However, the small world and random networks do have a pairwise difference between the data with a mean equal to zero. This suggests that there is a significant difference between the results of the lattice and random networks, but not between the results of the random and small world networks.

## 4.6 Discussion

The ANOVA test highlights that the type of network you run this model with will have an effect on the degree of song convergence in a population. Further analysis using t-test shows that there is a significant difference in mean level of convergence between lattice networks and random networks. However, a t-test comparing small world networks and random networks showed that there is no significant difference in the mean level of convergence in the population. Both networks showed a high mean of song convergence (90% for random networks and 95% for small world networks). Statistical analysis suggests that there is no difference between these networks. In regard to network structure, both networks have lower average shortest path lengths when compared to the lattice network. The small world network has a much higher level of average clustering. This suggests that short average path length may be what is most important for facilitating song convergence in a population, but further statistical tests are necessary to determine if this is true.

It is necessary to be wary when analysing tests of statistical significance in regards to simulation data (White, Rassweiler, Samhouri, Stier, & White, 2014). This is because that very large sample sizes can lead to incredibly low p-values since they are determined by replication, which can be very high in a simulation context. For this reason, choosing a sample size for the conditions presented here was difficult. Increasing the sample size may have resulted in changing the p-value, and White et al. have shown that by increasing sample size it is possible to achieve any desired p-value. They argue that it is more appropriate to focus on evaluating the magnitude of differences between simulations. This made it difficult to determine what is an appropriate sample size when carrying out the statistical tests in this chapter. It has been shown that as the number of samples collected increases, the distribution of the collected data will approach normality (central limit theorem). The samples used in each experimental setting when analysed were technically not normally distributed but increasing sample size would have resulted in the datasets approaching normality. However, this could have also lead to a decreasing p-value for our analysis. Furthermore, there is a common belief that the t-test is only valid for normally distributed outcomes, but researchers have shown that the results of a t-test are still valid even when used in extremely non-Normal data(Lumley, Diehr, Emerson, & Chen, 2002). Our statistical significance tests point to there being no significant difference between the results random and small world experiments. It should be noted though that even if we take the approach suggested by White et al and examine only the magnitude of difference between the small world model and random model, we do not obtain what appears to be a significant result (mean difference of 5).

In order to resolve this, it is useful to consider what happens in regard to real world humpback whales. The results show that small world networks and random networks achieve one of the most important aspects of humpback whale song through the convergence of a population that is in acoustic contact with each other onto the same song (R S Payne & McVay, 1971). Interaction between singing males appears to be involved in intrasexual interactions, with males observed approaching other singing males (Smith, Joshua N, Goldizen, Anne W., Dunlop, Rebecca A., Noad, 2008). This type of behaviour could be completely random, with the conspecific they approach being determined entirely random. This means that random networks seem more likely, as they facilitate high levels of song convergence, and do not enforce a strict social structure on the animals.

In regards to agent based models in other cetacean research, the most relevant research focussed on how cultural transmission may lead to the low levels of mtDNA diversity among matrilineal whales (Hal Whitehead, 2005), and the how social learning may be responsible for the clan social structure found in Sperm whales (Maurício Cantor et al., 2015). Matrilineal social networks focus on distinct family units where females remain in the same group, while the males leave to join other groups. In this regard, it is tempting to equate a clan in sperm whales with a cluster of nodes in a social network (our *k* variable), and male interactions with other clans could be viewed as the low re-wiring probability that connects these different social units. However, (Maurício Cantor et al., 2015; Mauricio Cantor & Whitehead, 2013) never explicitly refer to these networks as such, and further analysis would be required to determine if these are small world networks.

This model does not allow for the modelling of the other most important aspect of humpback whale song discussed in this thesis; revolution (Garland, Noad, et al., 2013; Noad, Michael J., Cato, Douglas H., Bryden, M. M., Micheline, Jenner, Jenner, 2000). In order to achieve this, it may be possible to generate two small world networks and have a few connections in order to determine how many connections are required to facilitate revolution. In regards the slow evolution of songs that are observed in the wild (Payne, K., Tyack, P., & Payne, 1983), this model suffers the same problem as the Migratory Model, in that once the population has converged on a song, the population is unable to change their song as the SR's have become fixed on 100% probability transitions in each row of the SR. The complete convergence observed in these models is also slightly unrealistic, as while the songs are highly similar, there are still differences in the song. This is likely due to production errors. How production errors may accumulate in this type model is unknown, and an area of further research.

In conclusion, while small world social networks returns return the highest levels of song convergence when compared to the other social networks studied in this chapter, the mean difference between the level of convergence when compared to random social networks is not significant. It may also require interactions between singing males more complex than approaching another conspecific. This suggests that out of the three social networks examined in this chapter, random social networks best facilitate song convergence in a population of humpback whales.

# Chapter 5 – Formal Grammars and Humpback Song

## 5.1 Introduction

This chapter discusses adapting the Sequitur algorithm for analysing sequences of humpback whale song, and the development of a sampler that utilises the tools generated using Sequitur in order to recreate humpback whale song. These algorithms are used to analyse song from the years 2002 and 2003. These years were chosen as 2003 was a revolutionary year, and using Sequitur allows us to compare the differences between these two years. Following this, the Sequitur algorithm and sampler are coupled with production errors, and implemented in a vertical cultural transmission model. A number of experiments are presented, each investigating a single type of production error. For the final experiment, all production errors are coupled in a vertical cultural transmission model. The results from these experiments are then discussed, as well as their implications for cultural transmission in humpback whale song.

## 5.2 Song Collection and Representation

The data used in this research project was provided courtesy of Dr Michael Noad. This data comes from Eastern Australia songs that were recorded at Peregian Beach, Queensland, using moored, radio linked hydrophone buoys. These were equipped with High Tech HTI 96 MIN hydrophones with a built-in _40 dB gain pre-amplified, and additional external custom built preamplifier. The signals were then transmitted using AN/SSQ-47A sonobuoy transmitters and received at the base station onshore using a type 8101-sonobuoy receiver. These signals were then recorded directly on to a computer in the wav file format (16 bit, 22kHz sampling rate).

The data provided by Noad consisted of 1773 hours recordings. Given the large number of recordings, it was necessary to be selective in our choices when building our database. In order to capture the changes that may occur in the song over the course of a mating season, three recordings from the start, middle and end of the mating season were chosen. The naming convention for the units here was developed in collaboration with Luca Lamoni and Jenny Allen. It was designed in order to highlight the hierarchical structure of the song, and be consistent with the descriptive naming convention developed by Dr Ellen Garland for her PhD thesis. Although automated classification methods were investigated, it was discovered that these relied on a high signal to noise ratio. The recordings from the Noad database vary greatly in quality, with many being affected by radio interference and anthropogenic noise. Due to this fact, recordings had to be analysed by hand.

Recordings were analysed using the Adobe Audition software (Adobe, 2017). Each unit was segmented using the marker tool in the software. The unit was then compared to the dictionary of units created by Garland (Appendix 1). This process is known as *song transcription.* After identifying the unit, it would then be named using the following convention:

1. The first letter, always in capitals, corresponds to the name of the theme. Theme names are based on their position in the earliest recording of the song. IE, the first theme, heard in the first recording of that particular year, will always be called theme A for that year.
2. The second letter is always a number, and corresponds with the number of that phrase. Once again, phrases are named based on their in the earliest recording.
3. The final letter, always in lower case, corresponds to the location of that unit with that particular theme.
4. After applying the hierarchal naming, an underscore is placed at the end of the name. The descriptive unit name is then applied.

Figure 43 is an example of the above method being used to name the individual units in a phrase. The names applied to each unit come from the unit naming convention developed by Garland included in Appendix 1.

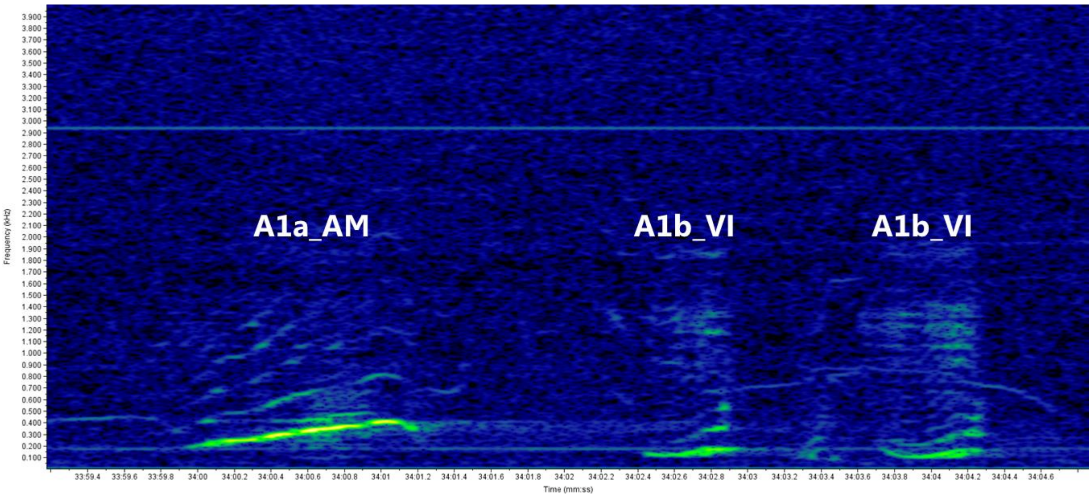


Figure 43: Spectrogram of a phrase from a song recorded in 2002. This shows the application of the naming convention developed for this thesis. If we take the unit on the far left, the capital A refers to theme A, the 1 to the phrase number and the lower case a refers to the unit position in the phrase, the capital AM means that this unit is an ascending moan, as based on Garlands stereotyped sequences found in Appendix 1.

## 5.3 Sequitur

There are several methods for inferring formal grammars. These methods however are generally associated with Linguistics. The Sequitur algorithm (Nevill-manning & Witten, 1997a) was developed to deal with sequences that have a hierarchal structure independent of any type of semantic mapping associated with language. Sequitur was designed to specifically identify recurring patterns in strings. It also accepts as input a single long sequence in order to infer hierarchy. Due to this lack of semantic mapping, Sequitur has been used to identify hierarchal structures in a passage of music (Figure 44), connecting it to the theme of applying computer music research to animal vocalisation research.

### 5.3.1 Sequitur Algorithm

The Sequitur algorithm generates a formal grammar from a single input string. In order to generate this grammar, the algorithm relies on two constraints: digram uniqueness, and rule utility.

Sequitur works by replacing repeating phrases in the given sequence with a new rule based on digrams. Here, a digram refers to two adjacent characters. When

Sequitur scans the input string, it will look for repeating paired digram's to replace with a new, non-terminal symbol. Consider the string, *S,* printed below:

$S \rightarrow aabbaa$

Since the two adjacent characters '*aa'* occurs twice in the sequence *S,* Sequitur generates the following formal grammar:

$S \rightarrow 1bb1$
$1 \rightarrow aa$

Sequitur implements a digram uniqueness constraint. This constraint ensures that no digram occurs more than once in the rule sequence it produce. This is shown in the following sequence analysis of string *S2.*

$S2 \rightarrow abaaba$

This will result in the following formal grammar

$S2 \rightarrow 1a1a$
$1 \rightarrow ab$

This occurs because when we scan the first four pairs of symbols in *S2,* we get the following digrams: *ab, ba, aa.* When sequitur reads the fifth symbol, it identifies that the digram *ab,* was generated earlier. This results in the generation of rule 1, and the process continues to the end of a string so that there are no repeated digram's.

The rule utility constraint makes ensures that all rules are used more than once on the right hand side of the grammar. This means that if a rule only occurs once, it is removed from the grammar, and is instead replaced with the symbols originally created it. This rule utility is implemented in order to decrease the number of rules in a grammar.

The method of implementing Sequitur is based on a port by Eibe Frank, and is available on the Sequitur Github[1].

**Figure removed due to Copyright Restrictions**

Figure 44: An analysis of two chorales harmonised by Bach. The sequitur algorithm identifies hierarchies within the passage of music (the shaded squares), and also identifies perfect and imperfect cadences in the piece. Taken from (Nevill-manning & Witten, 1997a)

### 5.2.2 Sequitur Sampler

The Sequitur sample adapts the formal grammar generated using the Franks Python port so that it can be used to recreate the original sequence, and generate new sequences based on the rules generated by Sequitur. This is achieved by incorporating edit operations inspired by Levenshtein Distance Analysis. These operations edit the sequence *S,* which informs the order the rules are to be carried out. This includes random insertion of new characters, random insertion of rules, deletion of characters, deletion of rules, substituting characters, substituting rules, replacing characters, and replacing rules. These operations can be viewed as being equivalent to learning and production errors. For example, a whale may insert a new unit into its sequence by accident or may miss a theme when learning a new song.

## 5.4 Analysing Humpback Song Using Sequitur

Here, Sequitur was used to analyse five different songs recorded off the coast of Eastern Australia in 2002 and 2003. Here we will examine only a single song from 2002 and 2003, as these were 'revolutionary' years, when the song of the eastern Australian population was replaced by the song (Garland et al., 2011), thus making it a very interesting case study. Other songs analysed using Sequitur from these two years are included in the supplementary material.

---

[1] *Available at https://github.com/craignm/sequitur*

Apart from the Sequitur analysis, a frequency count was taken for each unit in the song, as well as the number of rules, and the frequency of each rule in the Seq variable.

### *5.4.1 2002 Song*

Below is the formal grammar generated by Sequitur for one of the songs from 2002, recorded on the 21/09/2002. The songs shown here were generated from a single song cycle. Other songs from 2002 were analysed, and these are included in digital supplementary material. The analysis of different songs produced roughly the same number of rules. Although there were differences in the rule number assigned, many of the same rules were generated, and similar rule structure identified.

```
Original String =
'cncnoooocncnoooocncnoooocncnoooocncnoooocncnoooocncnoooocncncncncnc
nooookjjkjjkjjkbbbbbkbbbbbkfffkbbbbbbkbbbbbkfffkbbbbbbkbbbbbkfffaffa
ffafahmfagffagffagfahmaklaklakahmaklaklakahmaklaklakahmkddkddkddhmkd
dkddkddhmkdkddkddipcjjcjjcjjipcjjcjjcjjeooooc'

0 → 1 1 1 2 3 3 2 4 4 4 5 6 6 7 8 8 a 9 10 11 11 g 9 12 13 14 14 15 16 16
17 18 19 19 e 20 c

1 → 2 2                                             cncnoooocncnoooo
2 → 3 20                                            cncnoooo
3 → 21 21                                           cncn
4 → k 22                                            kjj
5 → k 23 23 b                                       kbbbbb
6 → 7 5 b                                           kbbbbbkfffkbbbbbb
7 → 5 k f 24                                        kbbbbbkfff
8 → a 24                                            aff
9 → 10 25                                           fahm
10 → f a                                            fa
11 → g 24 a                                         gffa
12 → 26 l                                           akl
13 → 12 26                                          aklak
14 → 15 13 l 26                                     ahmaklaklak
15 → a 25                                           ahm
16 → 18 17 25                                       kddkddkddhm
17 → 18 18                                          kddkdd
18 → k d d                                          kdd
19 → i p 27 27 27                                   ipcjjcjjcjj
20 → 28 28                                          oooo
21 → c n                                            cn
22 → j j                                            jj
23 → b b                                            bb
24 → f f                                            ff
25 → h m                                            hm
26 → a k                                            ak
27 → c 22                                           cjj
28 → o o                                            oo
```

Figure 45: Sequitur analysis of 2002 song.

Table 14 shows the frequency count for each unit in the song.

| Unit Name | Frequency Count | Percentage |
|---|---|---|
| c | 27 | 10% |
| n | 20 | 8% |
| o | 36 | 15% |
| k | 30 | 12% |
| j | 18 | 7% |
| b | 32 | 13% |
| f | 20 | 8% |
| a | 20 | 8% |
| h | 7 | 3% |
| m | 7 | 3% |
| g | 3 | 1% |
| l | 6 | 2% |
| d | 17 | 7% |
| i | 2 | 1% |
| p | 2 | 1% |
| e | 1 | 1% |

Table 14: Character frequency count for 2002 song.

Below is the table for the Rule frequency count from the Seq variable.

| Rule Name | Frequency Count | Percentage |
|---|---|---|
| 1 | 3 | 8% |
| 2 | 2 | 5% |
| 3 | 2 | 5% |
| 4 | 3 | 8% |
| 5 | 1 | 3% |
| 6 | 2 | 5% |
| 7 | 1 | 3% |
| 8 | 2 | 5% |
| a | 1 | 3% |
| 9 | 2 | 5% |
| 10 | 1 | 3% |
| 11 | 2 | 5% |
| g | 1 | 3% |
| 12 | 1 | 3% |
| 13 | 1 | 3% |
| 14 | 2 | 5% |
| 15 | 1 | 3% |
| 16 | 2 | 5% |
| 17 | 1 | 3% |
| 18 | 1 | 3% |
| 19 | 2 | 5% |
| e | 1 | 3% |
| 20 | 1 | 3% |
| c | 1 | 3% |

Table 15: Rule frequency count for 2002 song.

### *5.4.2 2003 Song*

Below is the formal grammar generated by Sequitur for the one of the 2003 songs, recorded on 18/09/2003.

```
original string  =
‘acacacacacacacacacacacacababababababacacdddddddacacddddddddacacdddddd
dddacacdddddddd’

0 → 1 1 2 3 4 4 4 5 6 6 d 6
1 → 2 2                                          acacacac
2 → 3 3                                          acac
3 → a c                                          ac
4 → 7 7                                          abab
5 → 2 8 8 8 d                                    acacddddddd
6 → 5 d                                          acacdddddddd
7 → a b                                          ab
8 → d d                                          dd
```

Figure 46: Sequitur analysis of 2003 song.

Below is a table showing the frequency count and what percentage each unit makes in the 2003 song.

| Unit Name | Frequency | Percentage |
|---|---|---|
| a | 25 | 30% |
| c | 19 | 23% |
| b | 6 | 7% |
| d | 32 | 39% |

Table 16: Character Frequency count for 2003 song.

Below is a table showing the frequency count and percentage quantity of each rule for the RuleSeq variable for 2003.

| Rule Name | Frequency | Percentage |
|---|---|---|
| 1 | 2 | 17% |
| 2 | 1 | 8% |
| 3 | 1 | 8% |
| 4 | 3 | 25% |
| 5 | 1 | 8% |
| 6 | 3 | 25% |
| d | 1 | 8% |

Table 17: Rule frequency count for 2003 song.

## 5.5 Experiments

These experiments investigate the effect different production errors have on the song as it passes from agent to agent in a vertical cultural transmission model. Finally, all the production errors are combined in order to investigate the effect they have on the song over time.

### 5.5.1 Control

In this control experiment, the teacher agent is trained using a song cycle from 2002. This provides the agents with the rules for recreating the song and the sequence in order for them to recreate the training song. This experiment simply shows that over multiple generations, the song does not change as it is transmitted from the teacher agent to the learner agent, even though the learner agent only receives a string from its teacher agent. Another important point to note is that the rules that each agent produces when it learns a song from its teacher agent are identical to the rules of its teacher agent. This ensures that any extensions to this model that may result in cultural evolution are a result of these extensions, and not the model itself. The only variables in this control experiment are the number of generations that the model carries out, and the initial starting song for the initial teacher agent.

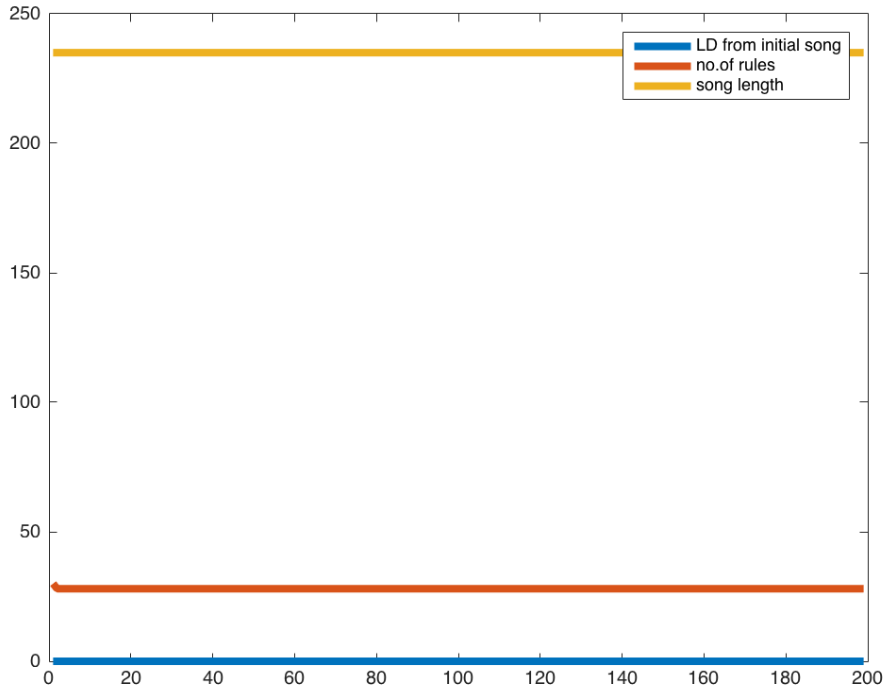


Figure 47: Results for Control Experiment.

### 5.5.2 Experiment 1 – Random Character Insertion to Sequence of Rules

In this experiment, before the teacher produces its song, there is a probability that its sequence of rules may have an extra random character added to it at a random index. A new parameter, *p,* is introduced in this model. This parameter controls the probability of these random character edits being carried. This parameter space was explored by generating an array containing 10 values ranging from 0 (no edit being carried out), to 1 (edits carried out for every teacher agent). Here we will analyse some specific cases of this parameter exploration.

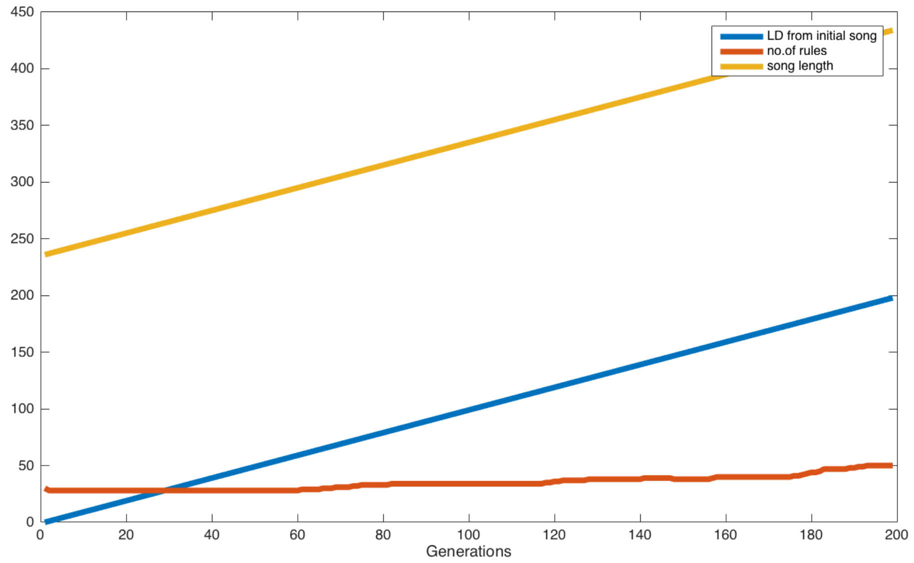


Figure 48: Results from insert character experiment

First, we examine the situation where $p = 0.9$. This means that there is a 90% probability that the song will be edited at every iteration, the most extreme case in this parameter exploration. This has a number of effects on the song. Firstly, the song length is constantly increasing, as shown in Figure 48.

Secondly, the number of rules tends to increase as these random edits are carried out. However, there are times where the number of rules decreases. It is necessary to point out that the number of rules in sequitur is determined by repeating patterns in the string, and if they can be replaced by a simpler rule.

### 5.5.3 Experiment 2 - Random Character Deletion From Sequence

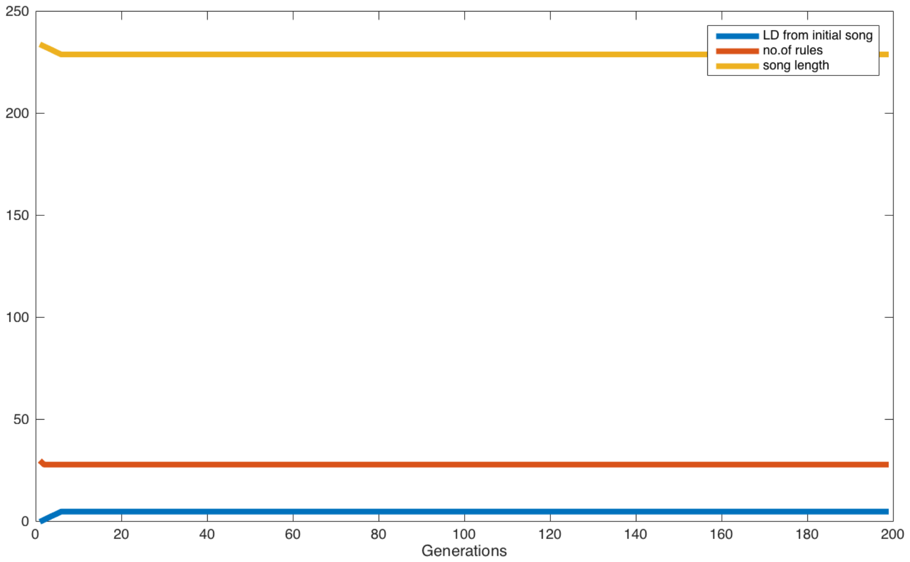


Figure 49: Results for random character deletion in sequence.

It is important to note here that all these value plateau due to the fact that there are not other characters in the sequence. If there were more characters, the trend of linear decrease for song length, and a linear increase for Levenshtein distance. The number of rules is less clear, bus is likely affected if a character deletion resulted in the merging of two rules, or the creation of a new rules, of if a character substitution in the rule would result in a simpler representation of the song according to Sequitur.

### 5.5.4 Experiment 3 – Random Character Substitution to Sequence

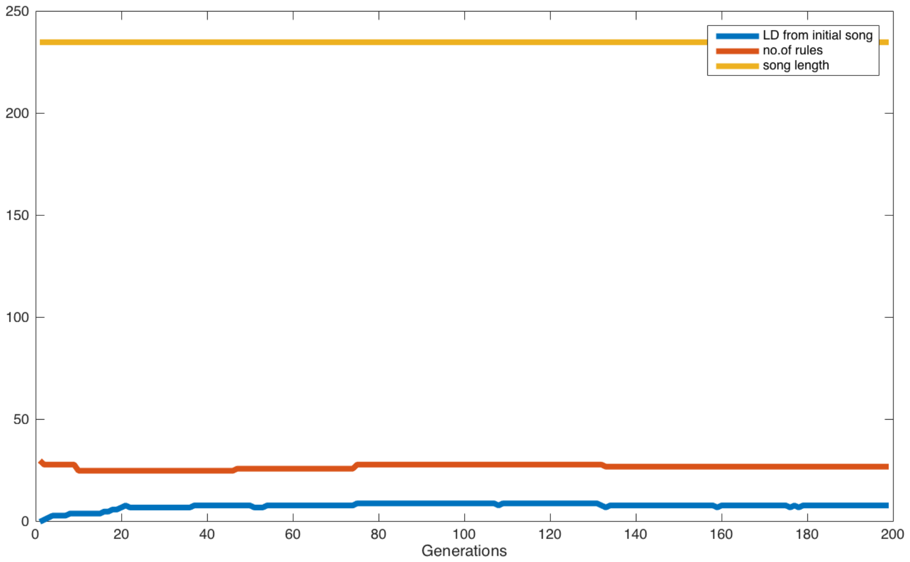


Figure 50: Results from random character substitution experiment.

From the character substitution results, we can see that the song distance remains the same. There is a slight fluctuation in the number of rules and Levenshtein distance, but these are not extreme.

### 5.5.5 Experiment 4 – Random Rule Insertion to Sequence

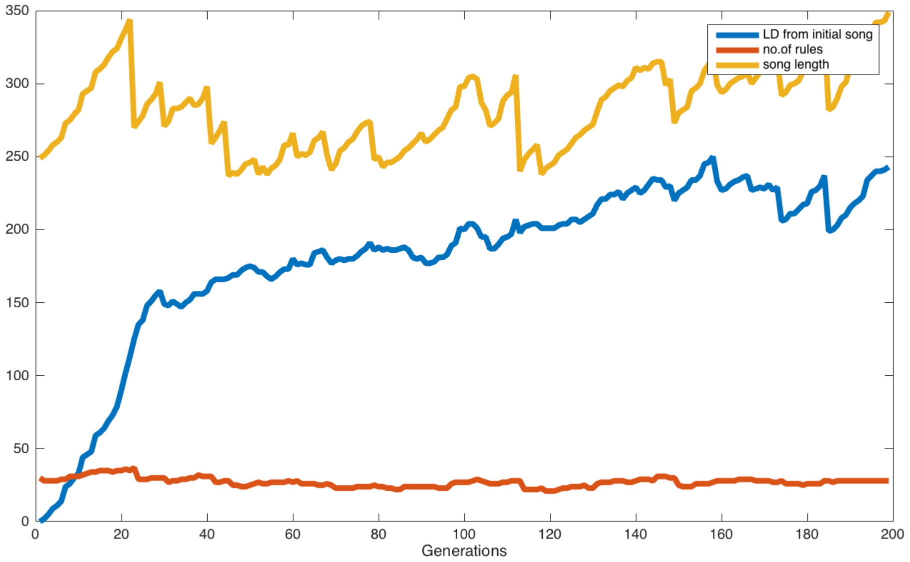


Figure 51: Random rule insertion results.

Random rule insertion shows the most dramatic fluctuation in the results. However, there appeared to be a decrease in the song length. This appeared to be caused by a glitch that occurred when songs reached a very long length (over 300 characters). The specific cause of the glitch has not been identified at the time of writing. The number of rules remained fairly consistent. Levenshtein distance generally tended to increase, with some downwards fluctuations occurring.

### 5.5.6 Experiment 5 – Random Rule Deletion to Sequence

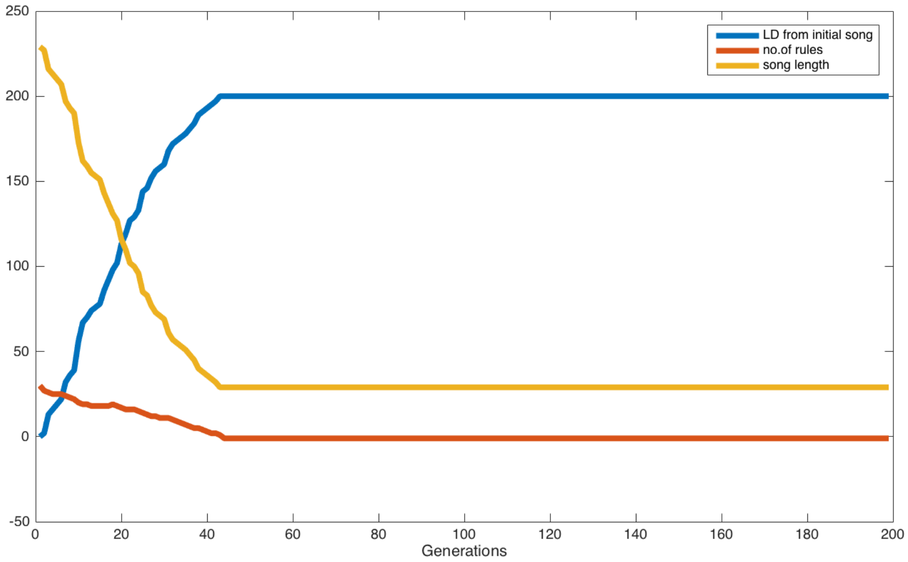


Figure 52: Results from random rule deletion experiments.

In the random rule deletion experiments, Levenshtein distance increased gradually until all rules had been deleted. The number of rules also decreased gradually until there were none left. Song length also decreased but did not reach zero as characters still remained in the agents Seq variable.

### 5.5.7 Experiment 6 – Random Rule Substitution

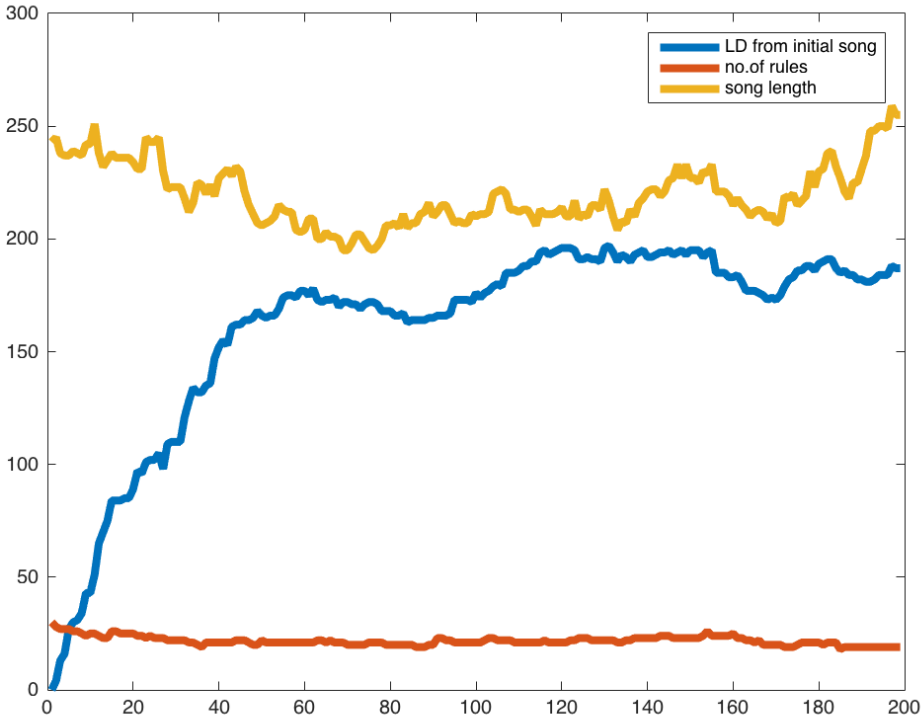


Figure 53: Results from random rule substitution experiments.

Random rule substitution resulted in Levenshtein distance increasing over the course of the model run. The song length could either increase or decrease, as the rules were of various lengths. The number of rules did not fluctuate dramatically.

### 5.5.8 Experiment 7 – Random Edits

In this experiment, all of the editing operations that were discussed in the previous experiments are carried out randomly. Once again, the only variable is *p,* that controls the probability of one of these edits being carried out. If the coin-flip determines that an edit should be carried out, a vector containing a uniform distribution is sampled from in order to determine which edit should be carried out.

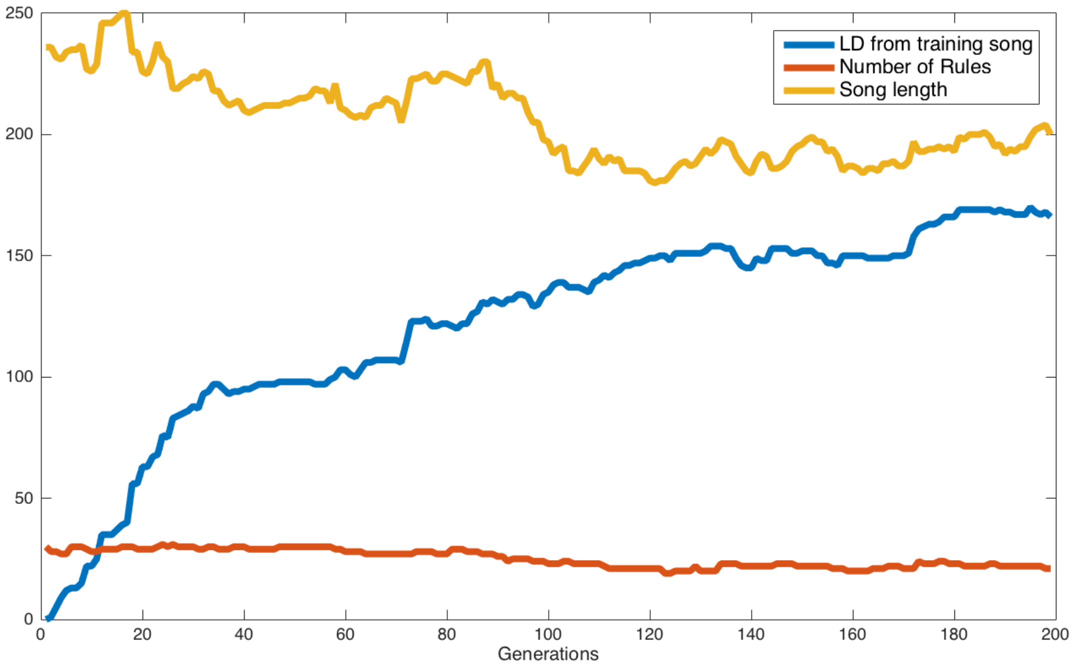


Figure 54: Results for song evolution when the probability of each edit operation is assigned uniform distribution. Here p = 1 and random edits are carried out every generation.

In order to determine the influence of the all the operations coupled together, the model was re-run with a *p* of 1, for 4'000 generations. These results are shown in Figure 55.

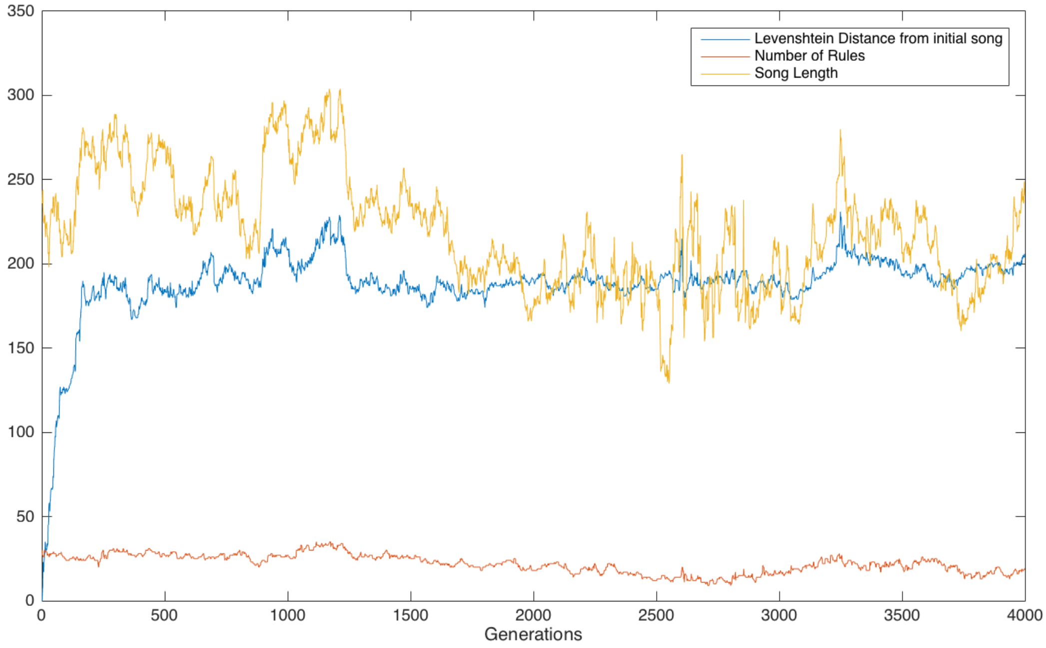


Figure 55: Long run with random edits.

One of the most interesting results from this experiment is the long run of experiment 7. The model was trained with string used for all these models (taken from **2002 song**). The string produced after 4’000 generations is printed below:

*Pvjjjjaaooaaooqhnjjjjaaooaaooaaoojjjjaaooaaooaaooaaooibrjjjjjaaoorjjjjjaaoogjjjrjjjcaaoojjjjjjfhqhngjjjjjjaaooaaooaaooaaoobaaooaaoojjaaoogjjjzfkjjjjjjcfjjcdzdqhngalkjjaaooaaoojjjjaaoogjjjnaaoopaaoogjjjjjaaooaaoojjjqhnglrjjjnyhncaaooaoohnqhnd*

From an initial qualitative examination of the string, we can see that its contents are very different to the sequence the model was initially trained with, but retains a structure. In order to verify the difference in the songs units, the frequency of each character was calculated for the training string, and final string at 4’000 iterations, as shown in Tables 11 and 12 respectively.

| **Value** | **Count** | **Percentage** |
|---|---|---|
| n | 20 | 8% |
| c | 26 | 10% |
| o | 37 | 15% |
| k | 30 | 12% |
| j | 18 | 7% |
| a | 52 | 21% |
| f | 20 | 8% |
| h | 7 | 3% |
| m | 7 | 3% |
| g | 3 | 1% |
| l | 6 | 2% |
| d | 18 | 7% |
| i | 2 | 1% |
| b | 2 | 1% |
| e | 1 | 0.5% |

Table 18: Unit frequency count for training song.

| Value | Count | Percentage |
|---|---|---|
| P | 2 | 1% |
| V | 1 | 05% |
| J | 73 | 30% |
| A | 56 | 23% |
| O | 56 | 23% |
| Q | 5 | 2% |
| H | 8 | 3% |
| H | 9 | 3% |
| I | 1 | 0.4% |
| B | 2 | 1% |
| R | 4 | 2% |
| G | 7 | 3% |
| C | 4 | 2% |
| F | 3 | 1% |
| Z | 2 | 1% |
| K | 2 | 1% |
| D | 3 | 1% |
| L | 2 | 1% |
| Y | 1 | 0.4% |

Table 19: Unit frequency count for final song in model long run.

| Rule Number | Rule | Rule Output |
|---|---|---|
| R1 | R19 R11 | Jjjjaaoo |
| R2 | q R17 | Qhn |
| R3 | R1 R4 | Jjjjaaooaaoo |
| R4 | a R18 | Aaoo |
| R5 | r R19 R16 | Rjjjjjjjjjaaooaaoo |
| R6 | g R10 | Gjjj |
| R7 | r R10 | Rjjj |
| R8 | c R4 | caaoo |
| R9 | R10 R10 | Jjjjjj |
| R10 | R13 j | Jjj |
| R11 | R4 R4 | Aaooaaoo |
| R12 | R13 R4 | Jjaaoo |
| R13 | j j | Jj |
| R14 | R2 g | Ohng |
| R15 | R12 R4 R10 | Jjaaooaaoojjj |
| R16 | j R4 | Jaaoo |
| R17 | h n | Hn |
| R18 | a o o | Aoo |
| R19 | R13 R13 | jjjj |

Table 20: Formal grammar for song at the end of a long run experiment.

**P, v, 1, 2, 3, 3, 4, I, b, 5, 5, 6, 7, 8, 9, f, h, 2, 6, 10, 11, 11, b, 11, 12, 6, z, f, k, 9, c, f, 13, c, d, z, d, 14, a, l, k, 15, 16, 6, n, 4, p, 4, 6, 15, 14, l, 7, n, y, 17, 8, 18, 17, 2, d – Long run experiment rule sequence.**

Table 20 shows a grammar analysis of the song using Sequitur. As we can see there appears to be a hierarchal structure similar to the grammar shown in **Error! Reference source not found.**.

One of the most interesting results is that the Levenshtein Distance in the long run appeared to plateau at around 500 generations. This is likely due to a maximum distance being reached from the initial training string. Since the song length plateaued, substitutions would be the only operation required in order to return to the original song, thus resulting in a fairly stable Levenshtein Distance score.

## 5.6 Discussion

This exploration of real world songs has only scratched the surface of what hierarchical analysis using Sequitur can achieve. This analysis focussed only on a comparative analysis of the structure of the song. A comparative analysis of the contents of the two songs would surely reveal a greater degree of insight into the differences of these two songs. For example, rather than assigning unique units when encoding a song, an alphabet based on the naming convention developed by Allen could be implemented (Allen, Garland, Murray, Noad, & Dunlop, 2017) in order to assign unique values to each unit. This would enable a Levenshtein distance comparison to be carried out between the rules of the two separate breeding seasons.

The seven iterated learning experiments reveal how these different operations can influence the song of the agents. It is important to note that in the generation of these songs, the agents never directly edit the rules themselves. They can repeat rules, substitute them with other rules, but they can never carry out a direct operation on their internal rules. Despite having access to only these operations, we see that the internal representation of the song (the rules) can still change.

Some of the operations presented here have a very predictable effect on the song. In the case of character insertions and deletions, the Levenshtein distance increases and decreases linearly respectively. How these may affect the internal rules of an agent depends on where the characters are inserted into the sequence, and if it results in Sequitur deleting or adding a rule. Examples of how these edits may affect the song are demonstrated in Figure 56 through to Figure 61. The effect each of these operations has on the song analysis is summarised in Table 21.

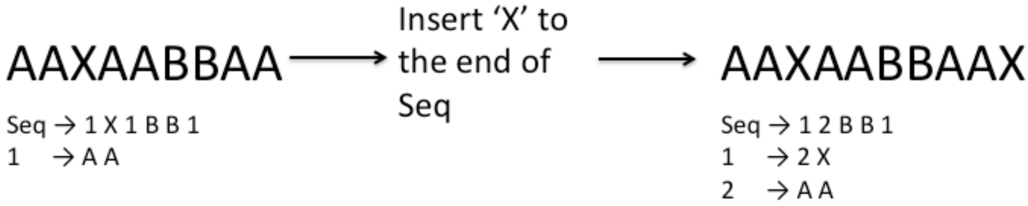


Figure 56: This shows how a character insertion can result in the creation of a new rule. The sequence AAX appears twice in the new sequence on the right, and can be re-created using previous rules, so the Sequitur algorithm adds another rule.

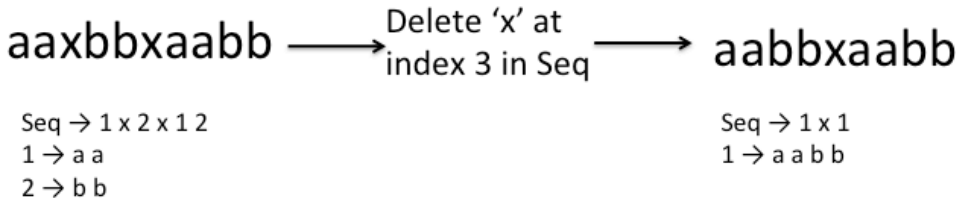


Figure 57: The effect deleting a character can have on a song.

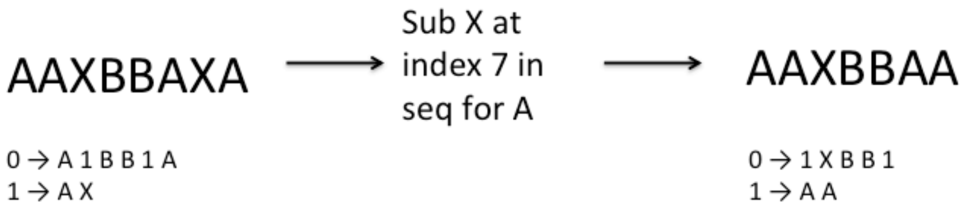


Figure 58: The effect of substituting a character in a song.

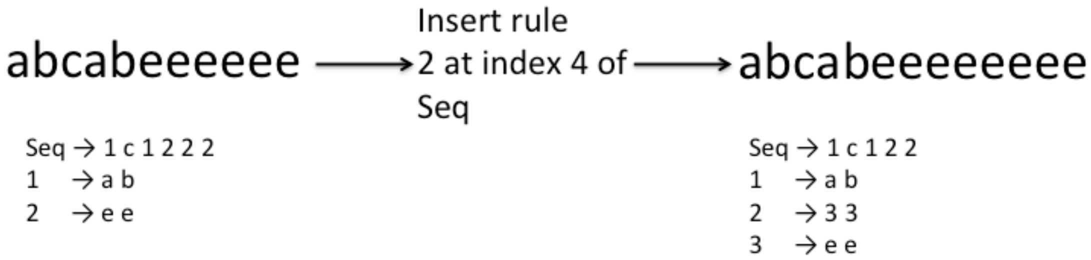


Figure 59: The effect character inserting a rule can have on a song.

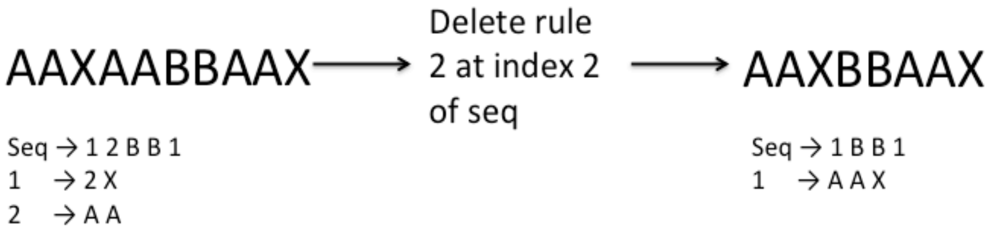


Figure 60: The effect of deleting a rule in a song.

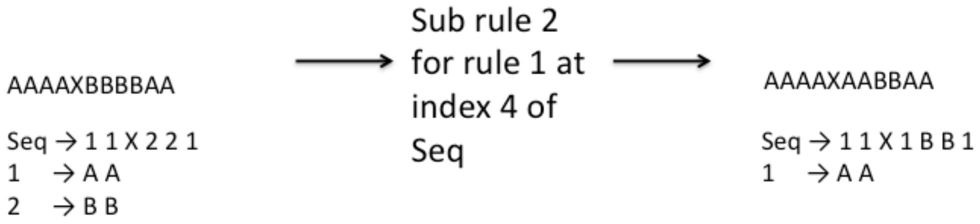


Figure 61: The effect of substituting a rule in a song.

| | Levenshtein Distance | Song Length | Number of Rules |
|---|---|---|---|
| Character insertion | **+L** | **+L** | **+** |
| Character Deletion | **+L** | **-L** | **-** |
| Character Substitution | **+L** | **x** | **+-** |
| Rule Insertion | **+** | **+** | **+** |
| Rule Deletion | **+** | **-** | **-** |
| Rule Substitution | **+** | **+-** | +- |
| Random | **+-** | +- | +- |

Table 21: This table shows the effect each operation has on the analysis methods. A + character implies that it generally increases the measurement, a – implies that it generally decreases the measurement. A combination of the two characters means it can both increase and decrease the measurement. An X means that it has no effect on that measurement. A capital L indicates that it influences the parameter linearly.

These processes were chosen as they were directly influenced by the concept of using Levenshtein distance in a synthesis form. These operations could be viewed as analogies to production and learning errors. For example, character insertion could be seen as a production error by accidentally adding an extra unit, and character deletion could roughly be equated to missing a unit when learning a song. These are processes that likely occur in wild humpbacks. They key point to take away from this is that when the ability to identify hierarchy is built into agents, minor edits carried out from agent to agent with these production and learning errors can result dramatically different songs given enough time in a form of cumulative cultural transmission. This occurs without any need to edit the agents' internal rules. New rules arise because of how the agents perceive the hierarchy in the song.

The syntactical abilities of animals are a matter of debate. An excellent discussion on this matter is given in (Cate, Okanoya, Cate, & Okanoya, 2012). However, in Cate et al. study, the primary discussion focussed on primates and birds, and their ability to distinguish between two different forms of strings with hierarchies relating to grammars of the form AB and AABB. The study focussed on birds and primates, and cannot be directly compared to the learning abilities of humpback whales. Furthermore, while playback experiments have been carried out on humpbacks (Tyack, 1983), the main response from the animal was one of aggression, with the whale responding by charging at the boat, making these types of studies impossible (or ill advised at the very least). This means that the syntactical abilities of humpback whales in relation to what type of grammars they are capable of reproducing are unclear at best. Another point made in the Cate et al. paper that may be challenged is that no grammar more complex than a probabilistic finite state grammar similar is required in order to recreate the vocal sequences of these animals. The most commonly used probabilistic finite state grammar used is the $n^{th}$-order Markov Chain. Studies have shown that there are many other models other than Markov chains that can recreate animal vocal sequences (Kershenbaum et al., 2014). This is further reinforced by studies that showed that Markov chains required an unfeasible amount of data in order to approach the capabilities of modelling humpback whale song (Suzuki et al., 2006), and as we have seen in Chapter 3, these models converge on short sequences and are not appropriate for modelling the song at the unit level.

The formal grammar generated using Sequitur is used as a placeholder for ability to recognise hierarchy in a sequence of symbols. It is possible that a simpler system could be developed. For example, after analysing a song, the rules could be trimmed, in order to remove rules that produce long strings such as R5 in Table 20. While this seems like a simple task in theory, implementing it proves to be quite challenging due to the nature of nesting in the Sequitur algorithm, and the deletion of rules requires reconfiguring not only the set of rules, but also the sequence itself. One possible avenue is to remove any rules longer than four characters. Then use the remaining rules to recreate the training sequence. However, current time constraints do not allow for this avenue to be explored in this thesis. Another alternative is to generate formal grammars based on the sequences analysed by human transcribers.

# Chapter 6 – Adapting the Models for Music

## 6.1 Introduction

Since the methods used to develop the models presented throughout the thesis orginate from analysis and synthesis methods for music, it is natural to want to revisit this, and investigate how these models could be used in music. This allows us to investigate the potential of the model to investigate how it could be applied to composing new music, but also allows us to investigate how these methods could be used in investigating cultural transmission in music. Cultural transmission plays an important role in how music changes over time. For example, in traditional Irish music, there is a large repertoire of songs that are shared amongst different performers. However, it is not uncommon for a performer to take a song and change to suit their needs, or to use existing material as the basis for new compositions.

For example, the famous song *On Raglan Road*, owes its origins to the Irish language song *Fáinne Geal an Lae*. The melody for *Fáinne Geal an Lae* became the melody for *On Raglan Road*. *Fáinne Geal an Lae* also experienced another form of change, when the melody was adapted into a march, and is one of the first songs many students of traditional Irish music will learn. While almost every student of Irish music will learn the melody of *Fáinne Geal an Lae*, many are unlikely to be aware of its origins, and will simply think of it as *On Raglan Road.* This is a dramatic example, but this change in music is a form of cultural transmission, and is an example of how a song may change over time. There was a large tradition of itinerant musicians in Ireland, such as the famous travelling harper Turlough O'Carolan (Rimmer, 1987), or the travelling pipers that existed as recently as the early 20th century (Tuohy & Ó hAodha, 2008). In order to collect the songs of these travelling musicians, Edward Bunting who was hired to transcribe Irish tunes for the Belfast Harp Festival. Buntings work displays an interesting aspect of cultural transmission, as he did not receive royalties for his work due to pirated editions, and the setting of his transcribed tunes to the popular songs of Thomas Moore (Cathcart, 2014).

There are claims that Buntings original transpositions actually contained errors as Bunting, a classically trained musician, did not account for the modal nature of Irish music, and “corrected” the melodies to put them in line with classical music theory. While these claims do not appear to be verified, it is not an unreasonable assumption, given the diatonic nature of the Irish harp. These “corrections” could be viewed as copying errors, not unlike those observed in birds (Fehér et al., 2009; Lachlan et al., 2016). While Moore’s setting of these songs can be viewed as an innovation, this example of transmission from O’Carolan, to Bunting, to Moore presents key opportunities for copying errors, and the songs likely changed due to these errors. Errors are not necessarily a bad thing, and in some instance may lead to genuine innovation that will give the individual an evolutionary advantage (L Rendell et al., 2010).

This gradual change may be something a composer may wish to recreate and suggests that models developed for studying vocal learning in animals would allow a composer to investigate how cultural transmission may transform melodies they have already written. For this reason, the models developed for this thesis may have use as a tool for composition. Similarly, they could even be used in some anthropological capacity to investigate how songs learned in an aural tradition may change over time. This move towards musical systems may also have some scientific value. One popular theory in language evolution suggests that language evolved through a musical proto-language (Fitch, 2010; Tallerman, 2013). This suggests that adapting models of vocal learning in animals for composition may reveal insights into the key differences between music and language.

In order to investigate the potential of these vocal-learning models for animals in composing music, this chapter attempts to connect the models back to their origins, and examine their potential to as a tool for algorithmic composition. The chapter begins by introducing an adaptation of the migratory model from Chapter 3 coupled with OSC, allowing it to communicate with Ableton Live software via MAX4Live. A number of experiments regarding the novelty algorithm from Chapter 3 are then presented, highlighting why it is inappropriate as a tool for composition. A new algorithm is then presented, called the Dynamic Novelty Algorithm. The model is used in a simple composition experiment, and its results and the potential of the model as a tool for composition is then discussed.

## 6.2 Adapting The Migratory Model

As it stands, the migratory model does not capture the full complexity observed in humpback whale song. While agents in our model do converge on a shared song in certain situations, it does not present any change once every agent in the population has learned the song. Furthermore, first order transition matrices are not capable of capturing the hierarchical structure of humpback whale song. Despite this, the model can be adapted for use in algorithmic composition. In the following section, the technical aspect of adapting the migratory model is explained.

### 6.2.1 Technical Considerations

Open Sound Control (OSC) (Freed, Adrian, and Wright, 2018) is a protocol used to transmit data between different audio software programs. This package allows for the quick adaptation of the migratory model to be used as a tool for composition. Using OSC, it can send data to and from the migratory model in order to generate new musical sequences in real time. This is achieved using the Max4Live API in Ableton Live(Ableton, 2017a, 2017b), so that the composer may introduce new songs sequences to the population, and play them back in order to generate new musical variations, based on this input and other parameter settings. This is illustrated in Figure 62.

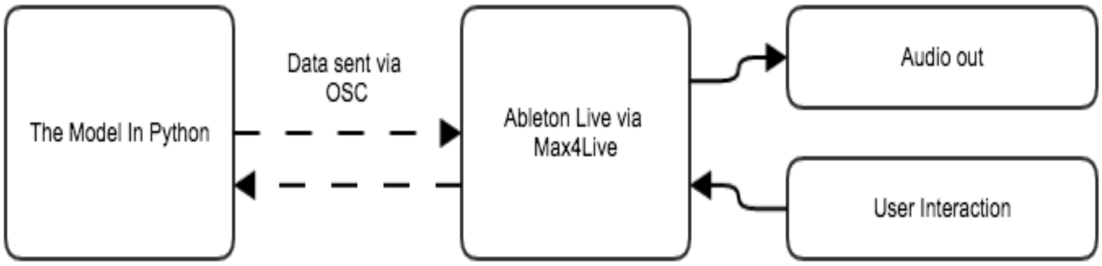


Figure 62: Signal flow from Ableton to the model.

In order to interact with the model, the composer uses the ComposerIn device (Figure 64) with a MIDI keyboard to create a sequence of notes and rhythms to be learned by a selected agent in the model. These notes are appended to a list in Max/MSP, where they are then formatted so that they can be used as an input to the model. The sequence is then sent via OSC to a selected agent, who estimates a first order transition matrix so that it may create variations on this theme. This interaction flow is demonstrated in Figure 63

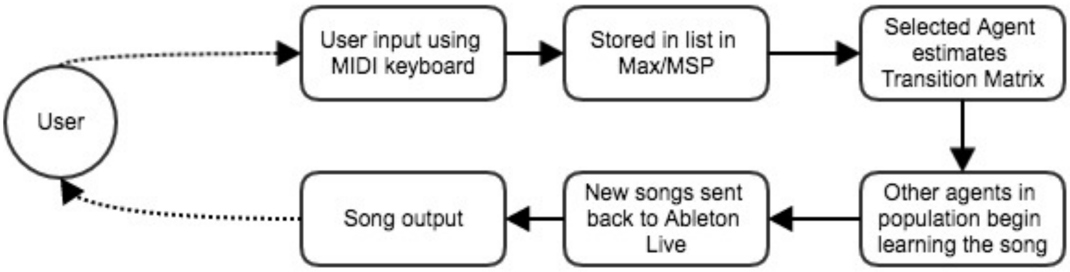
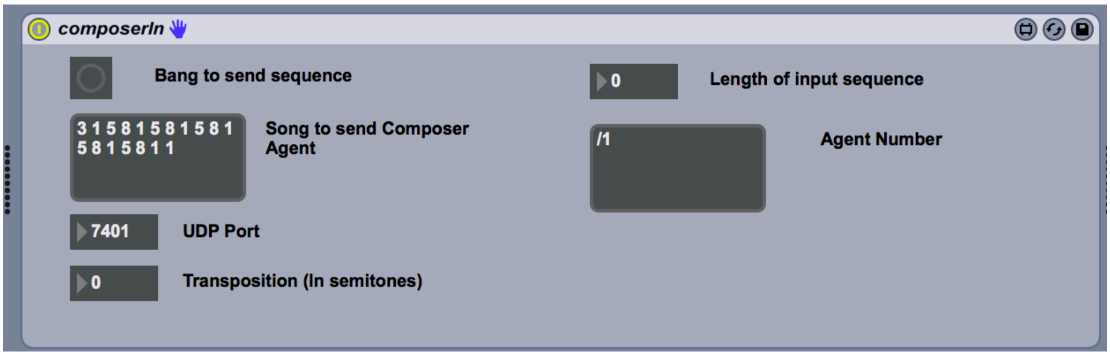


Figure 63: This shows how a composer interacts with the model.

At each iteration of the model, the song produced by each agent is sent back to Ableton Live using the modelOut device (Figure 65), where they are transposed in order for them to be formatted into MIDI notes. These are then stored in a message box and sequenced using a metro object. This allows the MIDI notes to be sent to any Live or Max4Live device that the composer wishes to use.



Bang to send sequence

Song to send Composer Agent: 3 1 5 8 1 5 8 1 5 8 1 5 8 1 5 8 1 1

UDP Port: 7401

Transposition (In semitones): 0

Length of input sequence: 0

Agent Number: /1

Figure 64: The ComposerIn Max4Live device.


ModelOut

/1 OSC Address for the agent you want

Click this bang after updating OSC Adress

250 Time between units

The song output by the agent

1 2 3

Stop the agent singing

200 Unit duration in milliseconds

0 Range (in semitones from c1)

7400 UDP Port for receiving messages.

Figure 65: The ModelOut device.

### 6.3 Developing a New Novelty algorithm

Since the model relies on the distance between agents to influence the transmission of the song, it is possible for the song input by the composer to be overpowered by the other songs in the population of agents if they are clustered closely together. This removes the power of the composer over the system, and requires some type of adaptation of the system. The solution to this was to try and incorporate a novelty algorithm, as discussed in Chapter 3. As mentioned in Chapter 3, the novelty algorithm did not work as expected, and agents tended to converge on highly random SR's. Here, a number of experiments focussing on the novelty algorithm are presented, and highlights why the algorithm is also inappropriate for generating musical ideas. Following these experiments, a new novelty algorithm is developed that allows the **composer** to have more interaction with the system.

#### 6.3.1 Experiment 1: Novelty and Levenshtein Distance(LD) Correlation (No learning)

In order to investigate how the novelty algorithm behaves when a listener agent has a sharply defined transition matrix, the following experiment was created.

1. Create two agents. A singer and an analyser.
2. At every iteration, both agents sing.
3. The analyser agent carries out a novelty calculation with its transition matrix. This value is appended to an array called nA.
4. A Levenshtein distance comparison is carried out between the singer song, and the analyser song. This value is appended to an array called lA.
5. Repeat steps 2 to 4 for N iterations.

After carrying out this experiment, the arrays nA and lA are normalised by the sum of their respective arrays.

The resulting arrays are then plotted. A number of transition matrices for the analyser and singer agent were tested. First, an analyser agent was given the following transition matrix:

$$aMat = \begin{vmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

The singer agent was initialised with a randomly initialised matrix. Figure 66 shows the result of this experiment.

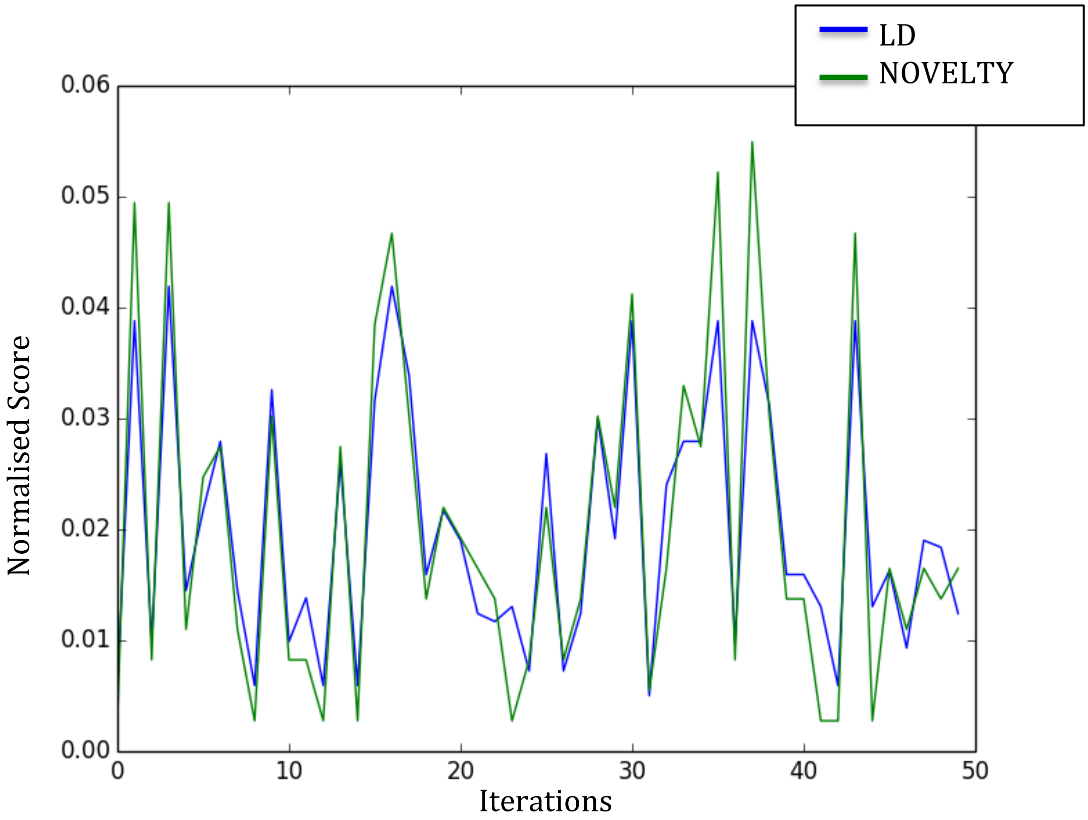


Figure 66: This shows the general behaviour of the novelty algorithm when the listener agent has a sharply defined transition matrix and the songs that are coming in are randomly generated.

From Figure 66, it is easy to see that novelty and Levenshtein distance generally correlate with one another when an agent has a sharply defined transition matrix, and the songs that it is analysing are random in nature. The experiment was repeated again, but the singers' transition matrix was changed to the following:

$$sMat = \begin{vmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

This resulted in a no variation in Levenshtein distance over the run of the model, returning a consistent value of 0.2 for both Levenshtein distance and novelty.

Finally, the experiments were carried out one more time, but with the agents initialised with random transition matrices. This produced varying results, with Levenshtein distance and novelty correlating in certain instances as seen in Figure 67.

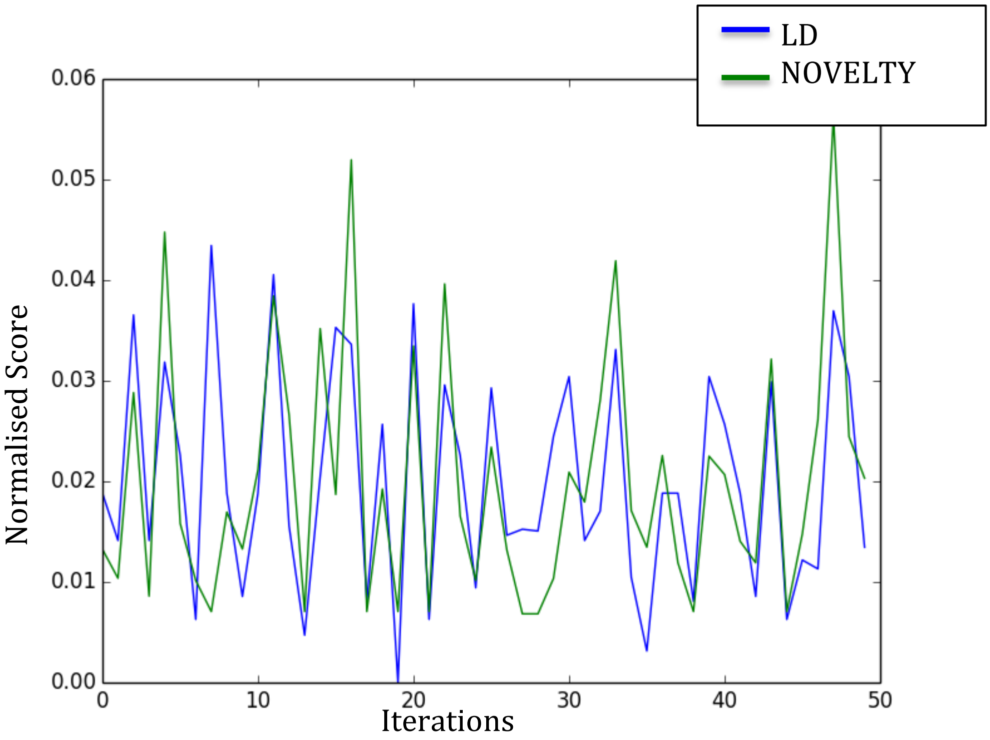


Figure 67: Correlation with random matrices.

At other times, it produced more chaotic behaviour, as seen in Figure 68.

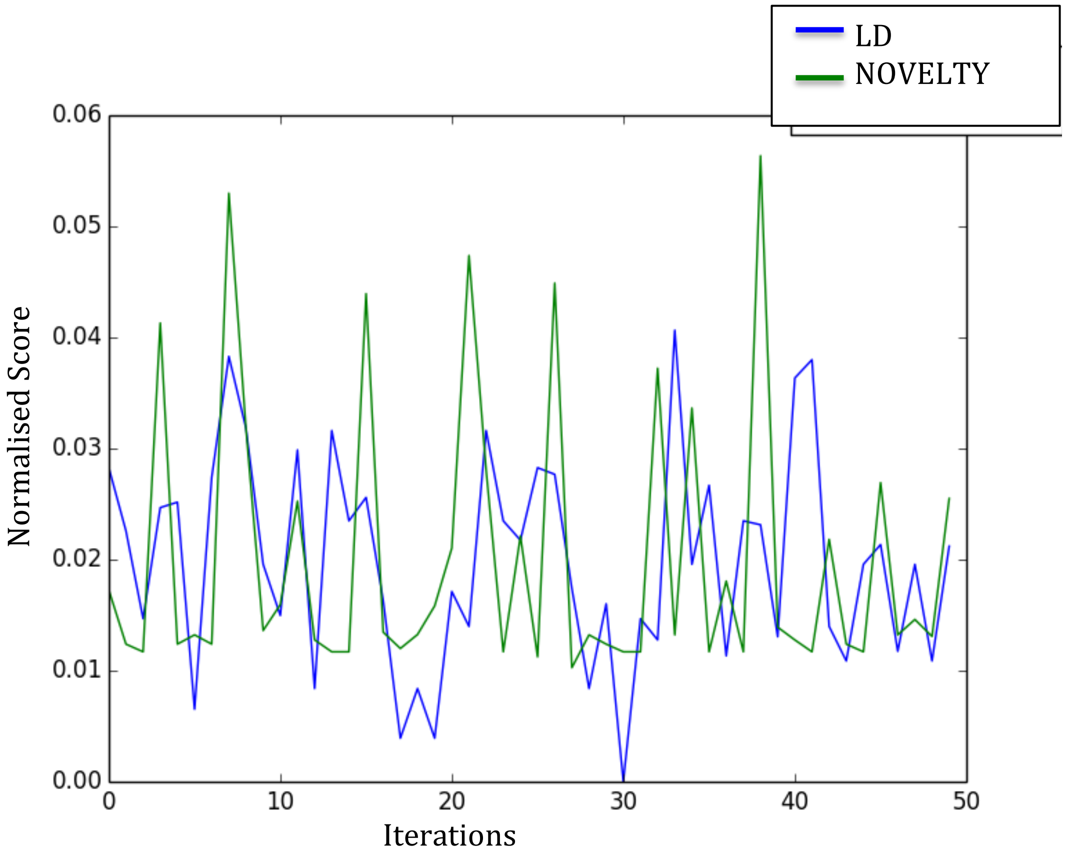


Figure 68: Another experiment with randomly initialised matrices for both agents. This displays slightly more chaotic behaviour.

Despite this, it generally seems that the majority of times Levenshtein distance and novelty correlated, even when matrices are randomised. The extent to which this happens needs to be investigated further.

### 6.3.2 Experiment 2: Novelty and LD Correlation (With Learning)

This experiment repeats experiment 1, but over multiple cycles. After steps 1 to 5 in experiment 1 are completed, the, rather than normalising the values, we sum *nA* array and append it to an array. We also sum the values in the *lA* array and append these values to a separate array. After this has been done, we update the analyser agents' transition matrix to approach the singer agents' transition matrix, using the same algorithm for learning in our model. A new learning rate variable, *lr,* is introduced to control how much of the singers' transition matrix the analyser agent will learn. This is show in Equation 8.

$$Ta = Ta * (1 - lr) + \ Sa * lr \tag{8}$$

Where *Ta* is the analyser agents' transition matrix, *Sa* is the singer agents' transition matrix, and *lr* is the learning rate. The results for this experiment are plotted in Figure 69.

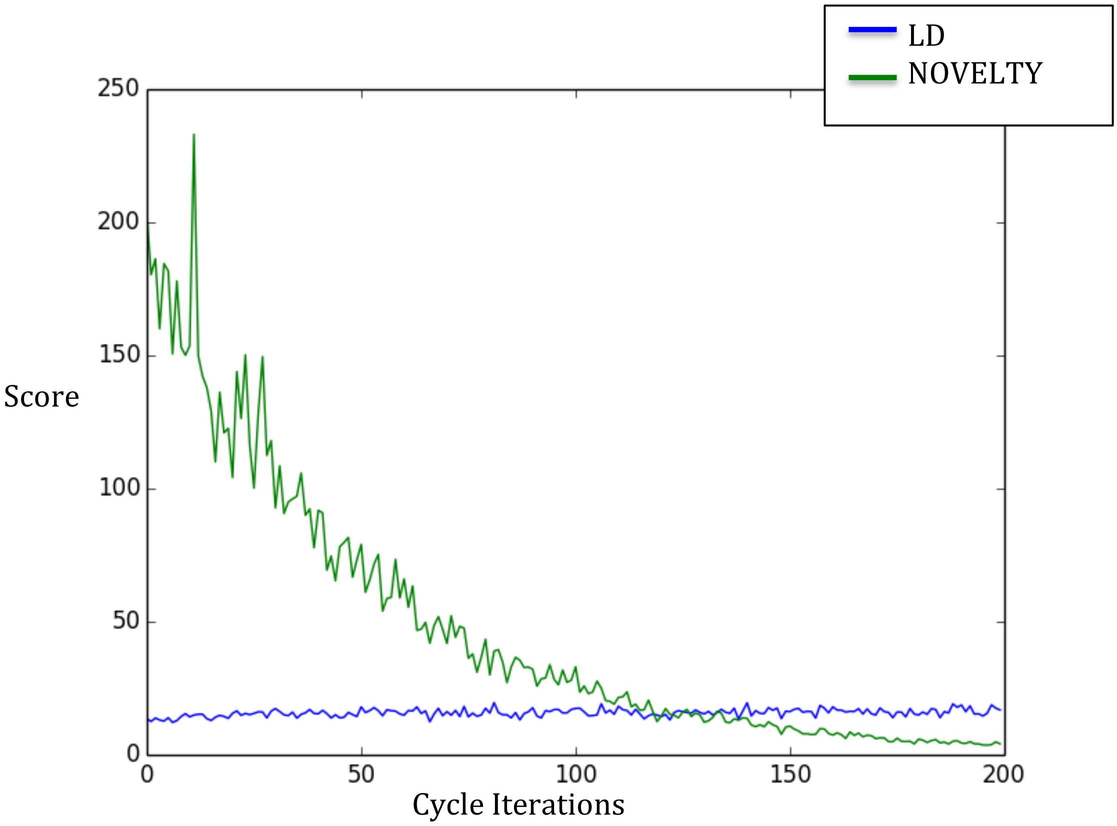


Figure 69: This shows novelty decreasing as the analyser agents' transition matrix approaches that of the singers' transition matrix.

As we can see from Figure 69, as the analyser agent approached the singer agents' transition matrix, novelty decreased. Because the songs were effectively random Levenshtein distance remained fairly constant throughout. Figure 70 to Figure 74 show the analyser agent’s transition matrices at different cycle iterations.

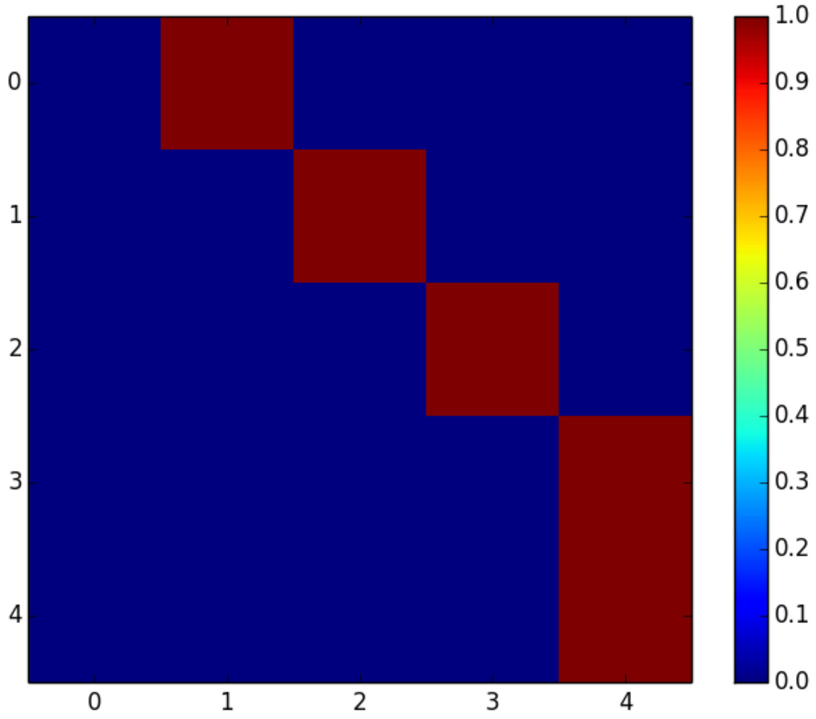


Figure 70: Analyser Agents matrix at 0

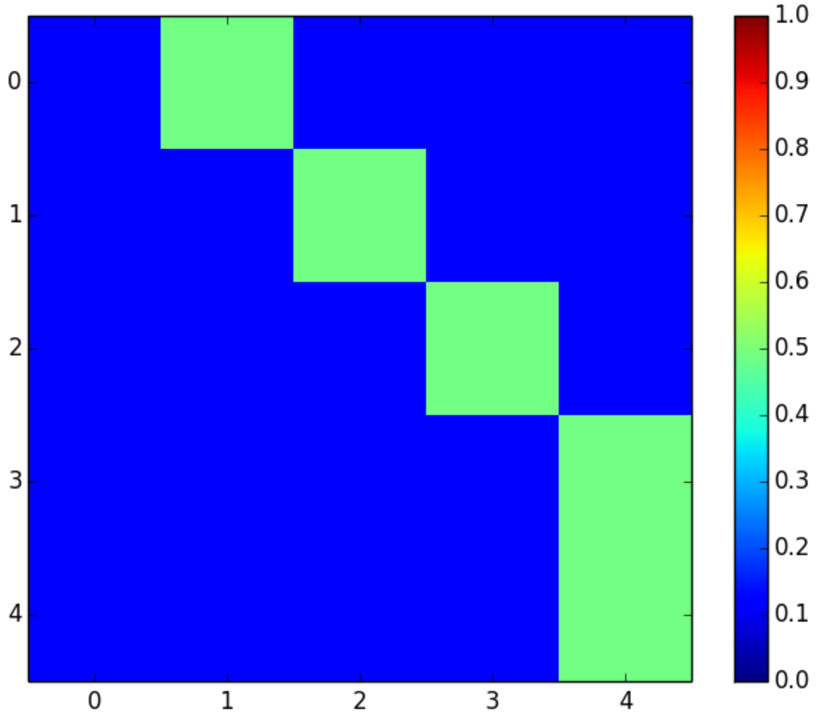


Figure 71: Analyser matrix at 50 iterations

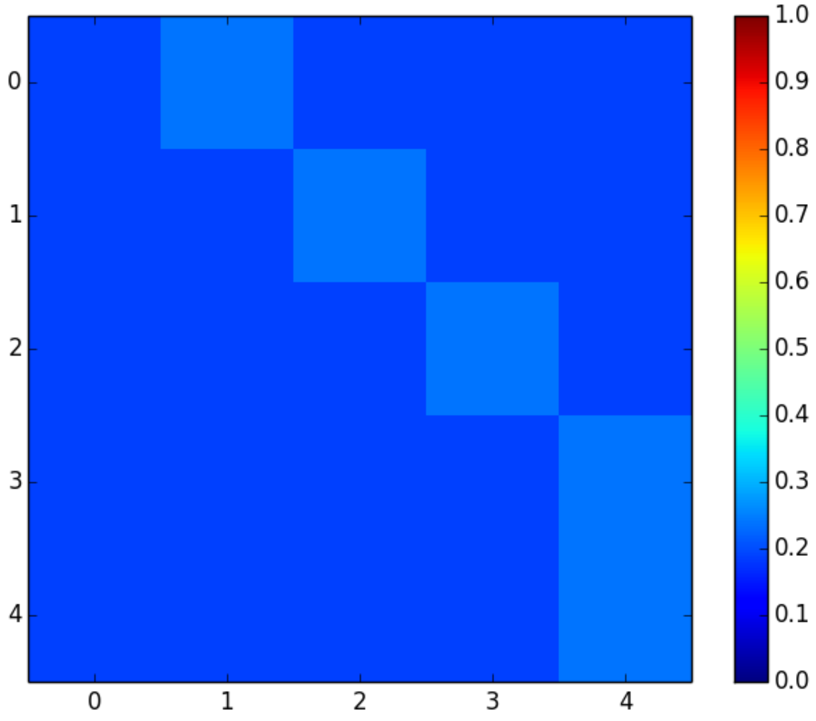


Figure 72: Analyser matrix at 150

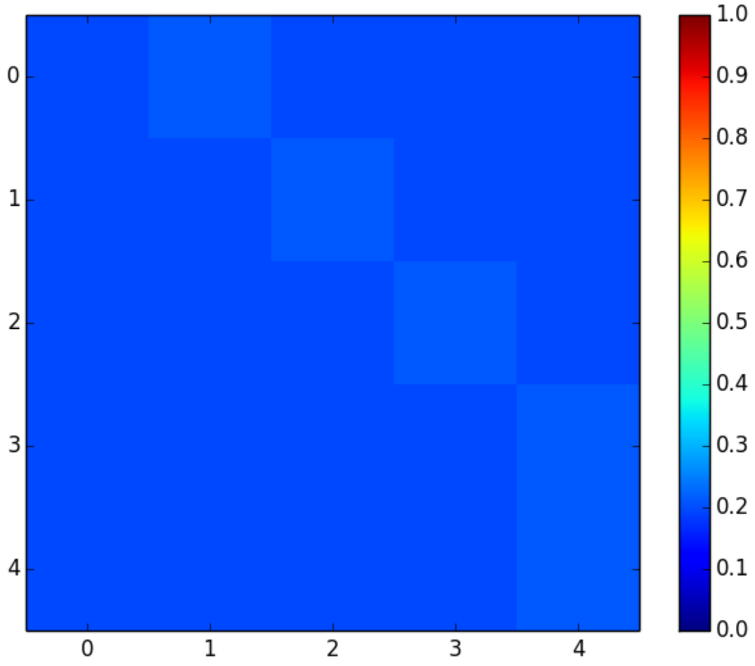


Figure 73: Analyser matrix at 199

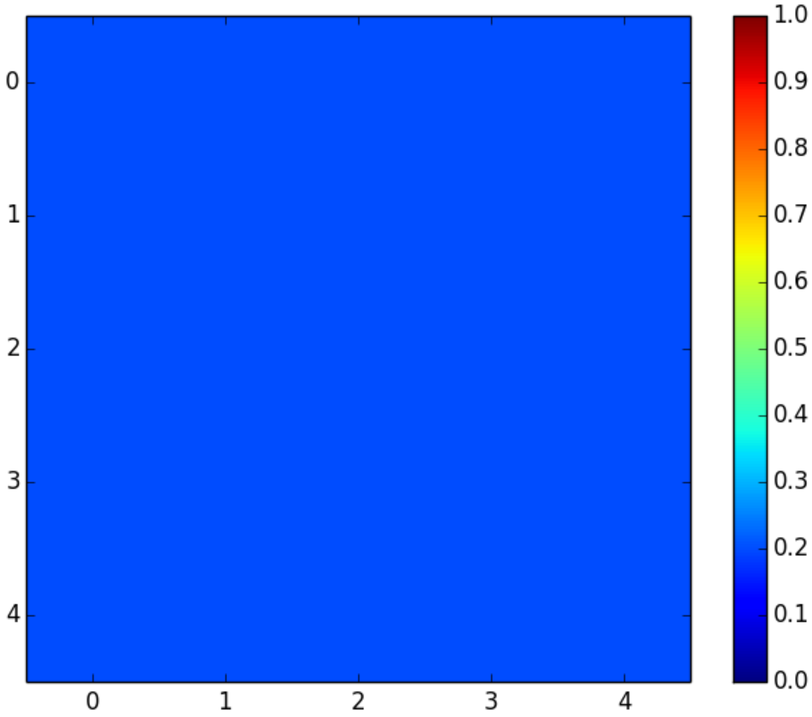


Figure 74: Singer matrix

## 6.4 Dynamic Novelty Weighting

The experiments in the previous section showed that while novelty and Levenshtein distance do correlate under certain circumstances, the agents eventually converged on uniform SR's if novelty is used as a weighting in their learning algorithm. Once all agents converge on uniform SR's, each transition is as likely as another, and the novelty of any incoming sequence is zero, and effectively stops learning. This type of stagnant, non-dynamic system does not allow the composer to influence the songs of agents. It would also be possible to create this type of system using a single agent with a uniform SR, making this type of complex model unnecessary, and highlights why the model with only novelty produces uninteresting musical ideas.

The original novelty algorithm effectively stops learning after a certain number of iterations, and the composer can no longer influence the system. To deal with this problem a new novelty algorithm was investigated. The reason for developing a new novelty algorithm is simply because it did not require dramatically changing how the agents learned, and kept the system as close as possible to the original model. This algorithm is called the Dynamic Novelty Weighting Algorithm, and it calculates novelty in relation to the songs of all agents in the population. This is achieved by calculating the novelty value for every agent in the population, and then dividing their scores by the highest novelty value in the population, as shown in Equation 9.

$$\alpha = \frac{\mathrm{nov(m)}}{\max(nov)} \qquad (9)$$

The dynamic weighting algorithm produces an oscillating effect on the probability of transitioning from one unit to another, as demonstrated in Figure 75. This shows the probability of an agent moving from unit A to unit B (the red line), and the probability of moving from unit A to unit C (the blue line). At the start of the model, the probability of going from unit A to B is 100%. Another agent in the population is trained with a probability of transferring from unit A to C 100% of the time. As our agents meet on the breeding grounds they hear the song, they hear this new song and deem it to be more novel than their own, thus applying more emphasis to learning it.

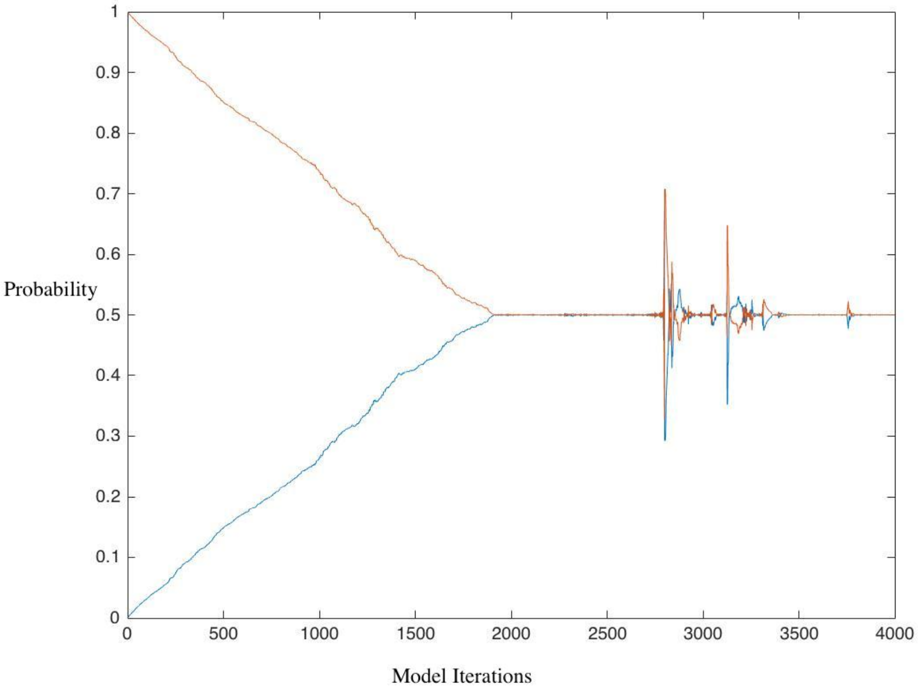


Figure 75: This figure demonstrates how the dynamic novelty algorithm creates an oscillation in the probability of transitioning from one unit to the other. (Transitioning from unit 1 to 2 in blue, transitioning from unit 1 to 3 in red).

## 6.5 MUSICAL DEMO

In order to test the model, testing was approached from a compositional point of view. First, four different musical themes were chosen to form the structure of the composition. These themes were chosen specifically because they have a high novelty value when compared to each other. They are also easily recognisable rudimentary musical themes. They consist of an ascending C major arpeggio (Theme A), a descending chromatic scale (Theme B), an ascending D minor arpeggio (Theme C), and a repeating G# (Theme D). These themes can be seen in Figure 76. At the start of our composition, every agent's transition matrix is trained by theme A. All subsequent themes were presented to only a single agent (agent 2). The songs being produced by agent 1 were recorded via MIDI.

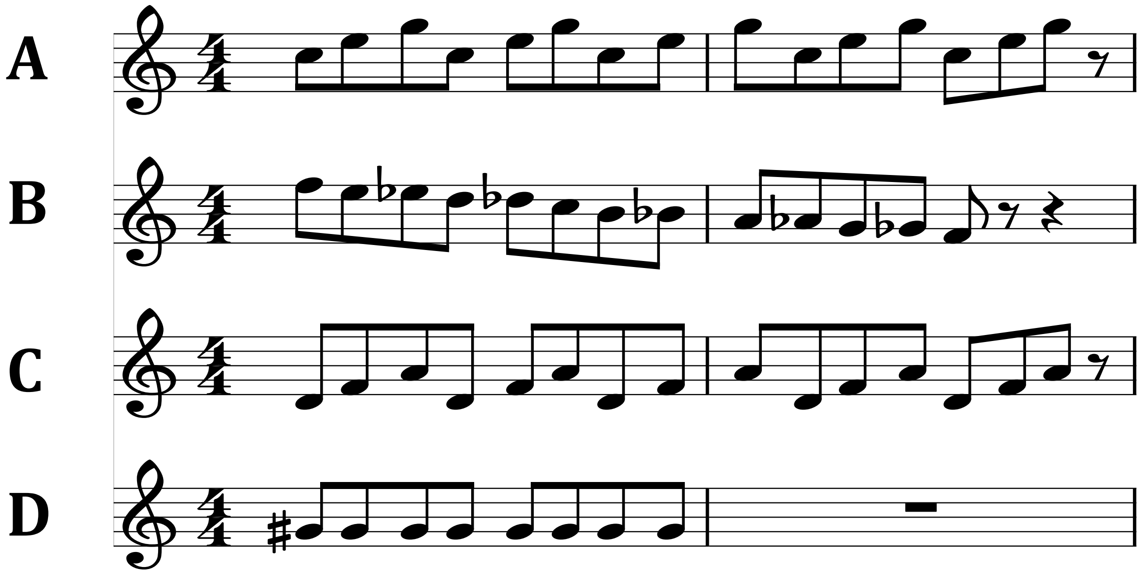



Figure 76: The four themes used in the composition.

The resulting composition is interesting, as the oscillatory nature described in Figure 75 of this chapter emerged not only for simple transitions as originally observed, but also for the structured themes presented to our population. Whenever a new theme was introduced, the agent would move between the newly introduced theme and the previous theme. This likely resulted because as after learning the SR for one theme, the theme would immediately cease to be novel. This meant that any other agent in the population still singing the previous theme would immediately have the most novel song. This created a type of musical tug of war between different themes.

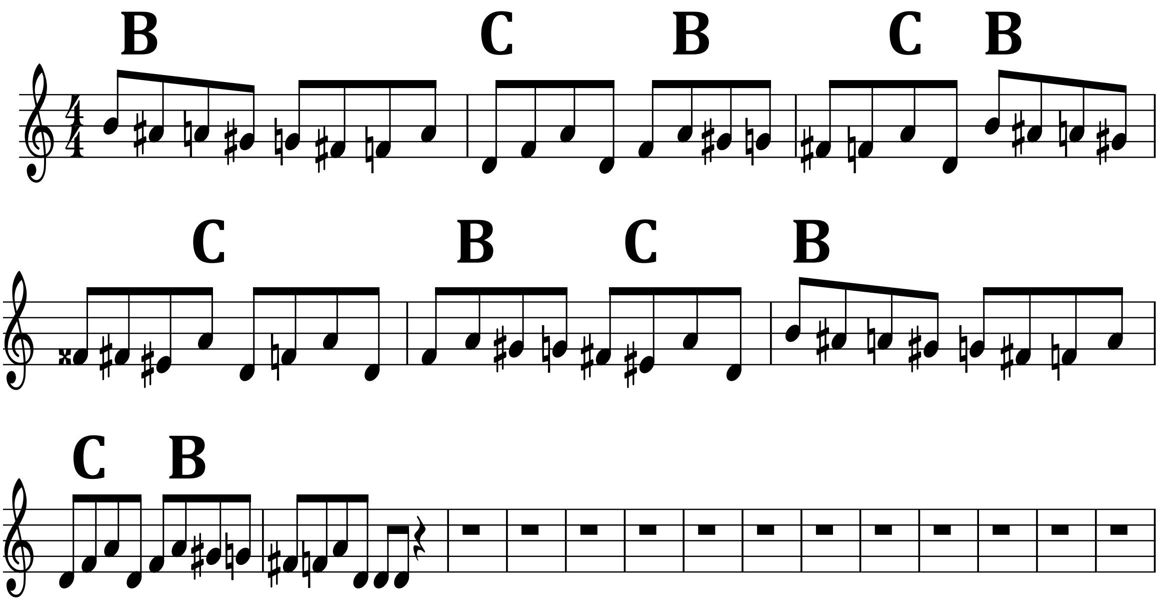


Figure 77: An example of theme oscillation. The agent alternated between a descending chromatic run and an ascending D minor arpeggio (theme C).

## 6.6 Discussion

From this chapter, we have seen that scientific methods for the analysis of animal vocalisations may be adapted for algorithmic composition. It explained the technical considerations necessary for interaction, and the development of a new novelty method that allows a composer to have a direct impact on the songs in the population. Emergent properties, such as theme oscillation and hybrid themes were also demonstrated through a compositional demo. It is necessary however to carry out a full investigation into the impact that the dynamic weighting novelty algorithm has on the evolution of songs in a larger population.

The emergence of oscillating themes is interesting, but this is a very simple musical concept. Gauging the musical success of these systems can be challenging, and the success of the system is often in the **ear** of the beholder. What is useful to one composer may not be useful to another. Similarly, the dynamic novelty algorithm generated a very simple musical idea, alternating between different themes. This does not make the system entirely useless though. In some cases, it forces a composer to work against what is currently the dominant song in the population. In order to generate a new theme that will influence the agents, a composer must consider transitions that the agents will not expect and forces them to adapt their input in creative ways in order to obtain a high novelty score. This constraint creates a musical game, where the composer is constantly working against a population of critics and forces them to reinvent their compositions in order to influence a population.

Finally, it would be interesting to incorporate production errors, and the Sequitur algorithm discussed in the previous chapter. Incorporating production errors would allow composer to simply input their themes and let the system run. If the agents incorporate production errors that are not dramatically different to the most popular song in the population, these edits would not have any effect on the songs in the population. If they incorporate a novel transition, their song would produce a high novelty score, leading this novel transition to be incorporated into the rest of the population. Incorporating Sequitur would allow for the generation of musical themes with more structure, but it is not entirely clear how the novelty algorithm could be adapted in order to facilitate the Sequitur algorithm. One solution is to have the agents SRs control the transition from rule to rule in the agents Seq variable. This would likely result in agents oscillating from a large variety of musical themes and would allow a composer to investigate how multiple themes in a single sequence could be remixed amongst a population.

Despite the ease in which a model may be adapted, this does not necessarily suggest that it is going to be useful for composing music. This model required a lot of work to generate what is arguably a very rudimentary musical concept. However, the model was trained with very simple musical phrases, and it is possible that given enough agents, and if trained with intelligent musical ideas, the model could produce an interesting result. This would require a talented composer simply highlights the human element in algorithmic composition. These models fall short compared to any human musical ability. However, this model may have alternative applications. One of the challenges in analysing humpback whale song is its long duration. It is difficult to identify the patterns aurally without the aid of a spectrogram. However, by assigning unique midi note values, it would be possible to sonify humpback whale song. This would allow for easy identification of the hierarchy of the song and may be of use as a tool for outreach. Furthermore, it would allow for a faster identification of how the song may change over time. In the original Fast Travel composition (Kirke, Freeman, Miranda, & Ingram, 2011a), changes in the song were sped up, reducing the time of change over the gradual drift of blue whale song from 10 years, to 10 minutes. By speeding up these rapid changes, we may be able to identify changes in a system over long periods of time. Furthermore, by adapting the model for use as a tool for composition brings it closer to use as a tool for studying the cultural transmission of music.

# Chapter 7 – Conclusions and Future Work

This final chapter discusses the contributions of the thesis and summarises the work presented.

## 7.1 Contributions

This thesis presented three multi-agent models that address specific issues surrounding humpback whale song. The first contribution is a model for investigating the role of migratory behaviour and spatial distribution in song learning (section 3.2), with extensions added to investigate the effect of novelty (section 3.3), production errors (section 3.4), and the coupling of these two parameters (section 3.5). Analysis of these models revealed the importance of feeding grounds as a location for cultural transmission that may facilitate cultural revolution (section 3.7 and 3.8), and also highlighted that simple Markov models are not capable of achieving song revolution as seen in wild populations of humpback whales (section 3.7 and 3.8).

The literature review reveals a distinct lack of knowledge surrounding the social networks of humpback whales on the breeding ground (section 2.2.2). In order to address this, a model was developed that synthesised social networks in order to control which agent's are capable of learning from each other (section 4.3). This model showed that social networks with small world properties consistently facilitate high levels of song convergence in a population (section 4.5) and outperform social networks that did not possess these properties. However, statistical analysis showed that there was not any statistical significance in the mean difference of convergence between random networks and small world networks. Since random networks facilitate high levels of song convergence as well as small world networks, they are also a likely candidate for the facilitation of song convergence, and do not rely on imposing as strict social network as small world networks. More research is necessary however in order to investigate the rate at which song convergence occurs in these networks.

The third model examined the issue of hierarchal structure in humpback whale song (Chapter 5). The literature review showed that Markov processes were inadequate for modelling humpback whale song (section 2.8.3 and 2.10), and the failure of Markov models to achieve true song revolution in the migratory model (section 3.7.1) suggested that Markov models were not capable of recreating the hierarchal structure of humpback whale song. In order to address this, the Sequitur algorithm was used to analyse the hierarchal structure of songs from 2002 and 2003 (section 5.3). This revealed the hierarchal structure of the song, showing that Sequitur could break down the song into rules resembling the hierarchal structure described in (R S Payne & McVay, 1971). Following this, a sampler was developed for the Sequitur algorithm in order to utilise the rules Sequitur generates in order to recreate humpback whale song. A variety of production errors were also implemented, and their individual effects on a real world song investigated in a vertical cultural transmission model (section 5.4.1 to 5.4.7). After identifying the effect that individual production errors had, they were coupled together in order to investigate how they may affect the transmission of the song (section 5.4.7). Over the course of a long run model. This resulted in a song that was significantly different to the original song that was used to train the model, but still possessing the hierarchal structure (section 5.4.8). This suggests that once some method of identifying the hierarchy of a song is in place, the changes observed in the song may simply be caused by relatively simple production and learning errors.

Finally, the migratory model was adapted in order to investigate its potential as a tool for composition (Chapter 6). The reason for adapting the model for music was to investigate how a composer may evolve a melody based on concepts in cultural transmission. Another reason is that adapting models of vocal learning in animals may help shed light on the concept of the musical proto-language theory, and highlight the distinct ways that animal vocalisations differ from music (section 6.1). The chapter showed how the model could easily be adapted using existing protocols in music technology (section 6.2.1). A new novelty algorithm was developed in order to address the shortcomings of the original novelty algorithm discussed in section 3.3. This lead to the development of a new novelty algorithm where the most novel song in the population has the largest impact on the overall song of the population (section 6.5.1). This new novelty algorithm caused agents to oscillate between the various themes they were trained with, and in some cases combine different themes in order to create hybrid themes (section 6.6), and paves the way for the using the model to investigate cultural transmission in music .

## 7.2 Research Questions Revisited

Here, the contributions and their achievements are related back to the research questions raised in section 1.3.

**RQ1:** Can methods used in computer music research be applied to the analysis of animal vocalisations?

Throughout this thesis, there has been a conscious effort to relate the methods used in the development of these models to the methods used in computer music research. Given the interdisciplinary nature of computer music, many of these methods have their origins in other disciplines, such as computer science, linguistics, and informatics. Despite these close relations, the thesis uses computer music research as a springboard for the development of the models, and there is indeed potential to draw on these methods for research into biology and cultural transmission. The ideas for this project originate in a composition (Kirke et al., 2011b), and a simple agent based on this compositional tool was developed in as a proof of concept for this project (Kirke et al., 2015). The first model developed for this project (the migratory model, discussed in section 3.2) was inspired heavily by an agent based model used to investigate cultural transmission in music (Todd & Werner, 1999). This included the use of Markov models, but also helped inspire the development of the original novelty algorithm discussed in section 3.3. The Sequitur model discussed in section 5.2 was designed to work with long strings of text, and has been used to identify structures in music (Nevill-Manning & Nevill-Manning, 1996). Sequitur has been compared with Probabilistic Context Free Grammars (PCFG) that have been used in composition, as discussed in section 2.8.6. The relative ease with which these methods could be adapted to examine cultural transmission in humpback whale song shows that these methods can easily be applied to animal vocalisation research.

**RQ2:** Can the phenomena of song revolution and evolution be addressed through a spatially explicit model that recreates humpback migratory patterns and fine-scale cultural transmission issues that may occur in individual interactions in humpback song transmission?

The spatial model in Chapter 3 of this thesis was designed to specifically address this question. It used the simplest method of song synthesis, and sound loss transmission as weighting for song learning (section 3.2.2). While the model itself did not succeed in recreating evolution or revolutionary patterns, it points the way towards what is required in order to achieve these phenomena (section 3.7 and section 3.8).

In regards to song revolution, Chapter 3 showed that feeding grounds are likely to play a key role in the phenomena of song revolution. Even with a low singing rate, it was possible for the geographically separated populations to achieve identical songs and song representations (section 3.7.1). However, this is not the same as song revolution, as the resulting SR's were more similar to a mixture of the two songs rather than one populations song replacing another. This suggests that other biases are necessary in order for song of a population to replace its neighbour, and the revolutionary pattern of western to eastern song replacement (Garland et al., 2011) is not explained. However, assuming that whales remain loyal to their respective feeding grounds, it is likely that revolutionary behaviour occurs on the feeding ground. Non-cognitive biases that may be responsible could be asymmetrical feeding grounds that facilitate song-learning opportunities. Another factor may be related to population size, and song replacement may occur when one breeding ground has a significantly higher population than the other. Another possibly important non-cognitive driving factor in song evolution could be the structure of the social networks of humpback whales, as shown in Chapter 4. This would be particularly true if the social structures were dynamic, and changed over the course of the whales migratory season.

In terms of individual interactions, the design of the agent based models in this thesis is always bottom up, meaning that when building these models, the role of individuals and how they learn and produce their songs is the main factor that drives cultural transmission and evolution. In this regard, the models have been successful in identifying sound loss transmission and production errors as important factors in song learning. More importantly, this model helps in eliminating cognitive biases that are not important in regards to song learning and production (section 3.8), such as the novelty algorithm developed in section 3.3. Novelty weightings caused agents to converge on unrealistic song representations, where the potential transitions that could occur had the same probability. This shows that in order for novelty to be a factor in song revolution, it may need to be coupled with other cognitive factors. At the time of writing, the author and his colleagues at University of St Andrews are developing a new model that incorporates song memory coupled with novelty that shows potential for revolutionary behaviour.

One of the most important factors individuals carry out appears to be production and learning errors, as exemplified by the models presented in sections 3.4 and 5.4. In a spatially explicit model, song changes would only occur when production errors were introduced, otherwise songs and song representations quickly converged on a single song and would not change. The most dramatic effect of production errors occurred when it was coupled with the hierarchical model shown in Chapter 5, and it was demonstrated that if an agent has the potential to identify hierarchy, simple production and learning errors might lead to dramatically different songs, given enough time (section 5.4.8).

**RQ3:** What different types of social networks could theoretically facilitate cultural transmission phenomena in humpback whales?

Chapter 4 was designed to address this question. It showed that social networks with small world properties allowed the highest level of song convergence in a population of whales (section 4.4 and 4.5). However, random networks can also facilitate song convergence just as well. In a population, all humpbacks generally conform to the same song. In order for this to occur, the social network of these animals must allow for consistently returning high levels of song convergence. This was related to social network analysis of transmitted feeding behaviours (section 4.6). Allen's analysis of the social networks found on the feeding grounds revealed an average of 50 associates for 653 whales examined. While this value was not tested due to time constraints, a model where agents were able to learn from as many as 30 conspecifics still returned a high average level of song convergence (section 4.5). It is important to note that while a whale may have many associates, it will not necessarily learn from every single one of them. This chapter also revealed that the lowest number of conspecifics an agent must learn from is 4 (section 4.4).

**RQ4:** Can formal grammar methods be used in the analysis of whale song, and synthesis of song when coupled with production and learning of humpback whale song?

As a tool for analysing humpback whale song, Chapter 5 demonstrated that algorithms that place an emphasis on hierarchy succeed in identifying the hierarchical patterns originally described in (R S Payne & McVay, 1971), with Sequitur dividing up the songs into its respective themes and phrases. It also helped demonstrate that is a wide variety in the complexity of the song. The song analysed in 2002 preceded a revolutionary event, and the song displayed a much lower higher degree of complexity than the song in 2003. This analysis was still relatively simple, and it should be noted that there are many more uses of where the Sequitur algorithm could be used in examining cultural evolution. Most importantly, following the transcription of song at the unit level, it provides a reliable method to identify the hierarchy of the song, without having to rely on humans in order to validate this structure of the song.

As discussed in section 2.8.6 and section 5.9, the use of formal grammars in the synthesis of animal vocal sequences is a controversial topic. However, the use of formal grammars in agent based models trained with sequences of humpback song have never been carried out, and the Sequitur algorithm combined with production and learning errors enabled the synthesis of evolving hierarchal songs similar to those seen in humpback whales. Some of the production and learning errors such as character deletion, substitution, and insertion produced predictable results (see sections 5.4.2 and 5.4.3) while others such as rule replacement, insertion and deletion were less predictable (sections 5.4.5 to 5.4.6). The most interesting results were obtained in long run models, where agents were allowed to carry out all these operations freely (section 5.4.8), suggesting that there is no one single operation that has an overpowering effect on the evolution of the song, and that the combination of these production and learning errors are necessary in order for these new songs to emerge.

**RQ5:** How can the methods developed to answer research questions 2 through 6 be used in order to create music?

As seen in Chapter 6, it was relatively easy to connect the model to Ableton Live software using OSC and Max4Live. This did not produce the most interesting music because of the models tendency towards short songs, and the model, while inspired by methods in computer music, still had a scientific focus. The musical results were unsatisfying, and required a more innovative approach in order to produce interesting results. The incorporation of population based novelty weighting presented in section 6.5.1 produced an interesting tug of war between different themes in the population (section 6.6). However, the model does succeed in showing that it is relatively painless in order to adapt an agent-based model for scientific research.

## 7.3 Suggestions for Future Research

In this section, potential avenues for future research are presented and discussed.

### 7.3.1 Migratory Social Network Models

A downfall of the social network model is that it does not incorporate the migratory movements of humpback whales. It also assumes that for these social network shapes are static, which is unlikely. Given the vast distances that humpbacks travel, it is much more likely that the networks through which they learn songs change as they migrate, being affected by geographical boundaries, and bottlenecks where whales from separate populations could possibly meet and learn each other's songs. It is also likely that a specific shape of social network could facilitate behaviour such as song revolution. However, combing these complex social networks with flocking and animal movement systems could prove unwieldy due to the already large parameter space increasing. The social network would also impose strict movement rules on the system. In reality, it is possible that the movement behaviour leads to these social structures, but the Migratory Model developed from Chapter 3 would require a complex analysis method in order to investigate if these social structures are emergent. In order to address this, a simpler grid based model could be implemented in order to investigate cultural transmission in social networks using the song learning and production methods discussed earlier.

This potential model would combine the social network methods outlined in Chapter 4 with the migratory patterns discussed in Chapter 3. In order to do this, social networks would be based on real world observations, of the number of humpback whales within each other's acoustic ranges at different points on their migratory cycle. Multiple unique social networks with different structures would then be synthesised to represent these different geographical locations. Agents would then spend a certain amount of time learning from each other in these social networks, before moving on to the next one.

### 7.3.2 Investigating the Role of Females

This thesis has focussed exclusively on male humpbacks, viewing their interactions as being the driving force behind the changes observed in songs. This was done so that the results would be consistent, across models and it would not be necessary to consider any type of inter-sexual interaction that is involved in song changes. One potential method would be to create a new type of agent, a female humpback, who can learn songs, but does not produce any. The female humpbacks could then choose whether to mate with a male humpback whale depending on how novel their song is. This would require male agents to adjust their songs to be the most novel in the population and would affect the songs that they produce. This is similar to the evolutionary pressure model implemented in (Todd & Werner, 1999).

### 7.3.3 Population Turnover and Genetics

None of the models here investigated the possible role a population turnover would have on the song of the whales. To implement this, it would first be necessary to implement female agents in the model as discussed above. If the agents are allowed to mate, they could be equipped with some type of genetic code in their structure. This could be a simple string. Following this, new agents could be created that possess some type of crossover of their parent's genetic string. If the role of females drove the song, this would result in a model to investigate how cultural transmission could affect the genetic make up of a population, and it would be possible to investigate if the songs and the genetic composition are related. This could be compared to other models that investigate the role of culture in evolution, such as those implemented in (Hal Whitehead, 2005; H Whitehead et al., 2002).

### 7.3.4 A Humpback Specific Formal Grammar

While the Sequitur algorithm was useful in recreating the hierarchical structure of humpback whale song, it is necessary to remember that it was not the original intention of the algorithm. It was designed to identify hierarchies in strings regardless of their origin. Sequitur points towards how it would be possible to develop a more realistic formal grammar for analysing and synthesising songs of humpback whales based on the description put forward by Payne (R S Payne & McVay, 1971).

In order to achieve such a model, it would be necessary to tie non-terminal symbols to specific terminology relating digrams to phrases, and how these phrases could be combined to create themes, and the order themes are performed in order to create songs. The result would provide a robust tool that would enable the comparison of songs, and allow researchers to identify hierarchies that do not rely on transcribing songs by hand. An example of a potential formal grammar is presented below.

*Song Structure* → ***T1, T1, T2, T2, T3, T3, T4, T4***

***T1*** **→ P1 P2**
**T2 → P3 P3**
**T3 → P1 P3**
**T4 → P4 P4 P5**
**P1 → a a**
**P2 → b b b b**
**P3 → c c d**
**P4 → d d d d**
**P5 → d a d d**

*Realised Song* → aabbbbaabbbbccdccdccdccdaaccdaaccddddddaddddddadd

### 7.3.5 Real Time Audio Models

This thesis investigated the songs of humpback whales from a discrete point of view. For every transcription used in this thesis, each sound was classified as a unit, and assigned a symbol for training the models. This means that we miss out on the finest level of detail in regards to vocal learning, and if there are mechanisms that could cause individual agents to change their individual units over time. Such a model would most closely resemble Mirandas' research into intonation discussed earlier in the thesis (Eduardo Reck Miranda, 2002) with particular attention being paid to the auditory and vocal production mechanisms.

In order to achieve this, a physics-based model of the humpback vocal system would need to be implemented. While research has been carried out in this area for toothed whales (Dubrovsky, Gladilin, Møhl, & Wahlberg, 2004), models for baleen whale vocalisations are still in their infancy. This is mainly due to a potential sound source (U-shaped folds) being identified ten years ago (Reidenberg & Laitman, 2007). However, there are currently researchers investigating potential methods for modelling the U-shaped folds (Adam et al., 2013; Cazau et al., 2016) so it may be possible to model this in the near future.

### 7.3.6 Development of Database of Song Analysis

One area that would benefit all future researchers in this area is the creation of a database the combines the different types of analyses discussed in this thesis. In order to structure such a database, it would be necessary to carry out analysis of all the songs in the Noad database using acoustic analysis, Markov chain analysis, Sequitur, and Levenshtein distance analysis. Researchers could type in the name of the recording they want, and the type of analysis they require. Many problems during the three years of this research involved extracting metadata, formatting songs, and hand based transcriptions. The development of a database such as this would enable research in this field to be carried out much more effectively.

### 7.3.7 Investigating the Influence of Humans

There have been efforts to investigate the role that humans have on marine mammals, these have mainly focussed on the role of whale watching boats and marine traffic (Chion et al., 2013, 2017; Parrott et al., 2011). The model could be adapted in order to investigate theories regarding the effect of anthropogenic noise on whale vocalisations. A recent study suggested that anthropogenic noise was creating a masking effect that was causing humpback whales increase repetition in their communication signals (R. A. Dunlop, Cato, & Noad, 2010). In order to investigate this theory, a new type of agent could be introduced to the migratory model. Specifically, a ship agent that is equipped with a noise model. If a ship agent was too close to a whale, the agent may respond by repeating the number of times they sing a phrase. This model is very simplistic, but would be useful as a demonstration tool, and also could be implemented in a more sophisticated fashion.

### 7.3.8 Application of Models to Other Species

In the broad scope, this work could be considered part of the field of evolutionary linguistics. One of the key topics in evolutionary linguistics is why certain species demonstrate vocal learning while others do not. For this reason, it would be interesting to see how the models presented in this thesis perform when they are trained using data from other animals. The Migratory model would not be suitable for such research, since its design is closely tied to theories surrounding the migratory movements of humpback whales. The social network and formal grammar model however could be trained using song sequences of birds, or even other cetaceans such as the bow-headed whale.

## 7.4 Conclusion

While these models shed light on factors such as the importance of feeding grounds as key locations for cultural transmission, and the importance of production and learning errors for song evolution, there are still many things we do not understand about humpback whales. Specifically, what is it that drives song revolution? What role (if any) do female humpbacks play in song evolution and revolution? How may realistic auditory and vocal models influence the evolution of the song? Why do humpbacks sing such complicated songs compared to other members of the cetacean family?

We are still seeking answers to these questions. When we do find them, more will emerge. However, this is one of the most profound aspects of researching these animals. Humpbacks have always been mysterious, and they evoke something deep inside us. We become enamoured with these charismatic animals and feel genuinely moved by them. Perhaps there is a certain amount of irony here. As scientists, there is a danger of anthropomorphising these animals because of the feelings that they stir in us. Yet the methods used in this thesis were developed in a computer music research lab, and the origins of this project lie in the creation of music. This is dangerous territory of course, and every effort has been taken throughout this thesis to avoid projecting human characteristics onto these animals, but the danger remains. However, this is the spirit of interdisciplinary research that makes it so exciting. To take ideas from other fields, pay attention to their academic rigour, and then apply them to the questions we have about the world. I hope that this thesis has achieved this, and I hope that, if anyone is still reading by now, that this work inspires the same type of passion for researching these animals.

# References

Abe, K., & Watanabe, D. (2011). Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience*, *14*(8), 1067–1074. http://doi.org/10.1038/nn.2869

Ableton. (2017a). Live. Ableton. Retrieved from https://www.ableton.com/en/

Ableton. (2017b). Max4Live. Retrieved from https://www.ableton.com/en/live/max-for-live/

Adam, O., Cazau, D., Gandilhon, N., Fabre, B., Laitman, J. T., & Reidenberg, J. S. (2013). New acoustic model for humpback whale sound production. *Applied Acoustics*, *74*(10), 1182–1190. http://doi.org/10.1016/j.apacoust.2013.04.007

Adobe. (2017). Audition.

Allen, J., Garland, E. C., Murray, A., Noad, M. J., & Dunlop, R. A. (2017). Quantitative classification of acoustic signals using exemplar units: classifying humpback whale song. *Journal for the Acoustical Society of America*.

Allen, J., Garland, E., Noad, M., & Dunlop, R. (2015). quantifying the rate of song change in the east australian humpback whale population. In *Society of Marine Mammology 21st Biennial Conference*.

Allen, J., Weinrich, M., Hoppitt, W., Rendell, L., James, R., Allen, J., … Eskelinen, H. (2013). Network-based diffusion analysis reveals cultural transmission of lobtail feeding in humpback whales. *Science (New York, N.Y.)*, *340*(6131), 485–8. http://doi.org/10.1126/science.1231976

Arriaga, G., & Jarvis, E. D. (2013). Mouse vocal communication system: Are ultrasounds learned or innate? *Brain and Language*. http://doi.org/10.1016/j.bandl.2012.10.002

Au, W. (2004). The sonar of dolphins. *Acoustics Australia*, *32*(2), 61–63. http://doi.org/10.1121/1.408312

Baker, S. C., Herman, L. M., Perry, A., Lawton, W. S., Straley, J. M., Wolman, A. a., … Ostman, J. (1986). Migratory movement and population structure of humpback whales (Megaptera novaeangliae) in the central and eastern North Pacific. *Marine Ecology Progress Series*, *31*, 105–119. http://doi.org/10.3354/meps031105

Beckers, G. J. L., Bolhuis, J. J., Okanoya, K., & Berwick, R. C. (2012). Birdsong neurolinguistics: songbird context-free grammar claim is premature. *Neuroreport*, *23*(3), 139–45. http://doi.org/10.1097/WNR.0b013e32834f1765

Berwick, R. C., Beckers, G. J. L., Okanoya, K., & Bolhuis, J. J. (2012). A bird's eye view of human language evolution. *Frontiers in Evolutionary Neuroscience*, *4*(APR), 1–25. http://doi.org/10.3389/fnevo.2012.00005

Berwick, R. C., Okanoya, K., Beckers, G. J. L., & Bolhuis, J. J. (2011). Songs to syntax: The linguistics of birdsong. *Trends in Cognitive Sciences*. http://doi.org/10.1016/j.tics.2011.01.002

Brainard, M. S., & Doupe, A. J. (2002). What songbirds teach us about learning. *Nature*, *417*(May), 351–358. http://doi.org/10.1038/417351a

Brenowitz, E. A., Margoliash, D., & Nordeen, K. W. (1997). An introduction to birdsong and the avian song system. *Journal of Neurobiology*. http://doi.org/10.1002/(SICI)1097-4695(19971105)33:5<495::AID-NEU1>3.0.CO;2-#

Cantor, M., Shoemaker, L. G., Cabral, R. B., Flores, C. O., Varga, M., & Whitehead, H. (2015). Multilevel animal societies can emerge from cultural transmission. *Nature Communications*, *6*, 8091. http://doi.org/10.1038/ncomms9091

Cantor, M., & Whitehead, H. (2013). The Interplay between Social Networks and Culture: Theoretically and among Whales and Dophins. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*(April), 1–10. http://doi.org/10.1098/rstb.2012.0340

Cate, C., Okanoya, K., Cate, C., & Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals : natural vocalizations and artificial grammar learning Revisiting the syntactic abilities of non- human animals : natural vocalizations and artificial grammar learning. *Philosophical Transactions of the Royal Society B*, (June). http://doi.org/10.1098/rstb.2012.0055

Cathcart, C. (2014). Pirates of the middle Cs: Edward Bunting and the Pirated Editions of His 1796 Collection. *Bulletin of the Historical Harp Society*, *24*(3), 14–22.

Cato, D., & McCauley, R. (2006). Passive acoustics for monitoring marine animals: progress and challenges. *… of ACOUSTICS 2006*, (November), 453–460.

Retrieved from http://www.acoustics.asn.au/conference_proceedings/AASNZ2006/papers/p117.pdf?origin=publication_detail

Cazau, D., Adam, O., Aubin, T., Laitman, J. T., & Reidenberg, J. S. (2016). A study of vocal nonlinearities in humpback whale songs: from production mechanisms to acoustic analysis. *Nature: Scientific Reports*, *6*(April), 31660. http://doi.org/10.1038/srep31660

Chion, C., Cantin, G., Dionne, S., Dubeau, B., Lamontagne, P., Landry, J. A., … Turgeon, S. (2013). Spatiotemporal modelling for policy analysis: Application to sustainable management of whale-watching activities. *Marine Policy*, *38*, 151–162. http://doi.org/10.1016/j.marpol.2012.05.031

Chion, C., Lagrois, D., Dupras, J., Turgeon, S., McQuinn, I. H., Michaud, R., … Parrott, L. (2017). Underwater acoustic impacts of shipping management measures: Results from a social-ecological model of boat and whale movements in the St. Lawrence River Estuary (Canada). *Ecological Modelling*, *354*, 72–87. http://doi.org/10.1016/j.ecolmodel.2017.03.014

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory*, *2*(3). http://doi.org/10.1109/TIT.1956.1056813

Clark, C. W. (1990). Acoustic behaviour of mysticete whales. In *Sensory Abilities of Cetaceans* (pp. 571–583).

Cleary, J. G., & Witten, I. H. (1984). Data compression using adaptive coding and partial string matching. *IEEE Transactions on Communications*, *COM-32*, 396–402. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.14.4305

Couzin, I. D., Krause, J., James, R., Ruxton, G. D., & Franks, N. R. (2002). Collective memory and spatial sorting in animal groups. *Journal of Theoretical Biology*, *218*(1), 1–11. http://doi.org/10.1006/yjtbi.3065

Croll, D. A., Clark, C. W., Acevedo, A., Tershy, B., Flores, S., Gedamke, J., … Suki, B. (2002). Only male fin whales sing loud songs. *Nature*, *417*(6891), 809. http://doi.org/10.1038/417809a

Darling, J. D., Jones, M. E., & Nicklin, C. P. (2006). Humpback whale songs : Do they organize males during the breeding season ? *Behaviour*, *9*(143), 1051–

1101.

Darwin, C. (1871). *The Descent of Man, and Selection In Relation to Sex*. John Murray.

Dawbin, W. (1966). The Seasonal Migratory Cycle of Humpback Whales. In K. Norris (Ed.), *Whales, Dolphins, and Porpoises*. University of California Press.

de Boer, B. (2002). Evolving Sound Systems. In A. Cangelosi & D. Parisi (Eds.), *Simulating The Evolution of Language* (pp. 80–97).

Deecke, V. B., Ford, J. K. B., & Spong, P. (2000). Dialect change in resident killer whales: implications for vocal learning and cultural transmission. *ANIMAL BEHAVIOUR*, *60*, 629–638. http://doi.org/10.1006

Delarue, J., Todd, S. K., Van Parijs, S. M., & Di Iorio, L. (2009). Geographic variation in Northwest Atlantic fin whale (Balaenoptera physalus) song: implications for stock structure assessment. *The Journal of the Acoustical Society of America*, *125*(3), 1774–1782. http://doi.org/10.1121/1.3068454

Dubrovsky, N. a., Gladilin, a., Møhl, B., & Wahlberg, M. (2004). Modeling of the dolphin's clicking sound source: The influence of the critical parameters. *Acoustical Physics*, *50*(4), 463–468. http://doi.org/10.1134/1.1776225

Dunlop, R. A., Cato, D. H., & Noad, M. J. (2010). Your attention please: increasing ambient noise levels elicits a change in communication behaviour in humpback whales (Megaptera novaeangliae). *Proceedings of The Royal Society - Biological Sciences*, *277*(1693), 2521–9. http://doi.org/10.1098/rspb.2009.2319

Dunlop, R. A., Noad, M. J., Cato, D. H., & Stokes, D. M. (2007). The social vocalization repertoire of east Australian migrating humpback whales (*Megaptera novaeangliae*). *Journal of the Acoustical Society of America*, *122*(5), 2893–2905. http://doi.org/10.1121/1.2783115

Eddy, S. R. (2004). What is a hidden Markov model? *Nature Biotechnology*, *22*(10), 1315–1316. http://doi.org/10.1038/nbt1004-1315

Eerkens, J. W., & Lipo, C. P. (2005). Cultural transmission, copying errors, and the generation of variation in material culture and the archaeological record. *Journal of Anthropological Archaeology*, *24*(4), 316–334. http://doi.org/10.1016/j.jaa.2005.08.001

Ellers, J., & Slabbekoorn, H. (2003). Song divergence and male dispersal among bird populations: a spatially explicit model testing the role of vocal learning. *Animal Behaviour*, *65*(4), 671–681. http://doi.org/10.1006/anbe.2003.2081

Fehér, O., Wang, H., Saar, S., Mitra, P. P., & Tchernichovski, O. (2009). De novo establishment of wild-type song culture in the zebra finch. *Nature*, *459*(7246), 564–568. http://doi.org/10.1038/nature07994

Fitch, W. T. (2010). *The Evolution of Language*. Cambridge, New York, Melbourned, Madrid, Cape Town, Singapore, Sao Paulo, Delhi, Dubai, Toky: Cambridge University Press.

Fitch, W. T. (2013). Musical Protolanguage: Darwin's Theory of Language Evolution Revisited. In *Birdsong, Speech, and Languag: Exploring the Evolution of Mind and Brain* (pp. 489–503).

Fitch, W. T., & Friederici, A. D. (2012). Artificial grammar learning meets formal language theory: an overview. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences*, *367*(1598), 1933–55. http://doi.org/10.1098/rstb.2012.0103

Foote, A. D., Vijay, N., Ávila-Arcos, M. C., Baird, R. W., Durban, J. W., Fumagalli, M., … Wolf, J. B. W. (2016). Genome-culture coevolution promotes rapid divergence of killer whale ecotypes. *Nature Communications*, *7*, 11693. http://doi.org/10.1038/ncomms11693

Ford, J. K. B. (1991). Vocal traditions among resident killer whales (Orcinus orca) in coastal waters of British Columbia. *Canadian Journal of Zoology*, *69*, 1454–1483. http://doi.org/10.1139/z91-206

Freed, Adrian, and Wright, M. (2018). Open Sound Control. Centre for New Music and Audio Technologies. Retrieved from http://opensoundcontrol.org/introduction-osc

Friedlaender, A., Weinrich, M., Bocconcelli, A., Cholewiak, D., Thompson, M., Ware, C., & Wiley, D. (2011). Underwater components of humpback whale bubble-net feeding behaviour. *Behaviour*, *148*(5), 575–602. http://doi.org/10.1163/000579511X570893

Frumhoff, P. (1983). Aberrant Songs of Humpback Whales (*Magaptera navaengliae*): Clues to the Structure of Humpback Songs. *Communication and Behavior of Whales*, 81–127.

*Fundamentals of Underwater Sound*. (2008). Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Fundamentals+of+underwater+sound#8

Fuß, H.-U. (2005). Introduction to Schenkerian Analysis. In *Nordamerikanische Musiktheorie* (p. 272).

Garland, E. C., Gedamke, J., Rekdahl, M. L., Noad, M. J., Garrigue, C., & Gales, N. (2013a). Humpback Whale Song on the Southern Ocean Feeding Grounds : Implications for Cultural Transmission. *PLoS ONE*, *8*(11). http://doi.org/10.1371/journal.pone.0079422

Garland, E. C., Gedamke, J., Rekdahl, M. L., Noad, M. J., Garrigue, C., & Gales, N. (2013b). Humpback whale song on the Southern Ocean feeding grounds: Implications for cultural transmission. *PLoS ONE*, *8*(11). http://doi.org/10.1371/journal.pone.0079422

Garland, E. C., Goldizen, A. W., Rekdahl, M. L., Constantine, R., Garrigue, C., Hauser, N. D., … Noad, M. J. (2011). Dynamic horizontal cultural transmission of humpback whale song at the ocean basin scale. *Current Biology*, *21*(8), 687–691. http://doi.org/10.1016/j.cub.2011.03.019

Garland, E. C., Lilley, M. S., Goldizen, A. W., Rekdahl, M. L., Garrigue, C., & Noad, M. J. (2012). Improved versions of the Levenshtein distance method for comparing sequence information in animals' vocalisations: tests using humpback whale song. *Behaviour*, *149*(13–14), 1413–1441. http://doi.org/10.1163/1568539X-00003032

Garland, E. C., Noad, M. J., Goldizen, A. W., Lilley, M. S., Rekdahl, M. L., Garrigue, C., … Robbins, J. (2013). Quantifying humpback whale song sequences to understand the dynamics of song exchange at the ocean basin scale. *The Journal of the Acoustical Society of America*, *133*(1), 560–569. http://doi.org/10.1121/1.4770232

Gedamke, J., Costa, D. P., & Dunstan, A. (2001). Localization and visual verification of a complex minke whale vocalization. *The Journal of the Acoustical Society of America*, *109*(6), 3038–3047. http://doi.org/10.1121/1.1371763

Gero, S., Whitehead, H., & Rendell, L. (2016). Individual , unit and vocal clan level identity cues in sperm whale codas. *Royal Society Open Science*, *3*, 1–12. http://doi.org/10.1098/rsos.150372

Gilbert, B. K. (1999). Opportunities for social learning in bears. In *Mammalian Social Learning: Comparative and Ecological Perspectives* (pp. 225–235). Retrieved from http://books.google.com/books?hl=ru&lr=&id=NS3So3ma1k8C&pgis=1

Gillick, J., Tang, K., & Keller, R. M. (2009). LEARNING JAZZ GRAMMARS. In *Sound and Music Computing Conference* (pp. 23–25).

Glockner, D. A. (1983). Determining the sex of humpback whales(Megaptera novaeangliae) in their natural environment. In R. Payne (Ed.), *Communication and behavior of whales* (pp. 447–464). Boulder: Westview Press.

Goodfellow, D. ., & Slater, P. J. . (1986). A Model of Bird Song Dialects. *Animal Behaviour*, *34*.

Gurney, M. E., & Konishi, M. (1980). Hormone-induced sexual differentiation of brain and behavior in zebra finches. *Science (New York, N.Y.)*, *208*(4450), 1380–3. http://doi.org/10.1126/science.208.4450.1380

Hamanaka, M., Hirata, K., & Tojo, S. (2015). σ GTTM III : Learning based Time-span Tree Generator based on PCFG. In *Proceedings of the 11th International Symposium on CMMR* (pp. 1–15). Plymouth, UK.

Hartman, C., & Beneš, B. (2006). Autonomous boids. In *Computer Animation and Virtual Worlds* (Vol. 17, pp. 199–206). http://doi.org/10.1002/cav.123

Hauber, M. E., Campbell, D. L. M., & Woolley, S. M. N. (2010). The functional role and female perception of male song in Zebra Finches. *Emu*, *110*(3), 209–218. http://doi.org/10.1071/MU10003

Hopcroft, J. E., Motwani, R., & Ullman, J. D. (2001). Introduction to automata theory, languages, and computation, 2nd edition. *ACM SIGACT News*. http://doi.org/10.1145/568438.568455

Hurford, J. (2011). *The Origins of Grammar: Language in the Light of Evolution II*. Oxford University Press.

Janik, V. M. (2014). Cetacean vocal learning and communication. *Current Opinion in Neurobiology*. http://doi.org/10.1016/j.conb.2014.06.010

Janik, V. M., & Slater, P. J. . (1998). Context-specific use suggests that bottlenose dolphin signature whistles are cohesion calls. *Animal Behaviour*, *56*(4), 829–838. http://doi.org/10.1006/anbe.1998.0881

Jarvis, E. D. (2007). Neural systems for vocal learning in birds and humans: A synopsis. In *Journal of Ornithology* (Vol. 148). http://doi.org/10.1007/s10336-007-0243-0

Jones, Eric and Oliphant, T. and Peterson, P. (2001). SciPy: Open Source Scientific Tools for Python. Retrieved from https://www.scipy.org/

Kato, Y., Hasegawa, T., & Okanoya, K. (2010). Song preference of female Bengalese finches as measured by operant conditioning. *Journal of Ethology*, *28*(3), 447–453. http://doi.org/10.1007/s10164-010-0203-7

Kavanagh, A. S., Owen, K., Williamson, M. J., Blomberg, S. P., Noad, M. J., Goldizen, A. W., … Dunlop, R. A. (2016). Evidence for the functions of surface-active behaviors in humpback whales (Megaptera novaeangliae). *Marine Mammal Science*, *33*(January), 1–22. http://doi.org/10.1111/mms.12374

Kershenbaum, A., & Blumstein, D. (2014). Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biological Reviews*. http://doi.org/10.1111/brv.12160

Kershenbaum, A., Bowles, A. E., Freeberg, T. M., Jin, D. Z., Lameira, A. R., & Bohn, K. (2014). Animal vocal sequences: not the Markov chains we thought they were. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1792), 20141370. http://doi.org/10.1098/rspb.2014.1370

King, S. L., Harley, H. E., & Janik, V. M. (2014). The role of signature whistle matching in bottlenose dolphins, Tursiops truncatus. *Animal Behaviour*, *96*. http://doi.org/10.1016/j.anbehav.2014.07.019

King, S. L., Sayigh, L. S., Wells, R. S., Fellner, W., Janik, V. M., Janik, V., … Prinage, S. (2013). Vocal copying of individually distinctive signature whistles in bottlenose dolphins. *Proceedings. Biological Sciences / The Royal Society*, *280*(1757), 20130053. http://doi.org/10.1098/rspb.2013.0053

Kirby, S. (2001). Spontaneous evolution of linguistic structure - An iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation*, *5*(2), 102–110. http://doi.org/10.1109/4235.918430

Kirby, S. (2002a). Learning, bottlenecks and the evolution of recursive syntax. *Linguistic Evolution through Language Acquisition: Formal and Computational Models*, 173–204.

Kirby, S. (2002b). The Emergence of Linguistic Structure: An Overview of the Iterated Learning Model. In *Simulating The Evolution of Language* (pp. 121–147).

Kirke, A., Freeman, S., Miranda, E., & Ingram, S. (2011a). Application of Intermediate Multi-Agent Systems to Integrated Algorithmic Composition and Expressive Performance of Music. In *Proceedings of International Computer Music Conference*. Huddersfield, UK. Retrieved from http://lib-srvr9.lib.plymouth.ac.uk:8080/handle/10026.1/895

Kirke, A., Freeman, S., Miranda, E., & Ingram, S. (2011b). Application of Multi-Agent Whale Modelling to an Interactive Saxophone and Whales Duet. In *Proceedings of International Computer Music Conference* (pp. 350–353). Huddersfield, UK.

Kirke, A., & Miranda, E. R. (2009). A survey of computer systems for expressive music performance. *ACM Computing Surveys*, *42*(1), 1–41. http://doi.org/10.1145/1592451.1592454

Kirke, A., Miranda, E., Rendell, L., & Ingram, S. (2015). Towards Modelling Song Propagation in Humpback Whales. In *Proceedings of 2015 Conference on Transdisciplinary Approaches to Cognitive Innovation*. Plymouth, UK.

Kogan, J. a. (1998). Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: A comparative study. *The Journal of the Acoustical Society of America*, *103*(4), 2185. http://doi.org/10.1121/1.421364

Kohonen, T. (1997). Exploration of very large databases by self-organizing maps. In *Proceedings of International Conference on Neural Networks (ICNN'97)* (Vol. 1, p. PL1-PL6). http://doi.org/10.1109/ICNN.1997.611622

Krause, J., Lusseau, D., & James, R. (2009). Animal social networks: An introduction. *Behavioral Ecology and Sociobiology*. http://doi.org/10.1007/s00265-009-0747-0

Lachlan, R. F., van Heijningen, C. A. A., ter Haar, S. M., & ten Cate, C. (2016). Zebra finch song phonology and syntactical structure across populations and continents-a computational comparison. *Frontiers in Psychology*, *7*(JUL). http://doi.org/10.3389/fpsyg.2016.00980

Le Cun Jackel, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L.

D., Cun, B. Le, Denker, J., & Henderson, D. (1990). Handwritten Digit Recognition with a Back-Propagation Network. *Advances in Neural Information Processing Systems*, 396–404. http://doi.org/10.1111/dsu.12130

Leighton, T. G., Finfer, D., Grover, E., & White, P. R. (2007). Spiral bubble nets of humpback whales: An acoustic mechanism. *Conference of Underwater Acoustic Measurements, Technologies and Results, Heraklion, Crete.*, *2*, 583–588.

Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The Importance of the Normality Assumption in Large Public Health Data Sets. *Annual Review of Public Health*, *23*(1), 151–169. http://doi.org/10.1146/annurev.publhealth.23.100901.140546

McDonald, M. A., Hildebrand, J. A., & Mesnick, S. (2009). Worldwide decline in tonal frequencies of blue whale songs. *Endangered Species Research*, *9*(1), 13–21. http://doi.org/10.3354/esr00217

McDonald, M. A., Mesnick, S. L., & Hildebrand, J. A. (2006). Biogeographic characterisation of blue whale song worldwide: using song to identify populations. *Journal of Cetacean Research and Management*, *8*(1), 55–65. http://doi.org/10.1029/2006WR005124.DOI

Miranda, E. R. (2002). Mimetic Development of Intonation. In *Proceedings of the 2nd International Conference on Music and Artificial Intelligence*.

Miranda, E. R. (2003). On the Music of Emergent Behavior: What Can Evolutionary Computation Bring to the Musician? *Leonardo*, *36*(1), 55–59. http://doi.org/10.1162/002409403321152329

Miranda, E. R. (Ed.). (2011). *A-life for Music: Music and Computer Models of Living Systems*. Middleton, Wisconson: A-R Editions, Inc.

Miranda, E. R., & Biles, J. A. (2007). *Evolutionary computer music*. *Evolutionary Computer Music*. http://doi.org/10.1007/978-1-84628-600-1

Miranda, E. R., Kirby, S., & Todd, P. (2003a). On computational models of the evolution of music: From the origins of musical taste to the emergence of grammars. *Contemporary Music Review*, *22*(3), 91–111. http://doi.org/10.1080/0749446032000150915

Miranda, E. R., Kirby, S., & Todd, P. (2003b). On Computational Models of the

Evolution of Music: From the Origins of Musical Taste to the Emergence of Grammars. *Contemporary Music Review*. http://doi.org/10.1080/0749446032000150915

Miranda, E. R., & Todd, P. M. (1999). A-Life and Musical Composition : A Brief Survey. *Adaptive Behavior*, (Degazio).

Morin, P. a, Archer, F. I., Foote, A. D., Vilstrup, J., Allen, E. E., Wade, P., … Harkins, T. (2008). Complete mitochondrial genome phylogeographic analysis of killer whales ( Orcinus orca ) indicates multiple species. *Genome Research*, *20*(858), 908–916. http://doi.org/10.1101/gr.102954.109.908

Mukherjee, S., & Mitra, S. (2005). HIDDEN MARKOV MODELS, GRAMMARS, AND BIOLOGY: A TUTORIAL. *Journal of Bioinformatics and Computational Biology*, *3*(2), 491–526. http://doi.org/10.1142/S0219720005001077

Murray, A., Antunes, R., Dunlop, R. A., & Noad, M. J. (2015). Hidden Markov models reveal the structural differences between simple and complex phrase types in the songs of humpback whales (Megaptera novaeangliae). In *21st Biennial Conference on the Biology of Marine Mammals, San Francisco, California, United States.*

NetworkX. (n.d.). Retrieved from https://networkx.github.io/documentation.html

Nevill-Manning, C. G., & Nevill-Manning, C. G. (1996). Inferring Sequential Structure. *Phdthesis*, (May), 212.

Nevill-manning, C. G., & Witten, I. H. (1997a). Identifying Hierarchical Structure in Sequences : A linear-time algorithm, *7*, 67–82.

Nevill-manning, C. G., & Witten, I. H. (1997b). Identifying Hierarchical Structure in Sequences : A linear-time algorithm. *Journal of Artificial Intelligence Research*, *7*, 67–82.

Newton, M. J., & Smith, L. S. (2011). Biologically-inspired neural coding of sound onset for a musical sound classification task. In *Proceedings of the International Joint Conference on Neural Networks* (pp. 1386–1393). http://doi.org/10.1109/IJCNN.2011.6033386

Noad, Michael J., Cato, Douglas H., Bryden, M. M., Micheline, Jenner, Jenner, C. (2000). Cultural Revolution in Whale Songs. *Nature*, *408*(537).

Ojima, S., & Okanoya, K. (2014). The Non-Hierarchical Nature of the Chomsky

Hierarchy-Driven Artificial-Grammar Learning. *Biolinguistics*, (23240033), 163–180. Retrieved from http://biolinguistics.eu/index.php/biolinguistics/article/view/348

Okanoya, K. (2004). Song Syntax in Bengalese Finches: Proximate and Ultimate Analyses. *Advances in the Study of Behavior*, *34*, 297–346. http://doi.org/10.1016/S0065-3454(04)34008-8

Oleson, E., Calambokidis, J., Burgess, W., McDonald, M., LeDuc, C., & Hildebrand, J. (2007). Behavioral context of call production by eastern North Pacific blue whales. *Marine Ecology Progress Series*, *330*, 269–284. http://doi.org/10.3354/meps330269

Oliveira, C., Wahlberg, M., Silva, M. A., Johnson, M., Antunes, R., Wisniewska, D. M., … Madsen, P. T. (2016). Sperm whale codas may encode individuality as well as clan identity. *The Journal of the Acoustical Society of America*, *139*(5), 2860–2869. http://doi.org/10.1121/1.4949478

Pace, F., White, P., & Adam, O. (2012). Hidden Markov Modeling for humpback whale (Megaptera Novaeanglie) call classification. *Proceedings of Meetings on Acoustics*, *17*(1), 70046. http://doi.org/10.1121/1.4772751

Parrott, L., Chion, C., Martins, C. C. A., Lamontagne, P., Turgeon, S., Landry, J. A., … Dionne, S. (2011). A decision support system to assist the sustainable management of navigation activities in the St. Lawrence River Estuary, Canada. *Environmental Modelling and Software*, *26*(12), 1403–1418. http://doi.org/10.1016/j.envsoft.2011.08.009

Payne, K., Tyack, P., & Payne, R. (1983). Progressive changes in the songs of humpback whales (Megaptera novaeangliae): A detailed analysis of two seasons in hawaii. In *Communication and Behavior of Whales* (pp. 9–57).

Payne, K., & Payne, R. S. (1985). Large scale changes over 19 years in songs of humpback whales in Bermuda. *Zeitschrift Für Tierpsychologie*, *68*(2), 89–114. http://doi.org/10.1111/j.1439-0310.1985.tb00118.x

Payne, R. S., & Mc, S. (1971). Sci Enc : E, *173*(3997).

Payne, R. S., & McVay, S. (1971). Songs of humpback whales. *Science (New York, N.Y.)*, *173*(3997), 585–597. http://doi.org/10.1126/science.173.3997.585

Prusinkiewicz, P. (1986). Score Generation with L-systems. *Proceedings of the International Computer and Music Conference*, 455–457.

Prusinkiewicz, P., & Lindenmayer, A. (1997). The algorithmic beauty of plants. *Plant Science*, *122*(1), 109–110. http://doi.org/10.1016/S0168-9452(96)04526-8

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257–286. http://doi.org/10.1109/5.18626

Reidenberg, J. S., & Laitman, J. T. (2007). Discovery of a low frequency sound source in mysticeti (baleen whales): Anatomical establishment of a vocal fold homolog. *Anatomical Record*, *290*(6), 745–759. http://doi.org/10.1002/ar.20544

Reiss, D., & McCowan, B. (1993). Spontaneous vocal mimicry and production by bottlenose dolphins (Tursiops truncatus): evidence for vocal learning. *Journal of Comparative Psychology*. http://doi.org/10.1037/0735-7036.107.3.301

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., … Laland, K. N. (2010). Why copy others? Insights from the social learning strategies tournament. *Science (New York, N.Y.)*, *328*(5975), 208–213. http://doi.org/10.1126/science.1184719

Rendell, L. E., & Whitehead, H. (2003). Vocal clans in sperm whales (Physeter macrocephalus). *Proceedings. Biological Sciences / The Royal Society*, *270*(1512), 225–31. http://doi.org/10.1098/rspb.2002.2239

Rendell, L., & Whitehead, H. (2001). Cetacean culture: Still afloat after the first naval engagement of the culture wars. *Behavioral and Brain Sciences*, *24*(2), S0140525X01613967. http://doi.org/10.1017/S0140525X01613967

Reynolds, C. (1995). Boids Background and Update. Retrieved from http://www.red3d.com/cwr/boids/

Riebel, K., Lachlan, R. F., & Slater, P. J. B. (2015). Learning and cultural transmission in chaffinch song. *Advances in the Study of Behavior*, *47*, 181–227. http://doi.org/10.1016/bs.asb.2015.01.001

Rimmer, J. (1987). Patronage, style and structure in the music attributed to turlough Carolan. *Early Music*, *15*(2), 164–174. http://doi.org/10.1093/earlyj/XV.2.164

Ritschard, M., Riebel, K., & Brumm, H. (2010). Female zebra finches prefer highamplitude song. *Animal Behaviour*, *79*(4), 877–883. http://doi.org/10.1016/j.anbehav.2009.12.038

Rohrmeier, M., Zuidema, W., Wiggins, G. A., & Scharff, C. (2015). Principles of structure building in music, language and animal song. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1664), 20140097-. http://doi.org/10.1098/rstb.2014.0097

Rosenbaum, H. C., Pomilla, C., Mendez, M., Leslie, M. S., Best, P. B., Findlay, K. P., … Kiszka, J. (2009). Population Structure of Humpback Whales from Their Breeding Grounds in the South Atlantic and Indian Oceans. *PLoS ONE*, *4*(10). http://doi.org/10.1371/journal.pone.0007318

Rowell, J. T., & M.R., S. (2012). Vocal Communications and the Maintenance of Population Specific Songs in a Contact Zone. *PLoS ONE*, *7*(5).

Schmitt, N. T., Double, M. C., Baker, S., Gales, N., Childerhouse, S., Polanowski, A. M., … Peakall, R. (2014). Mixed-stock analysis of humpback whales (Megaptera novaeangliae) on Antarctic feeding grounds. *Journal of Cetacean Research and Management*, *14*(1), 141–157.

Schusterman, R. J. (2008). Vocal learning in mammals with special emphasis on pinnipeds. In *The Evolution of Communicative Flexibility: Complexity, Creativity, and Adaptability in Human and Animal Communication* (pp. 41–70). http://doi.org/10.7551/mitpress/9780262151214.003.0003

Shannon, C. E. (2001). A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, *5*(1), 3. http://doi.org/10.1145/584091.584093

Shiffman, D. (2012). *The Nature of Code*. *The Nature of Code*.

Slater, P. J. (1986). The cultural transmission of bird song. *Trends in Ecology & Evolution*, *1*(4), 94–7. http://doi.org/10.1016/0169-5347(86)90032-7

Smith, Joshua N, Goldizen, Anne W., Dunlop, Rebecca A., Noad, M. J. (2008). Songs of male humpback whales , Megaptera novaeangliae , are involved in intersexual interactions. *Animal Behaviour*, 467–477. http://doi.org/10.1016/j.anbehav.2008.02.013

Smith, J. N., Goldizen, A. W., Dunlop, R. a., & Noad, M. J. (2008). Songs of male humpback whales, Megaptera novaeangliae, are involved in intersexual interactions. *Animal Behaviour*, *76*(2), 467–477.

http://doi.org/10.1016/j.anbehav.2008.02.013

Steels, L. (2015). *The Talking Heads Experiment*. Retrieved from http://langsci-press.org/catalog/book/49

Stowell, D., Gill, L., & Clayton, D. (2016). Detailed temporal structure of communication networks in groups of songbirds. *Journal of Royal Society Interface*, *13*, 1–34. http://doi.org/10.1098/rsif.2016.0296

Sturm, B. L. (2014). A simple method to determine if a music information retrieval system is a "horse." *IEEE Transactions on Multimedia*, *16*(6), 1636–1644. http://doi.org/10.1109/TMM.2014.2330697

Suge, R., & Okanoya, K. (2010). Perceptual chunking in the self-produced songs of Bengalese wnches (Lonchura striata var. domestica). *Animal Cognition*, *13*(3), 515–523. http://doi.org/10.1007/s10071-009-0302-4

Suzuki, R., Buck, J. R., & Tyack, P. L. (2006). Information entropy of humpback whale songs. *The Journal of the Acoustical Society of America*, *119*(3), 1849–1866. http://doi.org/10.1121/1.2161827

Tallerman, M. (2013). Join the dots: A musical interlude in the evolution of language? *Journal of Linguistics*, *49*(2), 455–487. http://doi.org/10.1017/S0022226713000017

Tervo, O. (2011). Bowhead whale songs sung by females in Disko Bay, Greenland. *19th Biennial Conference on the Biology of Marine Mammals*. Retrieved from http://www.forskningsdatabasen.dk/en/catalog/2185989857

Tervo, O. M., Christoffersen, M. F., Parks, S. E., Kristensen, R. M., & Madsen, P. T. (2011). Evidence for simultaneous sound production in the bowhead whale (Balaena mysticetus). *The Journal of the Acoustical Society of America*, *130*(4), 2257–62. http://doi.org/10.1121/1.3628327

Tervo, O. M., Christoffersen, M. F., Simon, M., Miller, L. A., Jensen, F. H., Parks, S. E., & Madsen, P. T. (2012). High Source Levels and Small Active Space of High-Pitched Song in Bowhead Whales (Balaena mysticetus). *PLoS ONE*, *7*(12). http://doi.org/10.1371/journal.pone.0052072

Tervo, O. M., Parks, S. E., Christoffersen, M. F., Miller, L. A., & Kristensen, R. M. (2011). Annual changes in the winter song of bowhead whales (Balaena mysticetus) in Disko Bay, Western Greenland. *Marine Mammal Science*, *27*.

Tervo, O. M., Parks, S. E., & Miller, L. a. (2009). Seasonal changes in the vocal behavior of bowhead whales (Balaena mysticetus) in Disko Bay, Western-Greenland. *The Journal of the Acoustical Society of America*, *126*(3), 1570. http://doi.org/10.1121/1.3158941

Thorpe, W. H. (1958). THE LEARNING OF SONG PATTERNS BY BIRDS, WITH ESPECIAL REFERENCE TO THE SONG OF THE CHAFFINCH FRINGILLA COELEBS. *Ibis*, *100*(4), 535–570. http://doi.org/10.1111/j.1474-919X.1958.tb07960.x

Thorpe, W. H. (1961). *Bird-song; the biology of vocal communication and expression in birds*. *Cambridge monographs in experimental biology*.

Todd, P. M., & Werner, G. M. (1999). *Frankensteinian Methods for Evolutionary Music Composition*. (N. Griffith & P. M. Todd, Eds.)*Musical networks: Parallel distributed perception and performanc*. Cambridge, MA: MIT Press/Bradford Books.

Tuohy, D., & Ó hAodha, M. (2008). *Postcolonial Artist: Johnny Doran and Irish Traveller Tradition*. Cambridge Scholars Publishing.

Turing, A. (1936). On computable numbers. *Proceedings of the London Mathematical Society*, *42*, 230–265. http://doi.org/10.2307/2268810

Tyack, P. (1983). Differential response of humpback whales, *Megaptera novaeangliae*, to playback of song or social sounds. *Behavioral Ecology and Sociobiology*, *13*(1), 49–55. http://doi.org/10.1007/BF00295075

Tyack, P., & Whitehead, H. (1983). Male Competition in Large Groups of Wintering Humpback Whales. *Behaviour*, *83*(1), 132–154. http://doi.org/10.1163/156853982X00067

Vu, E. T., Risch, D., Clark, C. W., Gaylord, S., Hatch, L. T., Thompson, M. A., … Van Parijs, S. M. (2012). Humpback whale song occurs extensively on feeding grounds in the western North Atlantic Ocean. *Aquatic Biology*, *14*(2), 175–183. http://doi.org/10.3354/ab00390

Wade, J., & Arnold, A. P. (2004). Sexual differentiation of the zebra finch song system. In *Annals of the New York Academy of Sciences* (Vol. 1016, pp. 540–559). http://doi.org/10.1196/annals.1298.015

Watkins, W. A. (1977). Sperm whale codas. *The Journal of the Acoustical Society of America*, *62*(6), 1485. http://doi.org/10.1121/1.381678

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of “small-world” networks. *Nature*, *393*(6684), 440–2. http://doi.org/10.1038/30918

White, J. W., Rassweiler, A., Samhouri, J. F., Stier, A. C., & White, C. (2014). Ecologists should not use statistical significance tests to interpret simulation model results. *Oikos*, *123*(4), 385–388. http://doi.org/10.1111/j.1600-0706.2013.01073.x

Whitehead, H. (1998). Cultural selection and genetic diversity in matrilineal whales. *Science*, *282*(5394), 1708–1711. http://doi.org/10.1126/science.282.5394.1708

Whitehead, H. (2005). Genetic diversity in the matrilineal whales: Models of cultural hitchhiking and group-specific non-heritable demographic variation. *Marine Mammal Science*, *21*(1), 58–79. http://doi.org/10.1111/j.1748-7692.2005.tb01208.x

Whitehead, H., Richerson, P. J., & Boyd, R. (2002). Cultural Selection and Genetic Diversity in Humans. *Selection*, *3*, 115–125.

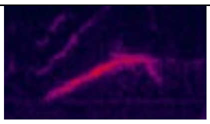
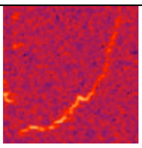
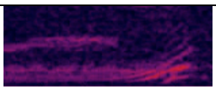
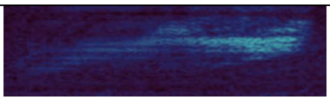
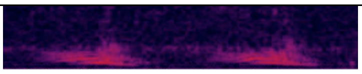
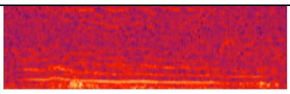
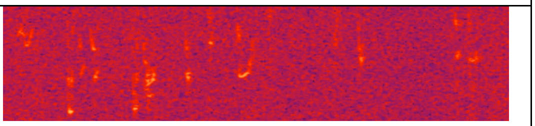
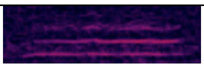
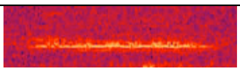
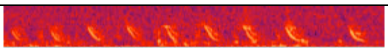
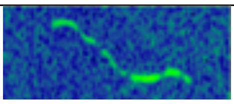
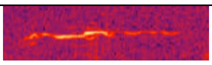
Whitehead, H., Waters, S., & Lyrholm, T. (1991). Social organization of female sperm whales and their offspring: constant companions and casual acquaintances. *Behavioral Ecology and Sociobiology*, *29*(5), 385–389. http://doi.org/10.1007/BF00165964

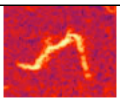
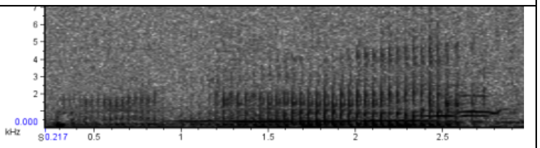
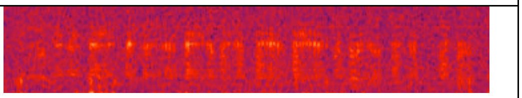
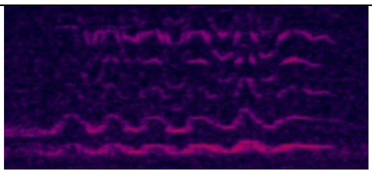
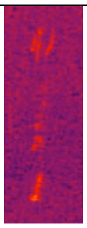
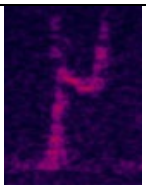
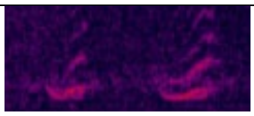
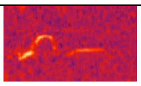
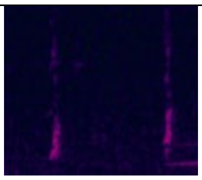
Whiten, A., Horner, V., & de Waal, F. B. M. (2005). Conformity to cultural norms of tool use in chimpanzees. *Nature*, *437*(7059), 737–740. http://doi.org/10.1038/nature04047

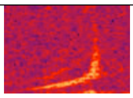
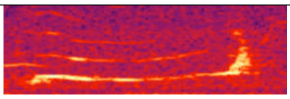
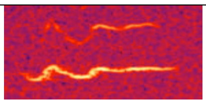
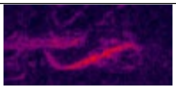
Williams, H., Kilander, K., & Sotanski, M. Lou. (1993). Untutored song, reproductive success and song learning. *Animal Behaviour*. http://doi.org/10.1006/anbe.1993.1084

Winn, H. E., & Winn, L. K. (1978). The song of the humpback whale Megaptera novaeangliae in the West Indies. *Mar. Biol.*, *47*, 97–114.

# Appendix 1 – Unit Naming Convention

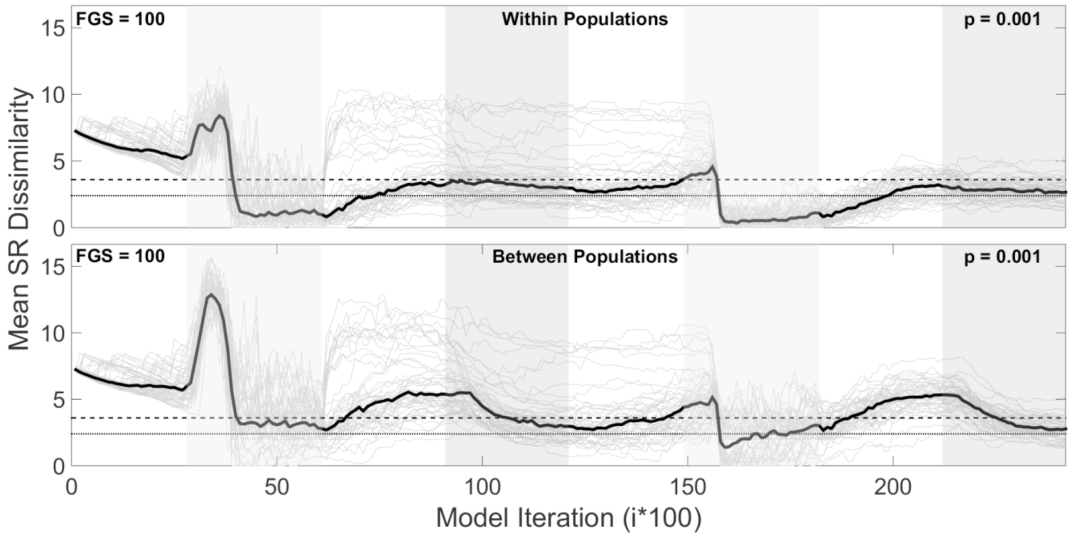
| Sound name | Sound ID | Picture |
|---|---|---|
| Ascending moan | am |  |
| Ascending shriek | ash |  |
| Bark | bk |  |
| Bellow | be |  |
| Croak | cr |  |
| Descending groan | dgr |  |
| Dolphin whistle | Dolphin whistle |  |
| Groan | gr |  |
| Growl | gr |  |
| Grunt | gt |  |
| High cry | Hc |  |
| Modulated moan | Mm |  |

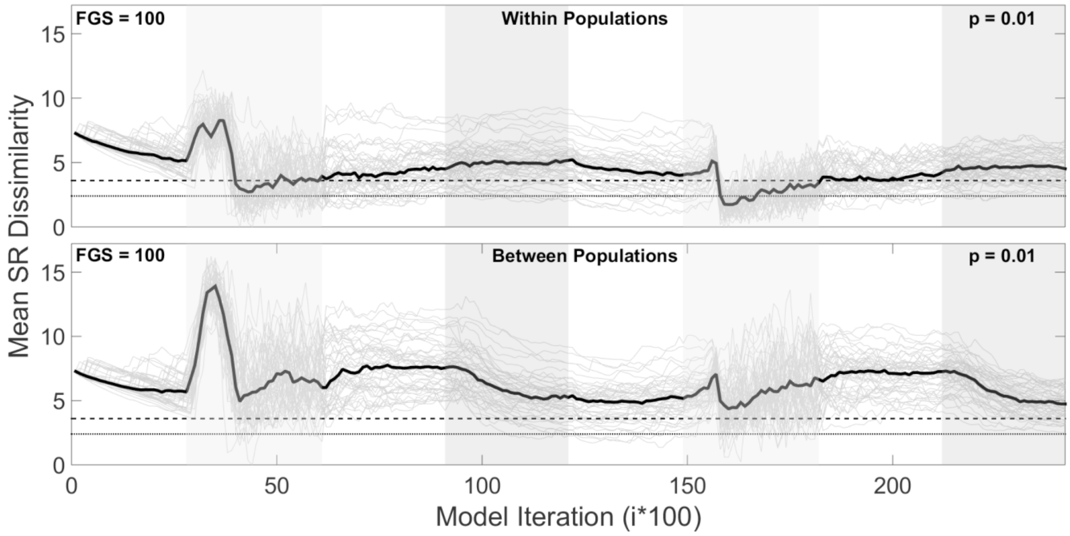
| | | |
|---|---|---|
| n-shaped high cry | Nhc |  |
| Purr | Pr |  |
| ratchet | ratchet |  |
| Siren | Si |  |
| Squeak | Sq |  |
| Trumpet | Tr |  |
| Violin | Vi |  |
| Wail | wail |  |
| Whoop | Wh |  |

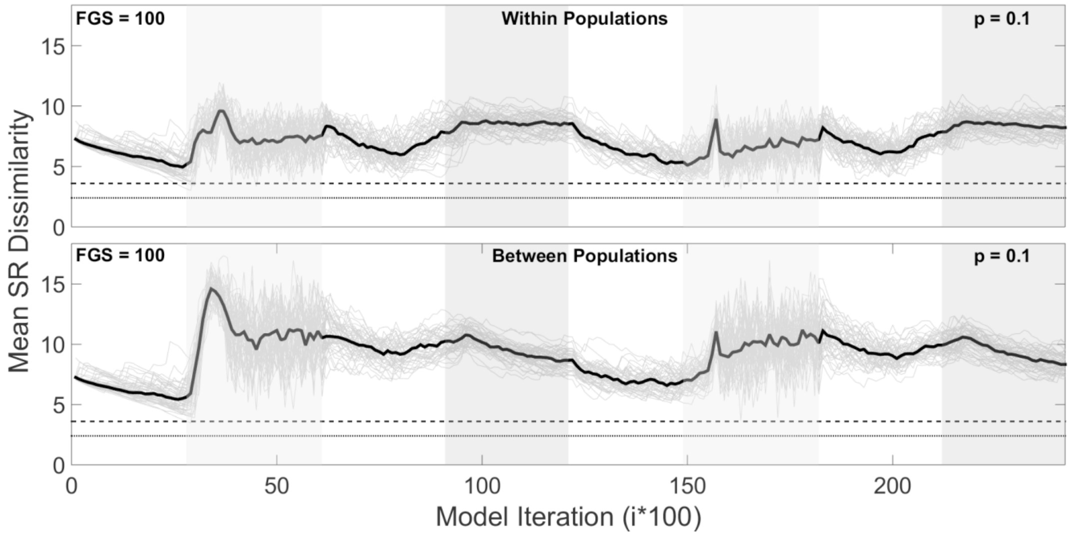
| | | |
|---|---|---|
| Wup | wp |  |
| Long wup | lwp |  |
| Wiggle moan | wm |  |
| u-shaped ascending moan | UAM |  |

# Appendix 2 – Chapter 2 Production Error and Production Error and Novelty Coupling Results (FGS 100 to 500)

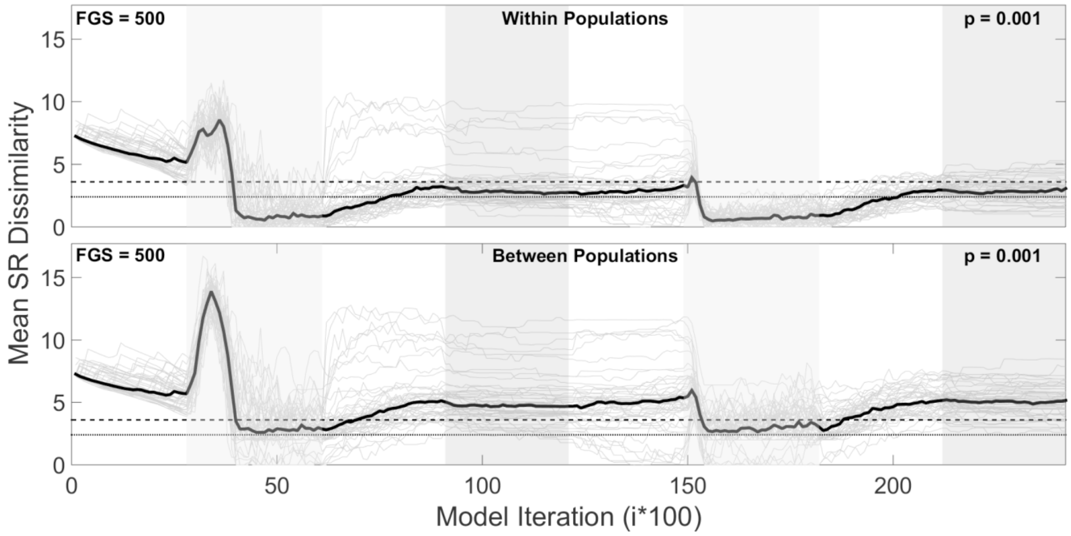
### Model 3, FGS 100

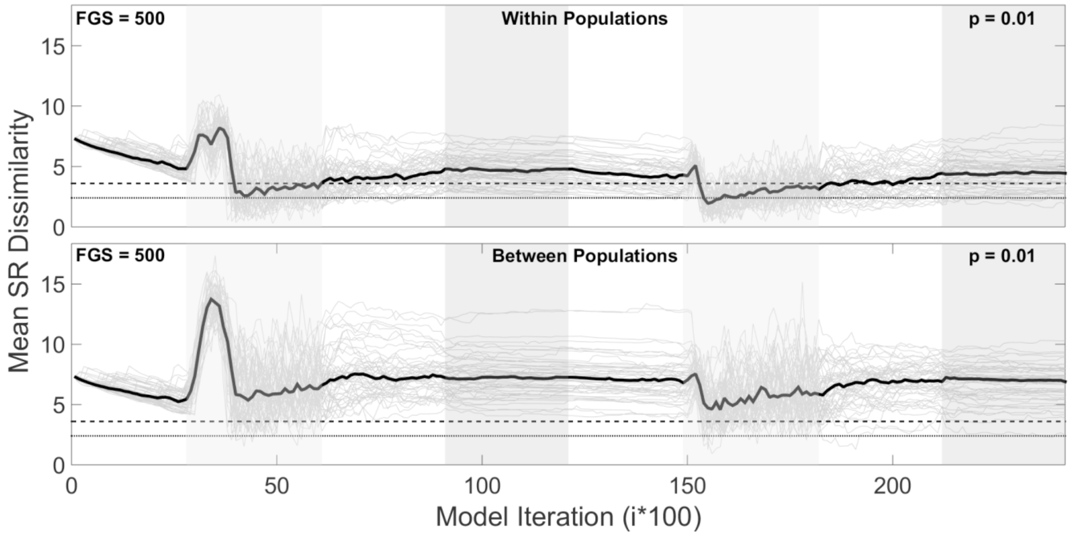


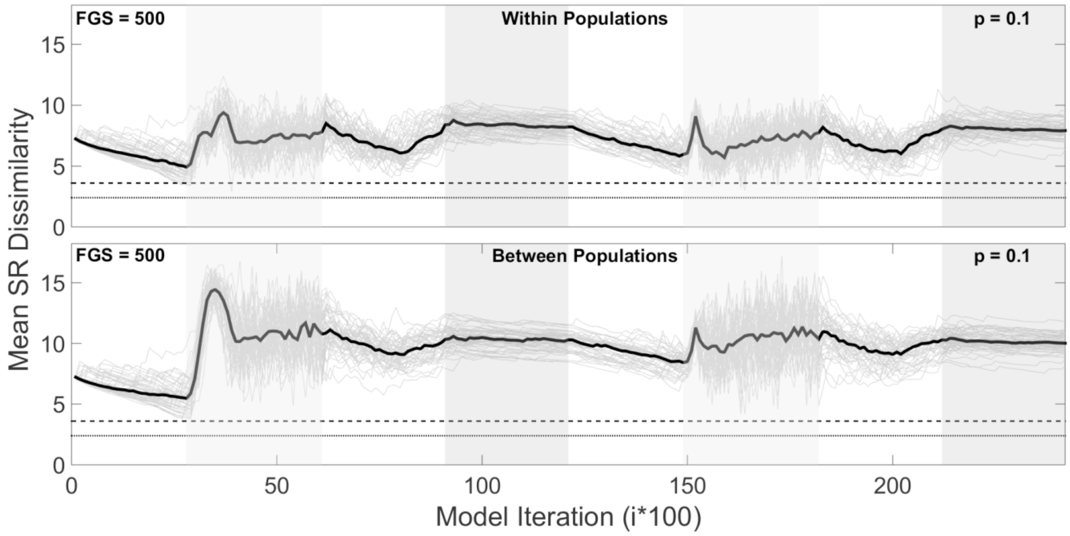




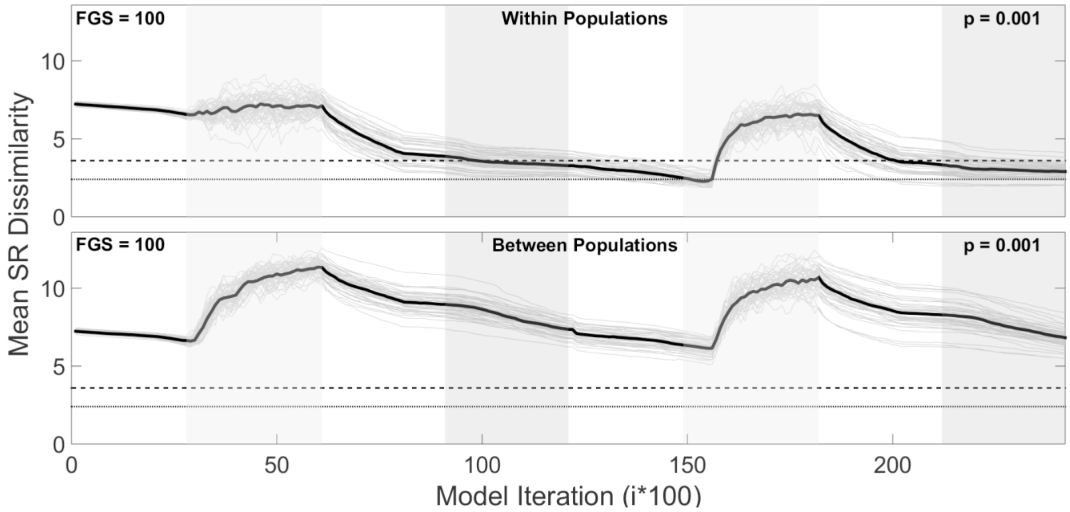
## Migratory Model 3 FGS 500

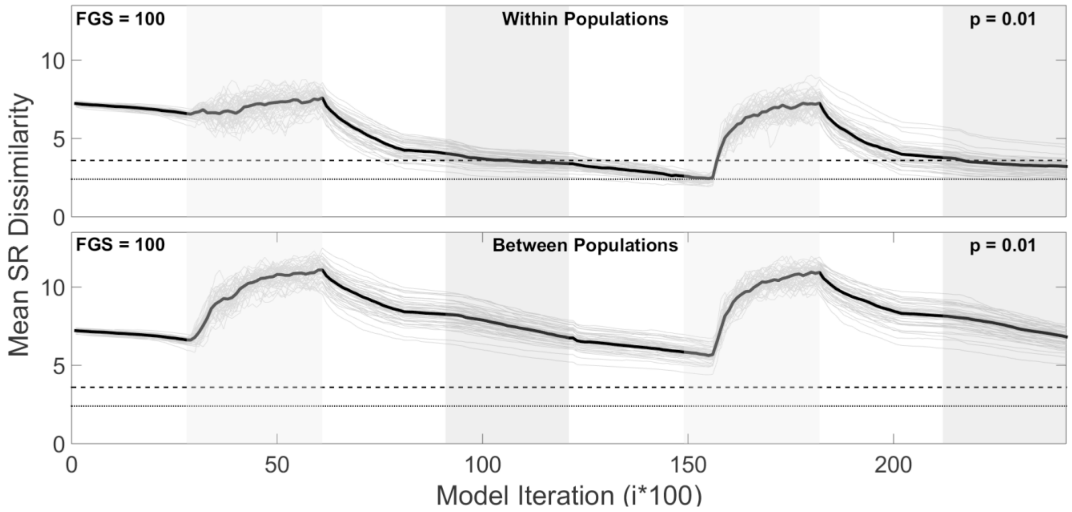


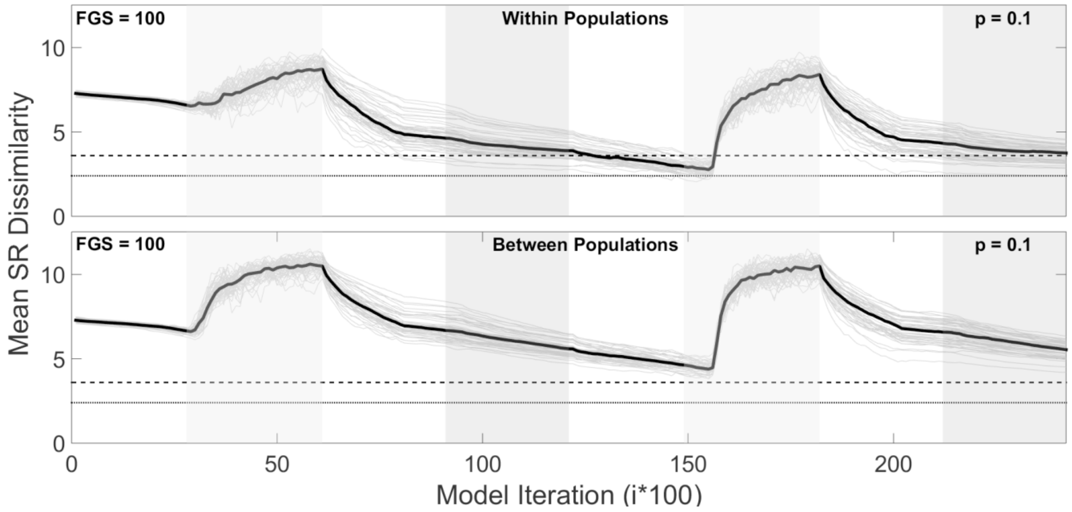




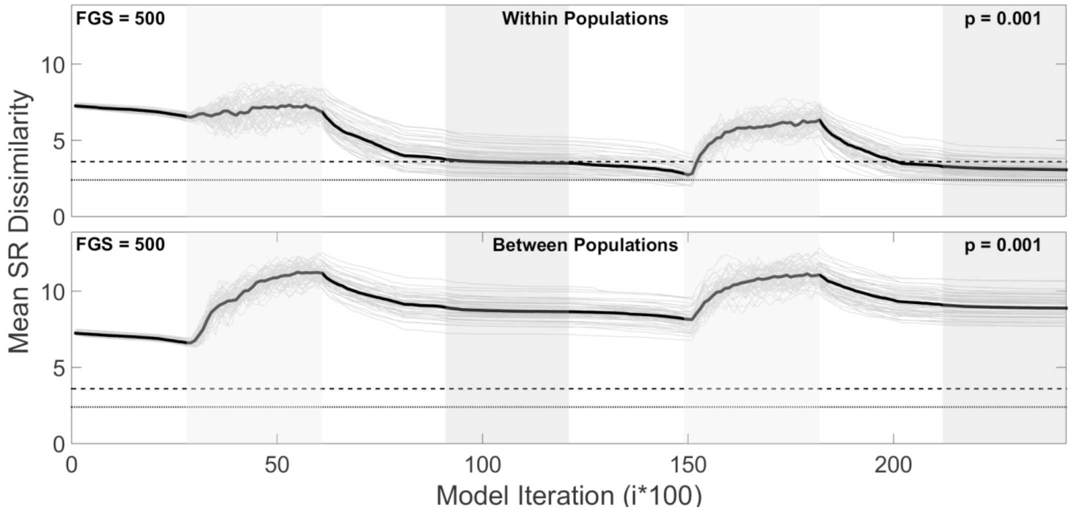
## Migratory Model 1.4, FGS 100

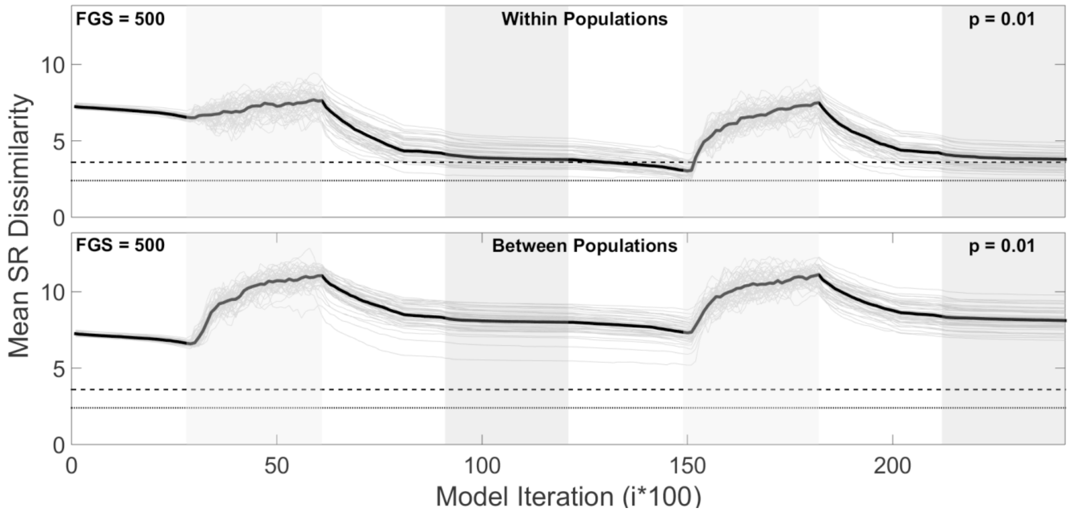


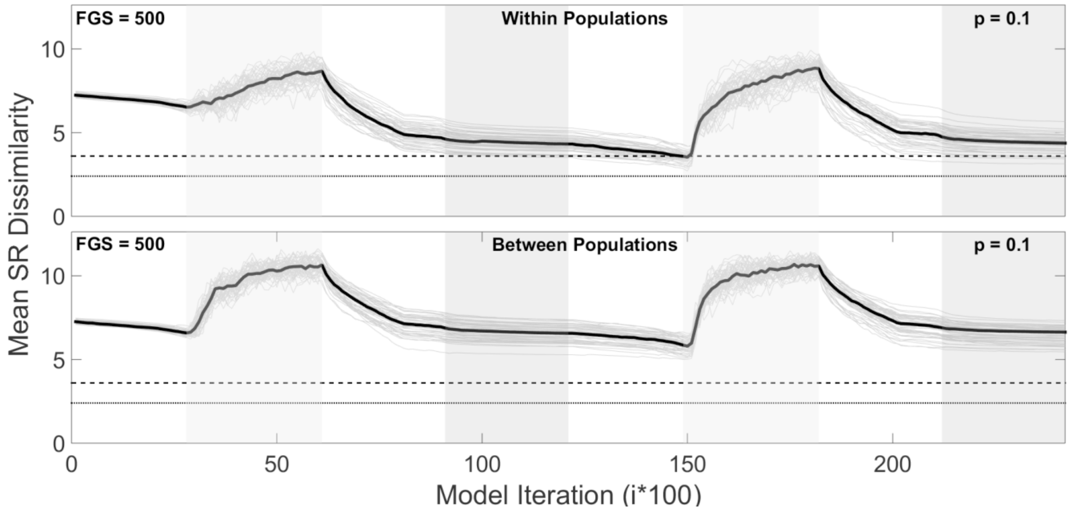




## Migratory Model 1.4 FGS 500




FGS = 500
Within Populations
p = 0.01
FGS = 500
Between Populations
p = 0.01
Mean SR Dissimilarity
Model Iteration (i*100)
0
5
10
0
50
100
150
200


FGS = 500
Within Populations
p = 0.1
FGS = 500
Between Populations
p = 0.1
Mean SR Dissimilarity
Model Iteration (i*100)
0
5
10
0
50
100
150
200

# Appendix 3 – Digital Supplementary Material

A CD with all agent based models and instructions on how to use them is provided with this thesis. The CD contains a folder with each model. In order to run the model, you will need Python Version 2.7.13 and the SciPy packages. The recommended distribution for running the scripts is Anaconda Ver 2.0.0. This is available at the link below:

https://repo.continuum.io/archive/

# Appendix 4 – Conference Proceedings

Attached at the end of this thesis are the conference proceedings where work from this thesis was published. This includes an abstract presented at the EvoLang conference, and a paper presented at the Sound and Music Computing Conference. The titles of the proceedings are printed below:

McLoughlin, M., Lamoni, L., Garland, E., Ingram, S., Noad, M., Rendell, L., Kirke, A., Miranda, E.R., (2016). "Adapting a Computational Multi Agent Model for Humpback Whale Song Research for use as a Tool for Algorithmic Composition." *In Proceedings of the Sound and Music Computing Conference*, 31 August - 3 September, Hamburg, Germany, 2016

Mcloughlin M., Lamoni L., Garland E., Ingram S., Kirke A., Noad M., Rendell L. and Miranda E. (2016). "Preliminary Results From A Computational Multi Agent Modelling Approach To Study Humpback Whale Song Cultural Transmission." In S.G. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér & T. Verhoef (Eds.) The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG11).